Cancer metastasis as a result of interactions between epithelial and mesenchymal gene programs

Metàstasis del càncer com a resultat d'interaccions entre programes gènics epitelials i mesenquimals

Antoni Celià Terrassa

UNIVERSITAT DE BARCELONA

FACULTAT
Farmàcia

DEPARTAMENT
Bioquímica i Biologia Molecular

Cancer metastasis as a result of interactions between epithelial and mesenchymal gene programs

---

Metàstasis del càncer com a resultat d'interaccions entre programes gènics epitelials i mesenquimals

Antoni Celià-Terrassa – 2012

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA

DEPARTAMENT
Bioquímica i Biologia Molecular

PROGRAMA DE DOCTORAT
BIOMEDICINA

Cancer metastasis as a result of interactions between epithelial and mesenchymal gene programs

Metàstasis del càncer com a resultat d'interaccions entre programes gènics epitelials i mesenquimals

Memòria presentada per Antoni Celià-Terrassa per optar al títol de doctor per la Universitat de Barcelona

Aquest treball ha estat realitzat en el departament de Biologia Cel·lular del Institut de Biologia Molecular de Barcelona (IBMB), Consell Superior d'Investigacions Científiques (CSIC), sota la direcció del Dr. Timothy M. Thomson

Director: Timothy M. Thomson Okatsu

Doctorand: Antoni Celià-Terrassa

Antoni Celià-Terrassa - 2012

Agraïments

Em considero molt afortunat per la gent amb qui m'he trobat i m'ha ajudat durant tots aquest anys... i que ha format part molt important d'aquesta experiència única en la vida. He d'agrair tantes coses a tanta gent que no sé quina extensió pot arribar a tenir aquest apartat.

Al primer que vull agrair i que sempre estaré indefinidament agraït és a en Tim. Vull ressaltar el teu compromís com a supervisor i director de tesi, ha estat espectacular, sempre has estat dispost a ajudar en qualsevol cosa, fins hi tot si calia corregir alguna cosa durant el cap de setmana així ho feies, sense dubtar i sense fallar. Això és un reflexa del que ets com a professional i com a persona. Tu vas confiar amb mi per començar un projecte que des del primer moment em va entusiasmar, moltes gràcies per aquesta oportunitat tant important per a mi. A més, he aprés moltíssim, almenys això crec... Gràcies perquè, a més de supervisar, ensenyes!... capaç d'ensenyar ciència i com interpretar-la, i moltíssimes altres coses... Moltes gràcies Tim, sempre t'admiraré i estaré molt agraït!

Vull agrair als col·laboradors que han confiat en el nostra projecte i han dedicat temps i esforç per anar endavant. Dintre d'aquest grup vull començar per la Dra. Nuria Rubio i Dr. Jerónimo Blanco de l'Hospital Sant Pau. Moltíssimes gràcies pel vostra enorme esforç, tant per iniciar com per ajudar fins el final d'aquest estudi, la vostra laboriosa ajuda ha estat essencial. I gràcies per la vostra simpatia que sempre heu tingut amb mi. També vull agrair al Dr. Pedro Fernández de l'Hospital Clínic pel gran esforç fet amb la part clínica i les divertides descripcions al microscopi p.e. les metafases com "bigotillos de Aznar". Al Dr. Juanjo Lozano, gràcies per la teva increïble rapidesa alhora d'analitzar dades i respondre les meves dubtes. A la Dra. Rosanna Paciucci de l'Hospital Vall d'Hebron per la teva simpatia i sempre estar disposta ajudar i col·laborar. Al Dr. Yibin Kang de Princeton per acollir-m'he al seu laboratori durant l'estada a l'estranger com a període formatiu de la tesi i voler implicar-se i ajudar en el nostra treball. Dr. Roger Gomis, Dr. Antonio Garcia de Herreros, Dra. Marian Martínez-Balbás, Dr. Ramon Vilella, gràcies per participar i ajudar en tot el possible.

També vull agrair a altres directors d'altres laboratoris on he estat abans de començar el doctorat i que foren claus pels meus inicis en la investigació. Primer de tot, a la Dra. Alicia Garcia-Arroyo del CNIC, per acollir-m'he al seu laboratori a Madrid com a estudiant de pràctiques, just abans d'acabar la carrera. Tu vas ser qui em vas donar l'impuls final que necessitava per decidir-m'he a fer el doctorat, moltes gràcies Alicia per omplir-m'he de confiança. També gràcies a la Dra. Paula Oliver, professora de la UIB que em va donar l'oportunitat de col·laborar en el seu laboratori, just quan estava al quart any de carrera, va ser el meu primer pas en el món de la recerca científica. Gràcies Paula. I també vull agrair al Dr. Damià Heine, professor de la UIB que em va aconsellar i ajudar per trobar un bon lloc on fer el doctorat.

I ara els companys de laboratori... amb els qui he compartit tantes hores i hores. Vull començar amb la primera persona que em vaig trobar al entrar al laboratori un 1 de març del 2007, la Marta Soler, moltíssimes gràcies Marta per tot el que em vas ensenyar. A la teva tesi deies que esperaves haver ensenyat alguna cosa durant els teus anys de doctorat, ara et puc afirmar que sí ho vares fer, a part d'aprendre tècniques també em vares ensenyar a ser prudent, cosa molt important per aquesta feina.

Óscar,... ara ja podem dir que fa molts anys que ens coneixem i hem passat per diferents moments. Des de fa molt que ens veiem diàriament, vàrem estudiar junts i em fet el doctorat junts, són dos èpoques úniques les que hem compartit i suposo que amb el temps encara lis donarem més valor. La veritat que em costa pensar que ara podem arribar estar llargues temporades sense veure'ns però crec que tard o prest el "equipo ganador" retornarà, tenim sa nostra estimada Mallorca. Vull destacar el teu gran sacrifici crucial pel projecte, gràcies per sempre està dispost a ajudar, la teva aportació ha estat clau. Molt content d'haver compartit grans moments, congressos, xerrades, viatges, festes,... Estic convençut que t'anirà molt bé perquè independentment de la sort, tu ho faràs. Mils de gràcies per tot!!

Marta Guerra, vaja un personatge agradable...que sàpigues que sempre seràs la nostra "capi". La veritat que, com diríem en Mallorquí, ets mel!...és a dir, de puta mare. Moltes gràcies per ensenyar-m'he tantes coses, pel teu bon humor, alegria, paciència... el ball de la palmera, etc etc... Raquel, enyor les teves sessions de música davant l'ordinador...era mel! A part d'això, dir-t'he que vaig aprendre moltíssim amb tu, sobre tot de la biologia del RNA... ets molt bona i m'encanta que t'encanti la teva feina. Moltes gràcies per ensenyar-m'he els passos a seguir al llarg del doctorat, pels teus consells i valuosa ajuda. Marta Isasa, pràcticament vàrem "nàixer" junts en això del doctorat, tesis, PhD... lo que sigui... quantes coses que hem passat des dels inicis eh!!, recordo molts bons moments, divertits i anecdòtics... sempre quedàvem fins les tantes i remugant... però insistint i insistint. Ha estat una gran sort recórrer en paral·lel aquest camí, moltes gràcies per ser tu mateixa i et desitjo el millor. Paki, moltes gràcies pel teu esforç i l'ajuda amb el paper, m'encanta la teva manera de treballar, diria que ets de les persones més eficients que he vist mai en un laboratori, estic molt content d'haver treballat amb tu i en poc temps m'has ensenyat moltes coses, tant de biologia com de maneres d'entendre aquesta feina, moltes gràcies. Lourdes, moltes gràcies per la teva ajuda quan vares començar fent pràctiques i sempre, encara que tu no ho creguis ets una de les persones que aprèn les coses amb més rapidesa i de forma més fiable, esper que en aquest tram final del teu doctorat assoleixis tot el que vols i que et vagi molt bé, gràcies. Bernat, ha estat un gran plaer conèixer-t'he, compartir laboratori durant aquests anys y a més poder jugar a futbol junts, gràcies per ser tant proper i agradable, ets un exemple a seguir tant com a científic com a persona, moltes gràcies per tot.

Johi, vares ser la primera persona que vaig veure fer una tesi doctoral, recordo els primers dies que anava als teus assajos i no m'enterava de gaire cosa, però em vas ensenyar a mostrar caràcter i ser forts, gràcies. Fabio*, tu llegaste casi al principio de mi tesis, y la verdad que nos lo pasamos muy bién, mil gracias por ser tan*

buena persona. Alexia*, muchas gracias por tu gran ayuda y implicación en el proyecto, me lo pasé muy bién contigo y estoy contento de haber compartido laboratorio. A parte de ser tan despistadilla, eres muy buena y debes aprovecharlo, mucha suerte.* Txell, quin gran fitxage que va fer el Tim, has estat una de les millor estudiants, ja sigui per la teva capacitat per treballar com per la persona que ets, sent una "adolescent" vares donar energia i vida al laboratori en molt poc temps,... recorda les tardes al bar triplet, moltes gràcies. Elijah, quin gran plaer trobar persones tant bones com tu, ets un crack! esper que tinguem l'oportunitat de veure'ns més sovint, gràcies per tot. Kristin, *eres una de las personas más espléndidas que he conocido, por tu forma de ser, tu valentia y sobretodo tu alegria... auténtica y fantástica, muchas gracias*. Simón, *gracias por tu ayuda y por siempre intentar innovar y admirar la ciéncia... Great "subsaharian dinners"*. Aida, una gran estudiant, intel·ligent, pacient, feinera i alegre... tens totes les qualitats perquè et vagi bé... moltes gràcies per la teva ajuda. A tots els estudiants que m'han ajudat durant aquests anys, i tècnics que hem tingut al laboratori, moltes gràcies a tots per la vostra ajuda: Maria Penzo, Cathy Ulloa, Magda Sekowska, Dani Álvarez, Ujué Fresan, Mark, Álvaro, Laura y Jaume Vaqué.

Lluís, quina gran sort que vos canviéssiu del CRG per venir al Parc Científic, la veritat és que ets un tio genial. Tot i que no fa molts anys que ens coneixem, tinc la sensació de que ens coneixem molt bé... som bastant semblants en moltes coses. Amb tu i en Jean Michel hem fet moltes coses... *si si mouchou que grande eres! sin duda eres la persona con más fuerza de voluntad y superación que he conocido nunca, acabo de mirar tu clasificación en la 27 ed. del Marathon des Sables 2012, posición 44 de 853 participantes; tiempo 29h10min en pleno desierto del Sahara, una auténtica proeza qué aquí también queda registrada, eres un super-man...aunque un poco "pedale". La verdad que con los dos, Lluis i JM, son muchas cosas: des de compartir cabina en cultivos hasta los partidos de futbol, Ricard parties, Pub tours, ir a correr con por la montaña, previas de champions,... mil cosas que han marcado estos años en Barcelona, muchas muchas gracias, sin vosotros dos esto no habría sido lo mismo. Visca Ricard!!*

Alice, hem compartit laboratori durant un any i ha estat un plaer, ets una gran persona i estic content d'haver-t'he conegut. Pilar, vaja una que estàs feta, ets tremenda, tens una personalitat extraordinari i una picardia molt divertida, m'ho he passat molt bé, esper que tot et vagui bé i que no canviïs, gràcies. Vero, *mucha fuerza... gracias por ayudar y hechar una mano siempre*. Conchi, ha estat una sort que després d'estudiar junts, hem fet la tesi en laboratoris veïns... o no, jeje! Gràcies per ensenyar-m'he a fer ChIPs i sobretot per aguantar les meves constants preguntes... tot i transformar-t'he en "super-guerrer". Treballa molt a San Diego que hem de tornar a Mallorca per muntar un centre d'investigació, jeje.

Al equipo de futbol, "All Blacks", "All Greys" y actualmente "All Organe", a todos los que hemos jugado juntos os agradezco mucho por haber compartido y sobretodo disfrutado de buenos momentos como equipo, espero que os hayais divertido tanto como yo. Gracias a todos: Marc, Ricky Lluís, JM, Albert, Bernat, Jordi,

Aleix, Tiago, Gwen, Xavi, Ivan, Márcia, Maria, Simón, Héctor,... y en especial al capitán Sergi por tomar la iniciativa de formar este equipo.

Gràcies al Carles Barceló, gràcies per posar-m'he en contacte amb el Tim, si no hagués estat per tu, no sé on seria ara mateix, t'agreixo molt que pensessis amb mi el dia que em vas telefonar per dir-m'he que en Tim cercava algú, moltes gràcies per això i per compartir altres moments sempre interessants. Gràcies a les tècnics de l'Hospital Clínic, Mónica y Laura, moltes gràcies per la vostra bona disposició a ajudar. Mónica, sempre dinàmica i divertida...enyoraré els cafetons a la terrassa del Clínic... fes bondat, petita Pumuki. Al tècnic Didac Domínguez de l'Hospital Sant Pau, gràcies per l'ajuda amb els ratolins, gràcies també al Juli per cuidar dels ratolins en ple estiu. Gràcies al tècnic Marc Guiu per ajudar amb les injeccions intracardiàques i també a l'Anna Arnal. Gràcies a la gent de Vall d'Hebron del laboratori de la Rosanna: Lide, Neus i Marta, gràcies. Gràcies al Ricky per les llargues xerrades de qualsevol cosa que em tingut durant aquests anys, moltes d'elles sobre el Barça, jeje... ets l'hostia... unes *palomitas*?... moltes gràcies.

Gràcies a tota la gent de fora del lab que he conegut durant aquests anys, ha estat molta gent, que sapigueu que heu format part d'aquesta experiència i que m'heu fet passar grans moments. Als amics mallorquins de Barcelona, als amics dels equips de futbol, als amics catalans, a tots... gràcies a tots. Especialment a na Marga, amiga de tota la vida, gràcies per sempre estar al nostra costat, podríem dir que has estat la nostra família a Barcelona. També, moltes gràcies, a uns altres veïns de feina, uns nois que es diuen Messi, Xavi, Iniesta,... i al Barça; sembla ridícul donar-lis les gràcies però ells han donat un toc màgic a aquests 5 anys que he viscut a Barcelona i almenys quan els experiments no sortien el Barça sempre guanyava!!

També vull mencionar i agrair als meus íntims amics de Sa Pobla que sempre m'han motivat per seguir endavant, moltes gràcies. També vull agrair a na Francisca Crespí, mil gràcies pels apunts i altres ajudes durant la carrera... no sé que hagués fet sense tu. Agrair a en Biel Nadal Porquer, per ajudar-m'he amb el disseny de la portada d'aquesta tesi. També vull agrair als amics de la meva família perquè molts d'ells es com si fossin part la meva família per lo molt que m'han apreciat sempre. A la família de na Violeta, en especial als seus pares, en Julian i na Carmen, per escollir Sa Pobla com a lloc per viure, d'altre forma no hauria conegut a na Violeta. I a Sa Pobla, al caràcter "pobler" de ser gent feinera, és un poble fort del que em sent orgullós, és el que he vist de petit i he après, gràcies

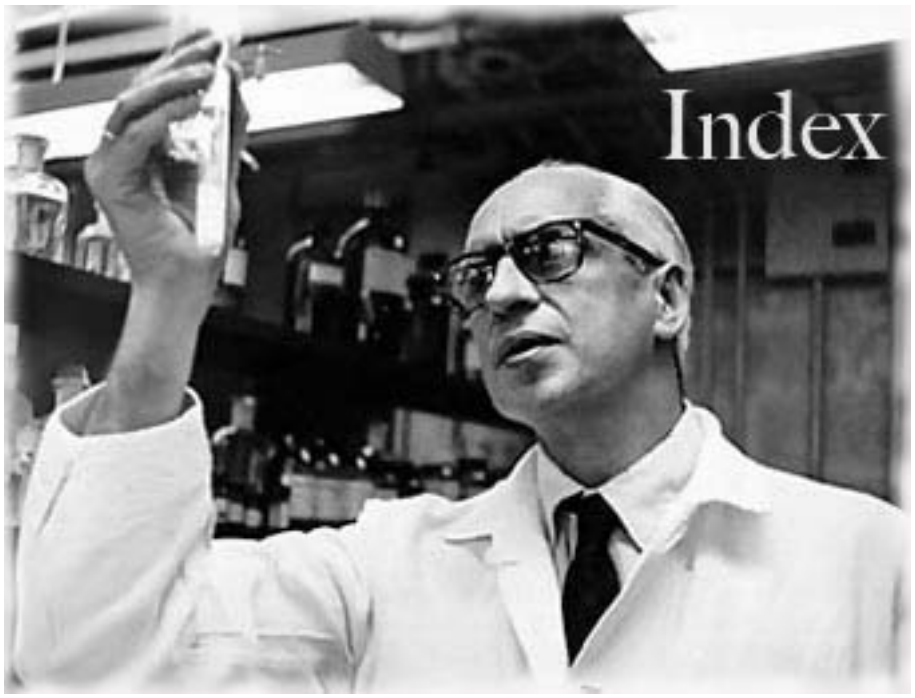
Lo més important, la meva família. Ells han estat els meus exemples a seguir i m'han procurat la millor infància que es pot tenir. Agrair als meus avis per part de mare, Catalina i Biel, només meu ensenyat bondat, és fantàstic pensar que lo únic que desitjàveu de mi és que fos una bona persona i per això sempre ho intentaré i vos tindré molt present. Als meus avis per part de pare, Joana i Toni, heu cuidat de mi de forma absoluta i increïble, esper poder fer-ho igual si un dia tenc nets, moltes gràcies per tantes alegries. Als meus padrins joves i altres tiets, gràcies per ajudar-m'he i estimar-m'he. Al meu cosí Biel pels grans moments quan érem petits i pels grans moments quan hem estat grans, sempre units. Al meu fillol, Joan Antoni

gràcies per estimar-m'he i aguantar les meves xarles. Al meu germà Biel, per haver-nos criats junts, ens em fet un a l'altre... estem lligats i he aprés moltes coses de tu, moltes gràcies. Als meus pares, Joan i Bel, ells són els màxims responsables de la persona que soc i de que hagi pogut fer el que he fet al llarg de sa meva vida. Ja he dit que no hi ha infància més feliç que sa meva, això ja ho diu tot de vosaltres dos, sou els millors pares del món, vos estim! Moltes gràcies per tot el que meu donat i el que m'esteu aportant contínuament.

Finalment, dedicar les ultimes línies a la persona més important durant aquests anys, a la meva parella, na Violeta. Tu has estat clau en tot això, tu ets qui em vas animar a sortir de Mallorca, tu ets qui m'has aguantat i m'has donat alegria en els moments més difícils. Si aquests 5 anys a Barcelona han estat tan especials i fantàstics, és perquè els hem compartit junts, era un projecte dels dos, ara ja consolidat i en queden de nous per venir que gràcies a tu em veig en forces d'afrontar. Gràcies per ser tu, per ser com ets i estimar-m'he. Aquesta tesi és tant cosa teva com meva, per això va dedicada especialment a tu!

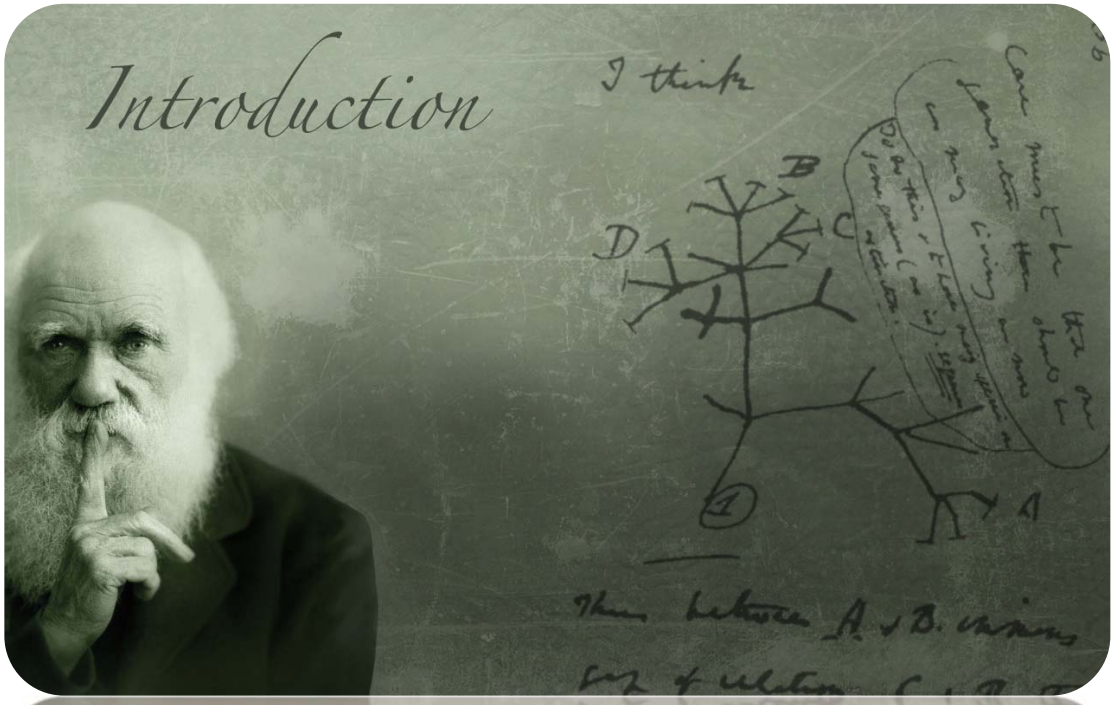
als meus pares
al meu germà

a na Violeta



Index

Severo Ochoa (Luarca, 1905 - Madrid, 1993)


Introduction
I think
B
C
D
A

Charles Darwin (Shrewsbury, 1809 - Kent, 1882)

1. Introduction to cancer progression

The progression of cancer into a disease is constituted by different biological capabilities by tumor cells. In 2000, Doug Hanahan and Robert Weinberg proposed six hallmarks of cancer [1]: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis and activating invasion and metastasis (Figure 1A). Recently, the two authors included two additional hallmarks, that have emerged from evidences accumulated over the last decade [2]: reprogramming of energy metabolism and evading immune destruction. Moreover, two enabling characteristics underlie and allow the acquisition of these hallmarks: genome instability, which generates genetic diversity, and inflammation, which foster multiple functional hallmark functions (Figure 1B). All of this needs to be considered in a context where tumor microenvironment plays a key role, tumors being a mixture of populations of neoplastic and non-neoplastic cells, the latter contributing to the acquisition of cancer hallmarks [2].

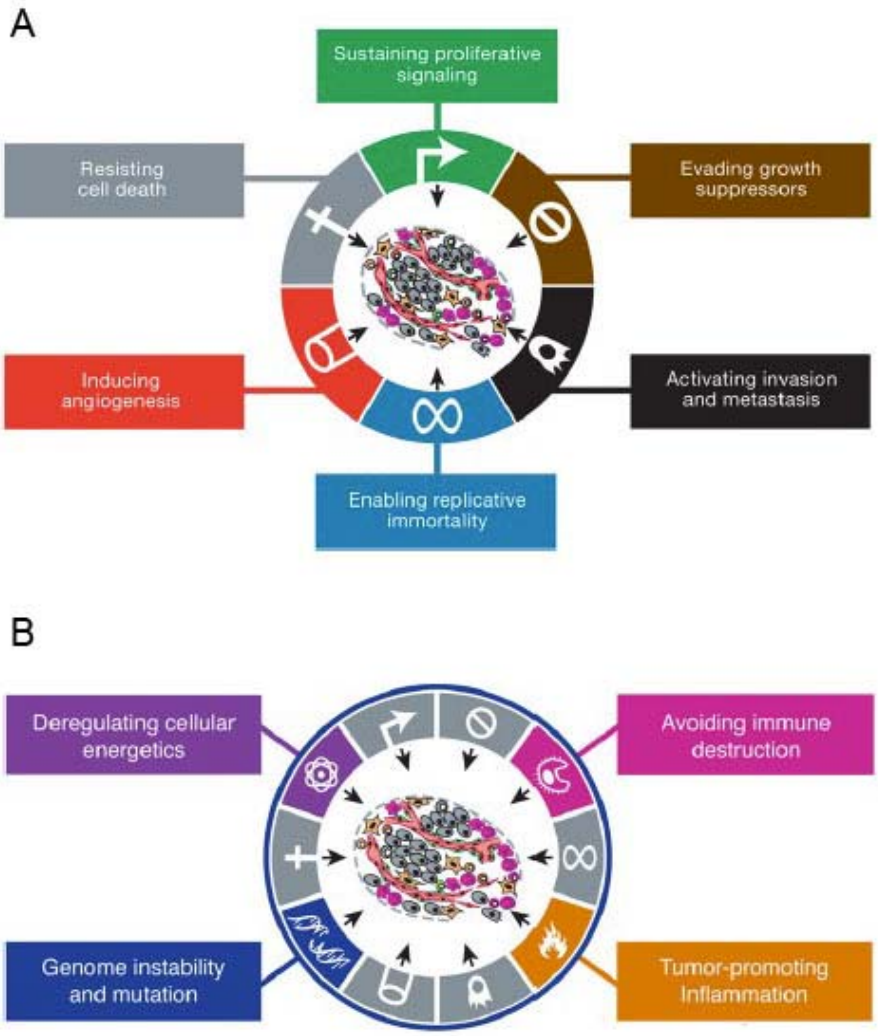


Figure 1. Hallmarks of Cancer. (A) Six biological capabilities for cancer progression established in 2000 by Hanahan and Weinberg. (B) New hallmarks for cancer progression, added as a consequence of its

biological importance reported by the new studies during the last decade. Hanahan and Weinberg 2011.

The major focus of this thesis dissertation is to study particular mechanisms of metastasis. **Metastasis** is the spread of cancer cells from primary tumor sites to distant organs and tissues and remains the cause of more than 90% of death in solid tumor cancer patients. The word *metastasis* means "displacement" in Greek, from μετά, *meta*, "next", and στάσις, *stasis*, "placement". In spite of its clinical importance, and recent critical advances from experimental models, the knowledge of the cellular and molecular mechanisms underlying cancer metastasis is still evolving. During decades cancer research has focused on the molecular bases of cancer, i.e., on what causes oncogenic transformation and the incipient emergence of tumors. However, progressively this shut perception has evolved and new frontiers became newly considered, given that tumors are more than just a mass of transformed cells and we now embrace the view that both neoplastic cell-autonomous and non-autonomous mechanisms determine the ultimate fate of tumors [3].

Molecular profiling of cancer using genomic-level approaches has revealed correlations among primary tumors and metastatic recurrence [4], leading to renewed questions on how, where and when cancer cells acquire genetic or epigenetic changes impacting their metastatic potential, and also reinforcing the hypothesis that cells with metastatic potential could exist in primary tumors, even at early stages of tumor development [5]. The evolutionary view of cancer progression, together with the numerous physical and physiological obstacles that normally prevent metastasis, entails a highly complex perspective of cancer progression. Millions of cells might be released into circulation every day from the primary tumor, but only a tiny minority of these cells will colonize a distant tissue. This implies that an evolutionary process with underlying selective pressures results in a selection of genetically or epigenetically heterogeneous lineages of cancer cells within the ecosystem of a given tumor [3]. In this way, cancer may progress as a disease of genetically heterogeneous cell populations with the potential to dynamically evolve by sequential environmental pressures. Although there are clear suggestions that diversity may be important in the establishment and progression of cancer [6], it is unclear how diversity should be defined. The problems of defining diversity and of characterizing the potentially many different neoplastic clones within any given tumor are not new and have parallels to problems encountered in other fields. Intratumoral heterogeneity has long been recognized, but has only recently been systematically investigated, and much can be gained from the application of both high-throughput and single-cell analytical approaches to the study of heterogeneity in cancerous and precancerous lesions [7].

2. Clonal origin of cancer, the Cancer Stem Cell hypothesis and origins of tumor heterogeneity: genetic and epigenetic aspects

There is a wealth of evidence that the acquisition of aggressive traits of cancer, or malignant progression, can be determined both by the occurrence of genetic mutations and by the imposition of heritable epigenetic marks on relevant genes [8]. Within a tumor, these newly acquired genetic and epigenetic events can emerge either sequentially within a single lineage, or in parallel in multiple, independent lineages [9]. In either scenario of cancer cell evolution, the final outcome is the coexistence in a given tumor of different subpopulations of tumor cells, each endowed with particular phenotypes (intratumoral heterogeneity). There is also evidence that transcriptional reprogramming in tumor cells can be induced in response to non-tumor environmental cues, that include factors such as TGF-β, PDGF or EGF [10], hormones, or hypoxic stress [11]. Therefore, cancer cells endowed with a capacity for indefinite self-renewal (cancer stem cells, CSCs), but still retaining some capacity for differentiation, could evolve into distinct phenotypes in response to environmental cues and to new mutations. It has been proposed that, like in any ecological niche [12], these subpopulations could interact among each other, either by competing for common resources [13], or by cooperating for mutual benefit [5] [10]. These tumoral subpopulations can also interact with, and use to their advantage, non-tumoral elements, as has been convincingly shown in many models of tumor progression and metastasis [15]. Those are important issues that deserve an extended comment later on.

Regarding the tissue of origin of human neoplasias, solid tumors account for the majority of the burden caused by cancers on the populations and epithelial cancers arising in tissues that include breast, lung, colon, prostate and ovary constitute approximately 80% of all cancers. However, the cellular origins of most solid tumors are unknown, although it has been speculated that different histological or molecular subtypes reflect distinct cells of origin at the time of tumor initiation.

The problem of complexity in cancer resides largely in the wide phenotypic and molecular heterogeneity observed even among tumors ascribed to the same histological type or subtype. Were it not for this huge variability, one might envision universal or near-universal therapies for a given tumor type, whereas the heterogeneous reality is prompting a recent flow of research towards personalized approaches. The gene-centric view has dominated the field of cancer research for decades, resulting in large efforts to define mutated oncogenes or tumor suppressor genes. The outcome has been that, despite the penetration shown by many such mutations, most of the identified alterations, with a handful of remarkable exceptions, have displayed low frequencies among patients. Cancer genome sequencing was proposed as a solution to the identification of high-frequency genetic alterations that might have escaped low-throughput detection approaches. Unfortunately, the results from several cancer genome sequencing projects have demonstrated once more that the vast majority of gene mutations found are not

shared among patients of a given tumor type [16]. With these results in hand, Heng *et al*. [16] have suggested that genome system replacement, as opposed to single-gene-centric mechanisms, drives cancer evolution, whereby cancer cells could follow unlimited numbers of genetic and epigenetic alternatives to progress towards malignancy. Therefore, the true challenge is to understand the behavior of the system (stability or instability) and the unpredictable replacement and switching that occurs between pathways during cancer progression, including the use of alternative pathways by cancer cells in response to medical intervention. In this highly dynamic process, there is a need to change our current approaches, by focusing more on monitoring the level of heterogeneity rather than attempting to identify specific patterns. There are strong arguments to indicate that strategies that attempt to reduce heterogeneity in order to study cancer mechanisms may represent flawed approaches. Without heterogeneity, there would be no cancer. That is the reason why many principles discovered using simplified homogenous experimental systems do not apply in the real world of heterogeneity. Thus, understanding the genome-centric concept of cancer evolution will help us to develop an experimental system to monitor and measure system dynamics. The genome-centric concept will guide us when applying genome level heterogeneity to the clinical challenges of cancer, as well as other common diseases [16].

Currently, two major models of tumor propagation have been proposed to account for heterogeneity and inherent differences in tumor-regenerating capacity: The “clonal evolution” model, and the more recently proposed “cancer stem cell” model. A third model, combining both models, may be more likely to provide a better explanation of the origin and evolution of cancer than either model in its own right.

2.1 The clonal evolution model

Nowell proposed the clonal evolution model in 1976 [17]. He observed that tumors generally seem to lose properties of differentiation as they progress and interpreted this removal of specialized functions as a way for them to maximize their proliferation and invasiveness.

This model of carcinogenesis states that cancer cells over time acquire various combinations of mutations within a tumor and that genetic drift and stepwise natural selection for the fittest yield the most aggressive cells and drive tumor progression (Figure 2). This model may, in some cases, involve a stochastic component, followed by genetic drift. According to this idea, tumor initiation takes place once multiple mutations occur in a random single cell, providing it with a selective growth advantage over adjacent normal cells. As the tumor progresses, genetic instability and uncontrolled proliferation allow the production of cells with additional mutations and hence new characteristics. These cells may produce a large number of offspring by chance, or the new mutations may provide a growth advantage over other tumor cells such as resistance to apoptosis. In either case, primarily the latter, new subpopulations of variant cells are born, and other subpopulations may contract, resulting in tumor heterogeneity [18].

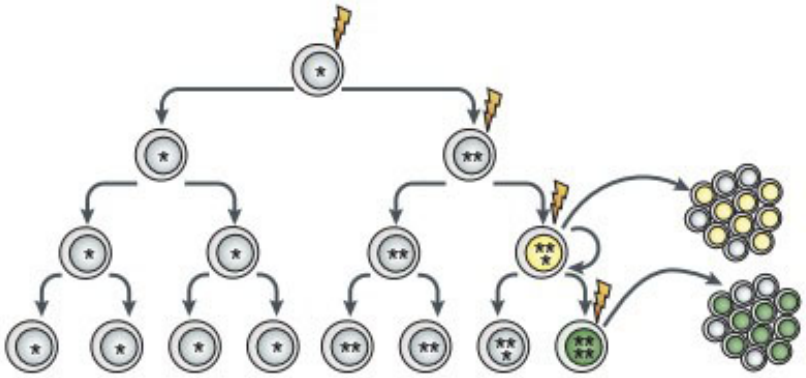


Figure 2. Clonal evolution model. The clonal evolution model postulates that mutant tumor cells with a growth advantage are selected and expanded, with cells in dominant populations having a similar potential for regenerating tumor growth. Asterisks denote number of mutation events accumulated. Modified from Visvader J.E. *et al.* 2008.

Perhaps the most convincing evidence for clonal evolution comes from studies of cancer cells taken from various regions within a single tumor using laser-microdissection. Such work reveals that every major type of human cancer and histological subtype contains at least several cell subpopulations with differing heritable abnormalities [12] [19]. The separation of these clones in distinct areas of a tumor occurs because sister cells usually remain contiguous in solid tumors. Examples of mutational heterogeneity that have been found are diploid and aneuploid clones within a tumor [20] and different allelic losses in cells from the same tumor [21]. Other evidences are: genomic instability enabling characteristics of cancer, unlimited proliferative capacity, various morphologies, metabolic differences, interplay with microenvironment, drug resistant clones after chemotherapy, or common genetic alteration between primary tumors, metastases and recurrences [18].

The acquisition of genetic events underpins this model but epigenetic differences (the main form of non-genetic heterogeneity) and microenvironmental changes are also likely to play important roles. Thus, genetic and epigenetic heterogeneity are key elements in cancer progression and drug resistance, as they provide the necessary population diversity, complexity, and robustness [16].

2.2 The Cancer Stem Cell hypothesis

Cancer stem cells (CSCs) are a subset of tumor cells with stem cell-like properties driving tumor initiation, progression and recurrence. CSCs share important properties with normal adult stem cells, including self-renewal (by symmetric and asymmetric division) and differentiation capacity, albeit aberrant. Their self-renewal and differentiation lead to the production of all cell types of a tumor, thereby generating tumor heterogeneity [22] [23]. At the same time, the other cells in a tumor do not have unlimited self-renewal capacity and cannot differentiate to produce all tumor cell types.

Importantly, CSCs may arise from malignant transformation of normal stem cells or from malignant restricted/differentiated progenitors that dedifferentiate and acquire self-renewing capacity. One implication of the CSC model is that cancers are hierarchically arranged, with CSCs lying at the apex of the hierarchy (Figure 3) [24]. The first evidence for the existence of CSCs came from acute myeloid leukemia [24], in which a rare subset comprising 0.01-1% of the total population could induce leukemia when transplanted into immunodeficient mice.

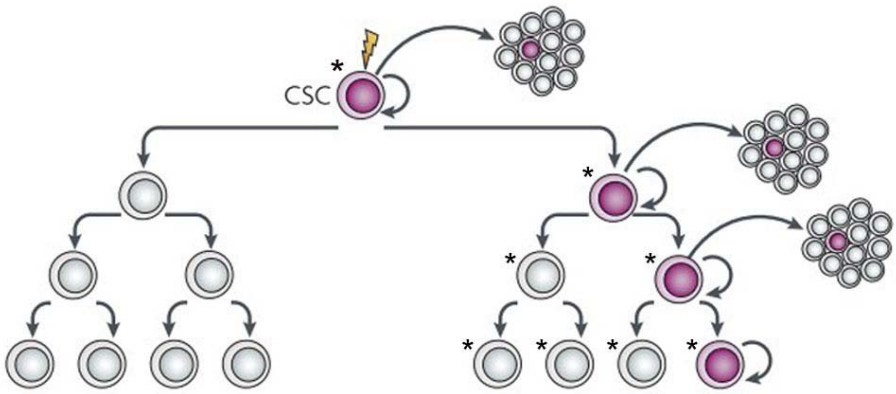


Figure 3. Cancer stem cell hierarchical model. Hierarchical organization of cells within the tumor, in which a subpopulation of stem-like cells is responsible for sustaining tumor growth. Asterisks denote number of mutation events accumulated. Modified from Visvader J.E. *et al.* 2008.

This is an attractive model to account for the functional heterogeneity that is commonly observed in solid tumors. Also, according to the cancer stem cell hypothesis, tumor progression is a result of the metastatic spread of these sort of cells, in this scenario called mCSCs [3] [27], and cancer recurrence is caused by their resistance to therapy [22] [26] [27]. In operational terms, the CSCs are, among all cancers cells, those able to initiate a xenotransplant tumor [28].

Interestingly, the notion of rare stem cells at the origin of cancer was already stated by Cohnheim in 1875. This concept is now supported by many observations, including the existence of normal stem cells in tissues where cancer often appears [29][30], that, being long-lived, are more prone to acquire mutations after different insulting events [31]. The self-renewal capacity and pluripotency of CSCs may explain tumor traits including monoclonality, unlimited proliferation and heterogeneity with a variety of differentiation states. Normal stem cells and cancer stem cells are known to be regulated by epigenetic mechanisms and microenvironmental influences. Moreover, they share other abilities which could explain metastasis and recurrence, including the induction of angiogenesis, apoptosis evasion, and drug resistance [31–34]. In fact, only a minor fraction of tumor cells are able to form a new tumor [32][35]. Important pathways in stem cell biology are involved in cancer, including Wnt, Hedgehog or Notch [26] [34]. In general, many key processes in cancer are reminiscent of embryonic developmental processes.

These two hypotheses of the origin and progression of cancer may be further refined in order to explain experimental evidences, as elaborated in the following sections.

2.3 Intersecting two models: Clonal evolution of Cancer Stem Cells

Although there are some differences in the basis of the above paradigms, they are not mutually exclusive, since CSCs themselves undergo clonal evolution, and thus, the two models can be coupled (Figure 4). Accordingly, both models postulate the origin of a tumor from a single cell that has acquired mutations and gained unlimited proliferative potential. In both cases, microenvironmental factors may influence tumor progression and furthermore in both cases the presence of stem-like properties would confer a selective growth advantage over the rest of the tumor cell population.

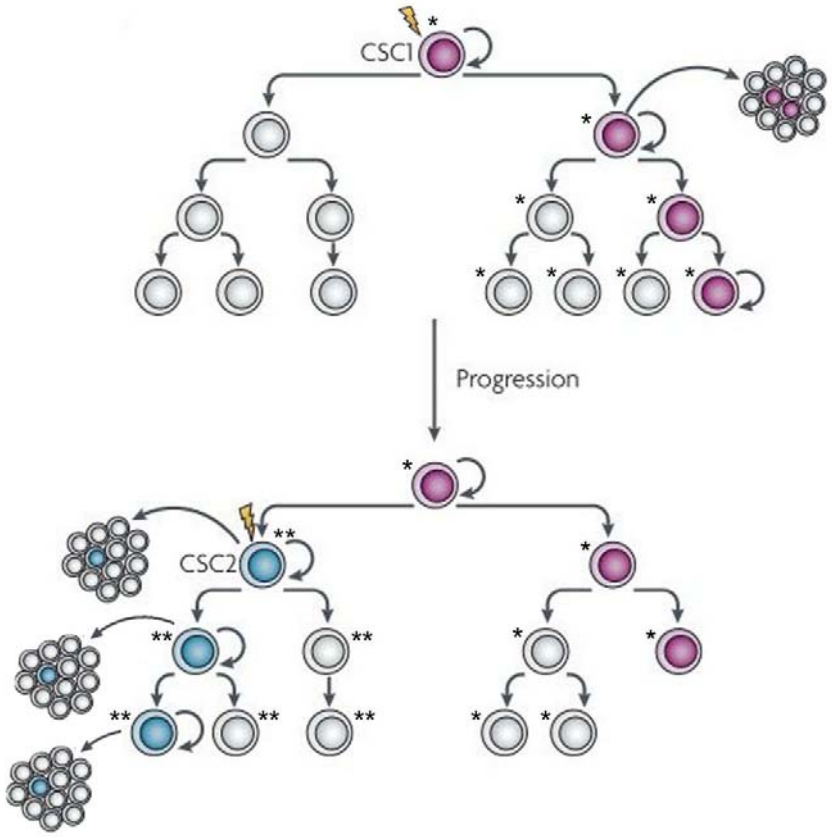


Figure 4. Cancer Stem Cell hierarchical model. Both models may underlie tumorigenesis. Initially, tumor growth is driven by CSC1, which may give rise to CSC2, due to clonal evolution of CSC1 as a result of mutations or epigenetic modifications. This more aggressive CSC2 becomes dominant and drives tumor progression towards malignancy. Asterisks denote number of mutation events accumulated. Modified from Visvader J.E. *et al.* 2008.

An example of this would be provided by breast cancer development, where $CD44^+$ normal mammary stem or progenitor cells acquire cancer-causing mutations. This results in $CD44^+$ cancer cells that can self-renew and differentiate into all tumor cell types, including $CD24^+$ cancer cells. Additional mutations may occur in any of these cells, leading to differentiation, dedifferentiation, or the acquisition of new characteristics. For instance, as indicated here, a $CD44^+$ cancer cell may mutate to become more differentiated and then again to acquire self-renewal ability and more stem cell-like traits. In addition, a $CD24^+$ cancer cell may become further mutated. Eventually, cell populations endowed with a growth advantage over other populations will dominate the tumor. The continued mutation and selection of cell

populations may lead to cancer cells that can self-renew and differentiate into all tumor cell types but do not express CD44 [18].

This variant model of the origin and evolution of cancer cells provides more versatility to the system and a better explanation of the emergence of a dynamic spectrum of tumor cells ranging from the most competitive to the disabled or doomed cells, with a positive selection of new CSC populations that acquire advantageous properties and that may be the main drivers of tumorigenesis. It is far from established, though, whether the latter cells are the only CSCs driving tumor growth and evolution or they do so together with the "earlier generation" CSCs. Some authors consider the second possibility and suggest an alternative stochastic model of tumor evolution.

2.4 Stochastic and reversible CSC models

Based on the previously stated hypothesis of CSC and clonal evolution, new models have been proposed that play down the importance of the hierarchical organization implicit in the cancer stem cell hypothesis, by which only the original malignant CSC is able to self-renew and give rise a to differentiated tumor cell progeny. The new models postulate an alternative stochastic and reversible mode of CSC evolution (Figure 5) [28]. Although these new models do not exclude a hierarchical organization, they consider it only as an extreme case of the stochastic model.

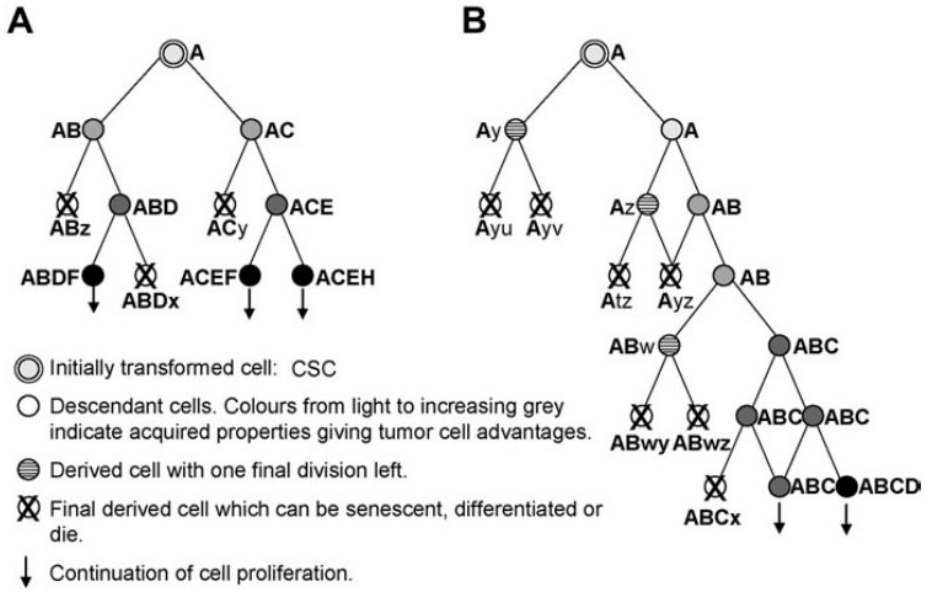


Figure 5. Two models for the generation of tumor heterogeneity. A) Alternative stochastic model in which successive functional, genetic or epigenetic events affecting the population of cells will lead to increasingly more diversified, more or less competitive cells. The scheme does not include the Darwinian selection forces that will drastically restrict the population diversity. B) The hierarchical model as generally and previously presented, with an evolving CSC population with some symmetrical divisions and many irreversible asymmetric divisions generating derived cells with limited life spans. ABCDEF represent the progressive addition of heritable changes that lead to increasing competitivity; wxyz are deleterious genetic or epigenetic events. Modified from Maenhaut *et al.* 2010.

Tumors are therefore probably constituted by a majority of less competitive, severely handicapped or doomed cells, and a minority of more competitive cells. The latter ones will to some extent behave as expected for CSCs. However, such cells would not constitute a unique homogenous but a heterogeneous population. They would be constituted of cells of the same origin having diverged and multiplied at different times in the evolution of the tumors and having survived a fierce competition. They would be stochastically and not deterministically selected [36] i.e. by a Darwinian selection [37], they could acquire some stem cell properties [38] and they would be quantitatively rather than qualitatively different from the derived populations [39]. There would be no fixed irreversible barrier between the two populations but rather a difficult but not impossible reversion to a more competitive state [40]. The state of these cells needs not to be permanent; cells could be more competitive at a particular site or a given time. Slow non-synchronized fluctuations that underlie the heterogeneity of even clonal populations [41] [42] could partly account for the diversity.

Different genetic or epigenetic events affecting different competitive cells at any stage could lead to new clones. While genetic events would only be reverted by other antagonistic events (e.g. an inactivation downstream from a constitutive activation), epigenetic events are reversible.

There are many arguments in favor of the concept of a dynamic equilibrium with some reversibility in transitions between cancer cell populations [43]. This model better fits a number of experimental findings, for example, the fact that more than 25% of melanoma cells are able to generate melanomas in severely immunocompromised mice and that this property has no relation to the expression of biomarkers [44]. Understanding how cancer cell states coexist and evolve within tumors is of fundamental interest and could facilitate the development of more effective therapies. Subpopulations of cells purified for a given phenotypic state return towards equilibrium proportions over time. These observations can be explained by a Markov model in which cells transition stochastically between states. A prediction of this model is that, given certain conditions, any subpopulation of cells will return to equilibrium phenotypic proportions over time. A second prediction is that breast cancer stem-like cells arise *de novo* from non-stem-like cells. These findings contribute to our understanding of cancer heterogeneity and reveal how stochasticity in single-cell behaviors promotes phenotypic equilibrium in populations of cancer cells [45] (Figure 6).

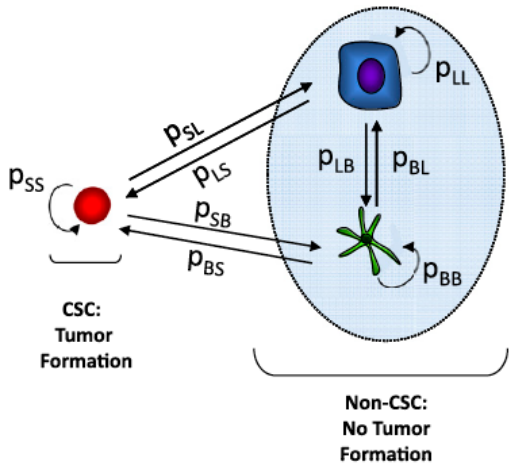


Figure 6. Alternative scenario in which there is a bidirectional interconversion between CSC and non-CSC states. The rates of transition between cell states, which vary across distinct cancer cell populations, can be computed with the Markov modeling approach. P = population; S = Stem-like phenotype; L = Luminal phenotype; B = Basal phenotype. Gupta *et al.* 2011.

3. Stemness, cancer, and cancer stem cells

Cancer cells that are proposed as the origin and the drivers of tumors have been variously designated cancer stem cells (CSCs), tumor initiating cells (TICs) and stemloids. Other authors propose that the operational term tumor-propagating cells (TPCs) is more suitable, based on the capacity of these cells to propagate *in vitro* and *in vivo*. However, CSCs being the most widely used term [45] [46], it will be the one generally used throughout this Introduction. The concept of cancer stem cells (CSCs) embodies intrinsically two aspects: the cancer stem cell as the initial target of the oncogenic process and the existence of two populations of cells in cancers: the CSCs and derived cells. It is important to mention that the name of CSC may generate confusion with mistaken interpretations regarding such cells as true stem cells, which they are not. The term CSC just indicates that these malignant cells share properties and gene programs with *bona fide* stem cells, but since they have been extensively transformed by genetic and epigenetic alterations, they are no longer stem cells and they do not have the capacity to give rise to wide varieties of differentiated progeny cells as embryonic stem cells do.

In some cases, cancer stem cells can arise from the mutational transformation of normal stem cells, whereas in other cases mutations might cause, in restricted progenitors or differentiated cells, the acquisition of properties of cancer stem cells, such as self-renewal potential. These pre-malignant stem cells would be the subject of genomic instability and clonal evolution, but they would be distinguished from other cancer cells by their tumorigenic potential, their ability to generate additional cancer stem cells (self-renewal) and their ability to generate phenotypically diverse non-tumorigenic cancer cells (with more limited proliferative potential) (Figure 7).

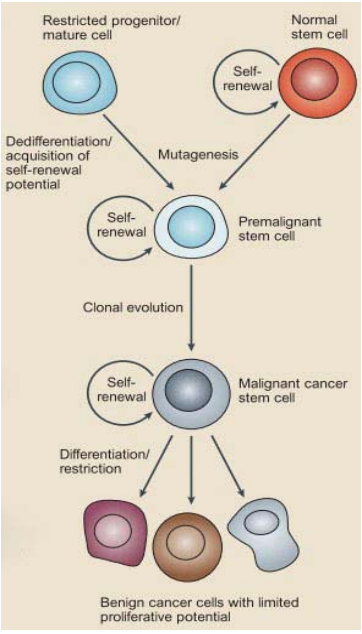


Figure 7. CSC origin, evolution and progeny. The mutational transformation of normal stem cells, or mutations in restricted progenitors or differentiated cells, might lead to the acquisition of properties of cancer stem cells, such as self-renewal potential. Genomic instability and clonal evolution may endow them with additional cancer stem cell attributes and permit them to generate non-tumorigenic cancer cells (with more limited proliferative potential). Modified from Pardal *et al.* 2003.

3.1. Embryonic and somatic stem cells

There are at least two general types of normal stem cells: the embryonic stem cells (ESCs), derived form the inner cell mass of the blastocyst-stage embryos, and the somatic stem cells (SSCs). Both are characterized by:

- Self-renewal and immortality (i.e. extended life span).

- Asymmetric divisions with irreversible sequential generation of a hierarchy of more differentiated descendants, with a limited life span, steadily reproducing the cell heterogeneity of tissues; this does not exclude some symmetrical divisions generating two stem cells.

- Homeostatic control, in which differentiation is a deterministic process. The fertilized oocyte can be considered as the initiator of all stem cells, the "primary stem cell". Stem cells of both types give rise to hierarchies of stem cells and derived pluripotential cells with a more restricted range of possible outcomes at each stage.

On the other hand, ESCs and SSCs differ in several other aspects. While ESCs are able to generate the three germ layers and ultimately all differentiated cells, SSCs have a much more restricted potential. Both types use different pathways for their maintenance. ESCs, which express telomerase, are immortal, whereas the exhaustion of the replicative capacity of adult somatic stem cells may contribute to aging. Importantly, while ESCs reproduce rapidly, the division rate of most SSCs is believed to be low although this is becoming controversial. To maintain their undifferentiated open state, SSCs require a very sophisticated but specific microenvironment with which they interact: the niche. The ESCs only need to be protected from differentiating agents. The characteristic and phenotype-directing transcription factors that are expressed in ESCs (OCT4, SOX2, NANOG) are not expressed in SSCs, each type of which expresses, or not, transcription factors specific of the stem cell status or of the cell identity-SSC that may exist in different tissues [28]. The distinction between ESCs and SSCs, and the general understanding of stem cell biology is currently under an intense focus of study, with many potential clinical applications. This and further discoveries will help to better understand the biology of CSCs because they too might be subject to reversible dynamics forces along the long path to metastasis.

First ESCs derivations, obtained from mouse, were achieved using a layer of mitotically inactivated fibroblasts (known as feeders) and a medium suplemented with serum. A crucial signal provided by feeders is the leukemia inhibitory factor (LIF) and serum can be replaced by bone morphogenic protein 4 (BMP4) [48]. This and others inhibitor or activators are recently showing up as technical procedures of improvement.

3.2. Transcriptional factor regulatory networks in ESCs

In embryonic stem cells (ESCs), the intricate interplay between transcription factors and their targets on the genomic template serves as building blocks for the transcriptional network that governs self-renewal and pluripotency. At the heart of this complex network is the transcription factor trio, OCT4, SOX2 and NANOG, which constitute the ESC transcriptional core, with feedforward and feedback loops of regulation [49]. Regulatory mechanisms such as autoregulatory and feedforward loops support the ESC transcriptional framework and serve as homeostatic control for ESC maintenance. In addition, genome-wide studies have further revealed additional players involved in pluripotency and the interconnectivity within the complex ESC transcriptional circuitry, in concert with epigenetic regulators that mantain the homeostasis of ESCs [49].

OCT4 or POU5F1, a key octamer transcription factor in ESCs, is downregulated upon differentiation and cells tend to lose their self-renewing state. It represses the differentiation specific-lineage gene CDX2. It is crucial for ESCs to mantain OCT4 at the appropiate levels in oder to mantain the pluripotency, increased or decreased levels may lead to differentiation [50].

SOX2, a SRY (Sex determining region-Y)-related transcription factor, a high mobility group box (HMGB) DNA-binding domain, can preserve ESCs stability, mantaining OCT4 expression at correct levels. In addition, OCT4 is a partner of heterodimerization with SOX2 with which it cooperate sinergistically for binding and regulation of many ESC-specific genes, including themselves. The cis-regulatory element to which SOX2-OCT4 complex is bound consists of neighbouring sox (5'-CATTGTA-3') and oct (5'-ATGCAAAT-3') elements [51].

NANOG is a homeodomain transcription factor and functions as a dimer. The loss of dimerization may compromise the preservation of sefl-renewal and pluripotency of ESCs. Moreover, dimerization of NANOG is crucial for its interaction with other plutipotency related proteins [52]. NANOG is essential for the establishment of pluripotency but is dispensable for its maintenance, provided when SOX2 and OCT4 are expressed.

Other factors relevant for ESC and pluripotency. KLF4 is a zing finger transcription factor with a defined role in mantaining self-renewal and thus contributing to the plutipotency and self-renewing framework in ESCs [53]. Also KLF2, KLF5 and KLF9 play roles in pluripotency [53]. RONIN, another zinc finger protein, can rescue the phenotype of OCT4 by repressing differentiation genes. RIF1, TCL1, TRIM28, CHD1 HDAC2 and others also interact and contribute to the plutipotent state [49].

Importantly, genome-scale binding analyses have revealed that OCT4, SOX2 and NANOG share a high degree of overlap in their binding to target genes. Moreover, they autoregulate themselves by binding to their own promoters forming autoregulatory loops [54]. This autoregulation can be either positive or negative in

nature over themselves or over other genes. This core regulatory mechanisms has a stabilizing effect in ESCs, keeping them poised for differentiation. KLF4 also plays important roles in this autoregulatory network because it is upstream of OCT4 and SOX2 with it shares common downstream targets, such NANOG, and also occupies the c-MYC promoter. MYC can support ESCs through functions distinct from those of the mentioned core genes, including positive cell proliferation regulation, negative regulation of differentiation, and regulation of chromosomal accesibility [55] (Figure 8 A). The co-occupancy of many different transcription factors of a gene promoters determines its state of activation or repression. Genes bound by more than four factors are generally transcriptionally active, whereas those bound by fewer transcription factors can be repressed [55]. The same study found that OCT4, SOX2, KLF4, NANOG, DAX1 and ZFP281 form a cluster that share many common genes targets while c-MYC and REX1 form a distinc cluster. The separate clustering of c-MYC from that of NANOG, OCT4 and SOX2 may have an additive effect required for the transcriptional activation of genes, whereby the MYC cluster binds to gene promoter regions and the core cluster to the enhancer regions forming enhanceosome complexes [56].

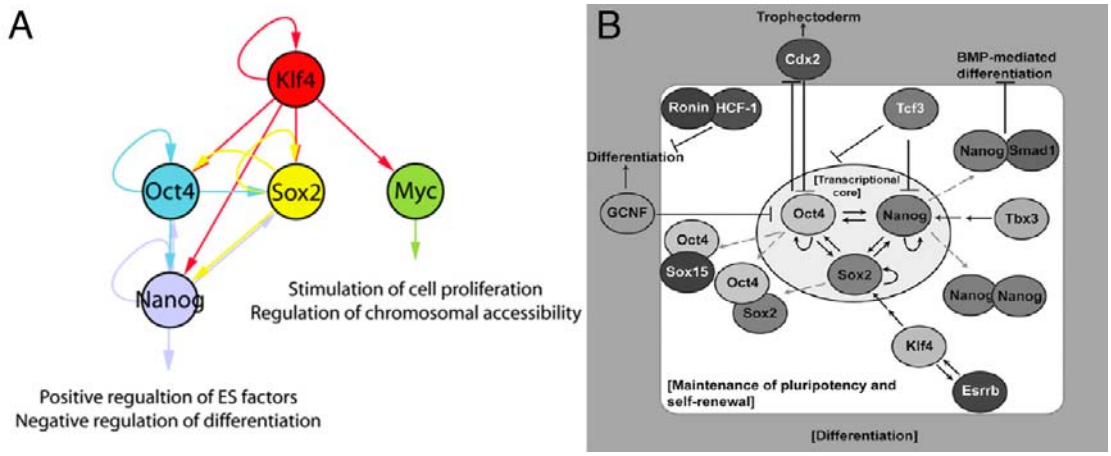


Figure 8. ESC transcriptional regulatory networks. A) Regulatory circuit with four somatic regulatory factors and Nanog. Modified from Kim *et al.* 2008. B) Transcriptional core and interplay with other positive or negative ESC regulating factors. Factors located in the white region are associated with pluripotency and self-renewal, whereas factors located within the grey region are associated with differentiation. Modified from Dominic and Huck-Hui, 2010.

The upstream activation pathways of these factors are essential for the maintenance of the pluripotent state of ESCs. LIF-STAT is essential for self-renewal in mouse ESCs, it activates KLF4 which in turn activates SOX2 [57], and it also promotes NANOG up-regulation by binding of STAT3 to its enhancer [55]. PI3K-AKT and MAPK pathways acting through LIF activates TBX3. The ectopic expression of KLF4 or TBX3 are enough to ensures ESCs pluripotency in the absence of LIF, and both factors are responsible of the LIF-STAT downstream transmision to the regulatory core network in ooder to mantain pluripotency [55]. The BMP/SMAD pathway also plays key roles in the maintenance of mouse ESCs. In contrast, LIF supplementation is not sufficient to maintain human ES (hES) cell pluripotency and neither BMPs which supplementation results in rapid differentiation [58]. On the other hand, FGF and TGF-β/Activin are central for self-renewal in human ESCs [58]. FGF seems to play a

more important role in human ESCs than in mouse EScs, inducing pluripotency and self-renewal through NANOG. However, in the presence of BMP4, FGF2 induces mesodermal differentiation through the activation of Brachyury and CDX2, while inhibiting ectodermal differentiation. When BMP4 acts without FGF2, it induces mesoderm and trophoblast differentiation through CDX2. This BMP induced differentiation can be blocked by the heterodimer formed by NANOG and SMAD1 [59] (Figure 8 B). Other studies have reported BMPs to act together with LIF to enhance self-renewal, and that the BMP4 receptor tirggers SMAD1 and activates Id (inhibitor of differentiation) factor family [56]. The Wnt/β-catenin signaling pathway is also involved in the maintenance of ES cell self-renewal. Interestingly, this pathway may be one of the few common pathways involved in self-renewal in both mouse and human ES cells [58]. Wnt signaling is pleiotropic, with effects that include mitogenic stimulation, cell fate specification, and differentiation of ESCs. Its role in stem cells is relatively complex and quite controversial. It has been described to maintain pluripotency through its inhibition over GSK3-β, either by activating self-renewal factors, such OCT4 and REX1 [60] or by blocking differentiation programs like EMT [61]. The maintenance of the ESC sate after Wnt signaling activation, could go through the non-canonical Wnt pathway, however evidence still lack and it is better reported that the pluripotent state seems to be governned by the participation of the canonical β-catenin pathway, which can binds to TCF1 to promote OCT4 or SOX2 gene expression and, importantly, can also bind to TCF3 and revert its repressor activity [48]. TCF3, co-occupies many genomic sites together with OCT4 and NANOG, and counteracts the activity of OCT4, NANOG and SOX2 resulting negative for pluripotency and self-renewal, as supported by the fact that and its depletion causes ESCs to be less prone to differentiation. When TCF3 is bound to β-catenin, its repressor activity is neutralized, and they exert as an activator complex promoting maintenance of ESCs [62–64] (Figure 7 B). Thus, WNT plays a key role in the maintenance of pluripotency of ESCs, although for long-term maintenance it may requires the presence of additional factors, such as, LIF or MEK inhibitors or FGF2, but some other authors reported that its alone activation it's enough for pluripotency maintenance of ESCs [60].

3.3. Reprogrammation to induced-pluripotent state

The paradigmatic irreversibility of stem cell differentiation was challenged by virtue of the successful transformation of fully differentiated adult cells into ESCs analogs, through the action of a few ESC transcription factors. In 2006, Yamanaka and colleagues reprogrammed adult mouse fibroblasts into induced pluripotent stem cells (iPSCs) similar to ESCs by introducing the transcription factors Sox2, Myc, Klf4 and Oct4 [65]. In 2007, human neonate-fibroblasts were reprogrammed using the same or slightly modified combination of genes (OCT4, SOX2, NANOG and LIN28), although in that report, reprogramming was not attempted with adult human fibroblast [66]. In 2008, adult mouse and human fibroblasts were reprogrammed into iPSC through the introduction of only 3 exogenous factors (SOX2, OCT4, KLF4) [67], even though the study provided evidence that endogenous MYC expression was required for such transformation, suggesting that MYC is still an indispensable factor for induced pluripotency. Thus SOX2, MYC, KLF4 and OCT4 are

considered the four core reprogramming factors for iPSCs. More recently, in 2010, it was reported that the epithelial phenotype, and more specifically, E-cadherin expression, is essential required for the reprogramming of fibroblasts into an induced pluripotent state [68] (this will be the object of discussion later on).

Although it is possible to reproducibly generate iPS cells by viral transduction of the above mentioned defined factors, with the current technology only a small proportion of the transduced cells become pluripotent (iPSCs), implying that the reprogramming of adult cells into a more pluripotent state is still inefficient. This low efficiency and partial reprogramming are technical hurdles that need to be overcome in order to achieve a routine application of human iPS cells in basic research, drug screening, toxicology, or regenerative medicine, and thus the understanding of CSCs biology in tumorigenesis and metastasis. There are two main models that attempt to explain the observed low efficiency of reprogramming [69], namely the "elite model" and the "stochastic model".

The "elite model" assumes that only a small subset of cells, "the elite population", is competent for reprogramming. This model contemplates two variants, the "predetermined elite model" (a small set of cells are competent for reprogramming even before the transduction of the four Yamanaka factors) and the "induced elite model" (cells need to be transduced with the four - or additional - factors and only some cells will be competent for reprogramming) (Figure 9).

In the "stochastic model", most, if not all, differentiated cells have the potential to become iPSCs after the introduction of the four Yamanaka factors. Cell differentiation often starts from a totipotent state, going through the pluripotent state, and rolling all the way down to a lineage-committed state (Figure 9), in the manner of Waddington's epigenetic landscape (or hill) metaphor. In normal development, pluripotent cells appear transiently. They cannot stop on the slope of the "hill" and are pulled down by gravity to rapidly differentiate into various lineages. In contrast, ESCs can self-renew and maintain pluripotency for a long time. Thus, it is as though ESCs are blocked by a bump or a roadblock formed by their particular epigenetic status. In this metaphor, the four reprogramming factors cooperatively push the cells up into the pluripotent state [69].

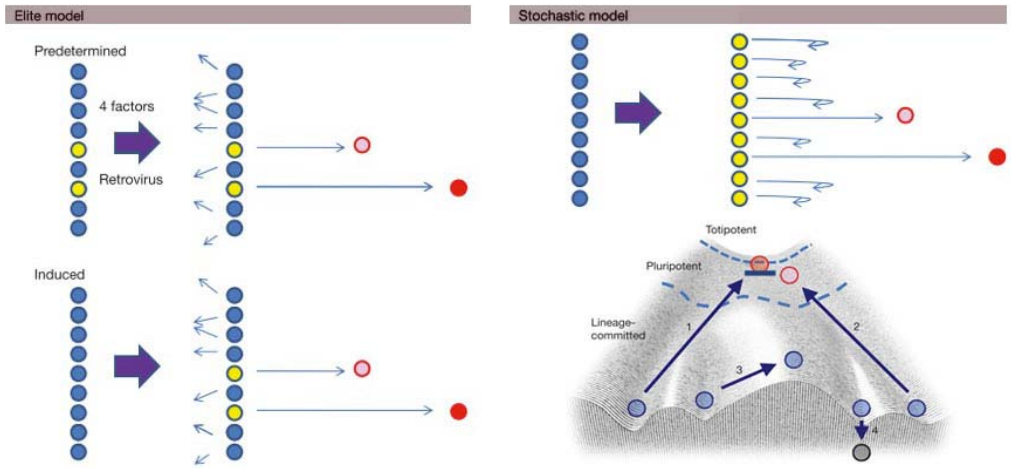


Figure 9. Two models explaining the low efficiency of iPS cell generation. In the "elite model", only a small number of cells, determined either before or after retroviral transduction, can be reprogrammed either partially or completely. In the "stochastic model", most cells initiate the reprogramming process, but only a few can achieve complete reprogramming. Yellow, cell competent for reprogramming; pink, partially reprogrammed cells; red, iPSCs. The "stochastic model" is represented in the context of the epigenetic landscape proposed by Conrad Waddington. The iPSC is like a ball rolling down the slope of a valley. The reprogramming factors cooperatively push cells up the slope to the pluripotent zone. Some cells are blocked by an epigenetic bump (closed rectangle) on the slope and thus become able to self-renew (1). Other cells are only partially reprogrammed and are not blocked by the bump; therefore, without the exogenous reprogramming factors, they would roll down again (2). When the expression of the reprogramming factors is not appropriate, cells may transform to other cell types (3), or even undergo apoptosis or senescence (4). For cells that are located in the middle of the valley (that is, somatic stem cells) it might be easier to go back to the pluripotent state. Modified from Yamanaka, 2009.

3.4. Cancer stem cells (CSCs)

Cancer stem cells are generally defined by three major properties [28]:

- They express a distinctive repertoire of cell surface markers that allow their reproducible and differential isolation. This is an operational prerequisite for the isolation of the CSC population. However, it should be noted that the markers proposed are different for different types of cancers and still controversial.
- They are endowed with tumor-initiating capacity as opposed to all other tumor cell subsets. This implies that cancers contain at least two populations of cells, one being the CSCs, which self-renew and are immortal, and the other being the derived population, which has a limited life span.
- They give rise to the growth of heterogeneous cancer tissues, recreating the full repertoire of cancer cell populations, as a consequence of their abilities to self-renew and differentiate.

The fraction of CSCs represented within a given tumor varies with the type of cancer and the method employed to identify them, anywhere from 0.1 to 30% of all neoplastic cells in a tumor [35] [40]. These variations are owed to the idiosyncratic and unpredictable nature of the genetic or epigenetic perturbations, and their consequences, that result in detectable populations of CSCs. In experimental terms, the CSC hypothesis is reinforced when it is shown that by previous isolation of presumptive CSCs, e.g. by using a biomarker or combinations of biomarkers, or by serial transplantation, the tumor yield is greatly increased after xenotransplantation into immunodeficient mice. This certainly suggests that a fraction of the tumor cells have a superior ability to generate a tumor or, in other words, they are enriched in tumor-initiating cells. On the other hand, the fact that, in most tumors, CSCs represent a minor population, with the majority being derived cells implies that all bulk determinations of any tumor property (gene expression, proteome,…) reflect mainly the majority of derived cells rather than the original cancer stem cells. These and other important CSC properties are summarized in Table 1 [28].

A major issue pending resolution is whether CSCs are quiescent or proliferative. It is important to clarify if this sub-population contains the most competitive cells relative to poorly competitive derived cells, and have the capacity to drive primary tumor and secondary tumor growth. As mentioned above, most of the recent literature and studies in breast and ovarian cancers and in gliomas [28] suggest that CSCs are highly proliferative. This concept fits in well with the strong inverse correlation between the number of multiplying cancer cells and the survival of breast cancer patients, but of course other models could explain this correlation. The fraction of multiplying cells was measured by thymidine labeling or by the expression of proliferation markers, their number by the product of the fraction with the volume [27].

Another issue is whether CSCs and derived cells are qualitatively distinct populations. A dynamic view has been proposed that takes into account quantitative, stochastically determined and at least partially reversible characteristics. One possibility of maintaining over time a fractional population of CSCs would be the existence of control feedback mechanisms such as those operating to maintain the population of somatic stem cells (SSCs) [73]. Other new arguments are that embryonic and somatic stem cells may themselves be in a dynamic state. Pluripotent ESCs lines present a high degree of heterogeneity and interconvertibility [74]. Moreover, there are new arguments in favor of the reversibility of the derived cell status. For instance, it has been shown that a population of senescent human keratinocytes gives rise to new emergent tumorigenic cells [75].

This concept of CSCs *per se* implies at least a dichotomy between two cell populations. However, as a matter of fact, tumor populations are very heterogeneous, with many more than two populations with distinct features. For example, the genetic diversity evident in different varieties of breast cancers is also observed within lesions and invasive regions (intratumoral heterogeneity), and is predictive of progression [7]. This heterogeneity of cancer cells in space is further complicated by a possible heterogeneity in time, attributable to tumor cell plasticity. Thus, the available evidence suggests that CSCs represent a dynamic space-temporal continuum of proliferating cancer cells with varying growth potentials, whose biomarker expression may be either shared or distinct from one cancer to another and in the same cancer at different times [27].

Table 1. Comparative list of properties between CSC and normal stem cells (SSCs and ESCs) [28]**.**

CSCs properties	Stem cells (SSCs or ESCs) properties
Self-renewal and immortality - Tumorigenic capacity - Disorganized niche - Resistance to apoptosis - Self-sufficiency from growth signals - Anchorage-independent growth	Self-renewal and immortality - Not tumorigenic - Well-structured niche - Resistance to apoptosis? - Requirement for growth factors - Anchorage-independent growth
Sustain the growth of tumor - Only a fraction of the population - Recreate the full repertoire of cancer cells	Consistent minority - Subpopulation of stem cells - Give rise different tissues and/or cell types
Undifferentiated - Asymmetric divisions? - Irreversibility? - Limited life span of derived lineages - Disorganized derived cells	Undifferentiated - Differentiation with asymmetric divisions - Irreversible under normal conditions - Organized derived cells - Limited life span of derived lineages
Repertoire of characteristics markers - Large variability	Repertoire of characteristics markers - Well defined
Resistance - Radiotherapy - Chemotherapy - Hypoxia	Resistance - Radiotherapy - Chemotherapy - Hypoxia
No homeostatic control	Homeostatic control
High proliferation rate	Proliferation rate - High in ESCs - Low in SSCs
Expression of pluripotent transcription factors	Expression of pluripotent transcription factors

The mathematical modeling of the evolution of the proportion of CSCs versus derived cells is far from evident but, importantly, it suggests that if CSCs exhibit some stem cell characteristics, they would be those of ESCs, with their rapid multiplication rate, rather than those of the infrequently dividing SSCs.

A clinically important related property is that CSC would be resistant to chemotherapy and radiation, thus being a major factor in cancer recurrence after anti-tumoral therapy [51] [52]. This resistance is believed to stem from the high activity or expression of multidrug resistance, anti-apoptotic proteins and enhanced DNA repair mechanisms. Thus, emerging therapies should target not only signaling pathways and gene networks regulating self-renewal, but also such resistance mechanisms [78]. Moreover, at certain stages of disease, the neoplastic progression may occur either as a consequence of intrinsic tumor pathogenesis and/or as escape mechanisms from chemotherapeutic challenge (Figure 10), similar to the development of bacterial resistance to antibiotics. In this way, selective pressures exercised by treatment may be associated with neoplastic progression and may lead to the development of a higher frequency of functionally defined CSCs in secondary or metastatic stages, as well as to inter-patient and intra-patient variability of CSC properties.

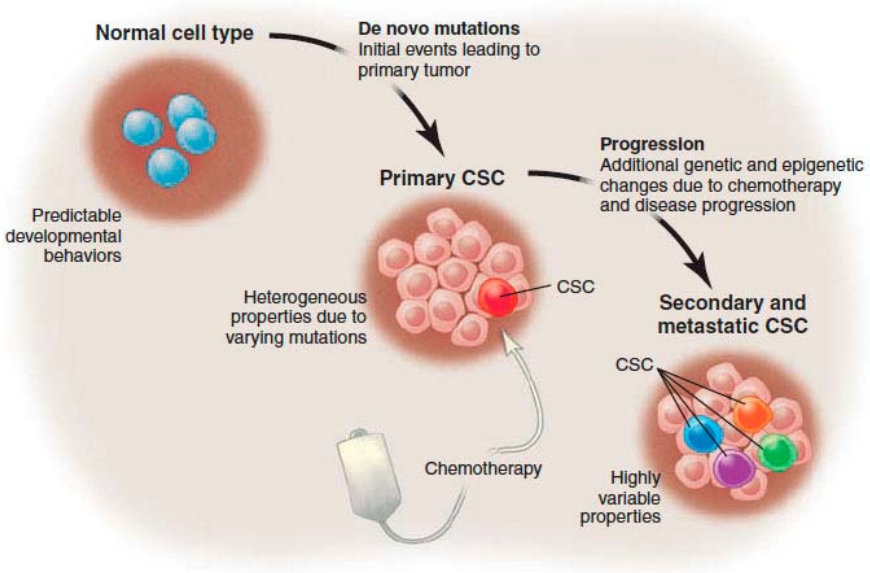


Figure 10. Clinical and therapeutical implications of CSCs. CSCs appear to be influenced both by the specific genetic abnormalities of a given tumor and by the stage of disease progression and the types of drugs used to challenge tumor growth. Rosen *et al.* 2009.

Consequently, for any particular type of cancer, the patient-to-patient variability of CSCs may be quite substantial. Taken together, these issues make any consistent definition of CSC properties difficult. Furthermore, the variability in CSC properties introduces problems when developing new therapies. This increasing complexity of the CSC paradigm requires further studies that better define their nature, attributes and identifying features in different tumor types.

3.5. The stem cell niche and the premetastatic niche

In 1978, Schofield suggested that stem cells live in a niche, i.e. a physiologically defined supportive microenvironment. Normal stem cells reside in a "stem cell niche" that maintains them in a stem-like state. In mammals, niches for adult stem cells have been characterized in the bone marrow, skin/hair follicle, intestine, neural system and testis [79]. Normal stem cells are niche-dependent and their expansion out of the niche is limited to transit amplifying progenitor cells before finally yielding differentiated cells (Figure 11). Cell types and architectures of niches for specific stem cells are variable from tissue to tissue. Nonetheless, many key players in stem cell niches have evolutionarily conserved functions in both normal and malignant tissues, and thus may play key roles in tumorigenesis and metastasis to different target organs.

Recent data suggest that CSCs also rely on a similar niche, dubbed the "CSC niche," which controls their self-renewal and differentiation. Furthermore, and very interestingly, the localization of secondary tumors also seems to be orchestrated by the microenvironment, which is suggested to form a pre-metastatic niche. Thus, the microenvironment seems to be of crucial importance for primary tumor growth as well as metastasis formation [80].

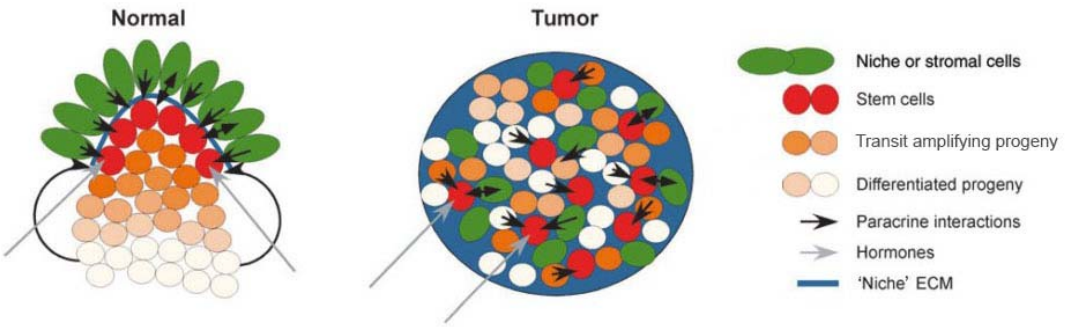


Figure 11. Hypothetical model depicting the organization of stem cells in normal organs and tumors. In normal tissues, the niche cells provide signals enabling normal stem cells to self-renew. Transit-amplifying progenitor cells do not receive this signal and their proliferation is constrained by cellular mechanisms that count the number of mitotic divisions. With each cell division, the proliferation capacity of these daughter cells declines (programmed decline in replication potential), and their degree of differentiation increases. Normal stem cells and their progeny are localized in a highly organized manner in relation to each other and the stem cell niche. In contrast, in tumors, cancer stem cells and their progeny may be randomly distributed. The tumor stem cell niche may in this case be provided both by more differentiated tumor cells and/or stromal cells, and the whole extracellular matrix (ECM) of the tumor could function as a specialized "niche" ECM. Modified from Polyak and Hahn, 2006.

In short, the niche regulates stemness, proliferation, and resistance to apoptosis of stem cells. It has a complex architecture and is composed of diverse stromal cells, such as mesenchymal and immune cells, a vascular network, soluble factors, and extracellular matrix components. Analogously, tumors involve not only the tumor cells themselves, as they also result from the complex interplay between tumor cells and nonmalignant cells that make up the tumor environment. Like normal stem cells, CSCs seem to depend on a similar, permissive environment, the CSC niche, to retain their exclusive abilities to self-renew and give rise to more differentiated progenitor cells, while staying in an undifferentiated state themselves [81]. Moreover, the CSC niche also has a protective role by sheltering CSCs from diverse genotoxic insults, the niche contributing to their enhanced resistance to therapy [56] [57]. This could be more relevant in some tumors, brain and colon cancers being good examples of malignancies in which CSCs seem to rely on a specialized microenvironment.

Genetic and epigenetic changes may occur in stem cells, leading to the generation of CSCs and to the expansion of cells within the niche. This may eventually result in an altered niche as the cells become independent of normal regulatory signals and produce extrinsic factors that deregulate niche-forming cells and disturb the surrounding ECM. And this may further result in an inappropriate production of growth signals and thus generating additional CSCs [23]. Because of this scenario, there is increasing interest in the possibility of exploiting the putative CSC niche for drug targeting.

Thus, CSCs and the niche may interplay through bidirectional regulatory signals towards cancer progression. The niche, also referred to as the microenvironment, might also be implicated in the final steps of the metastatic cascade. It is known that metastases selectively occur in certain organs such as lungs,

liver, brain, and bones. In recent years, evidence has suggested that the primary tumor itself is actively involved in adapting to these so-called premetastatic niches during tumor cell homing, by secreting systemic factors and directing bone marrow-derived cells and macrophages to certain tissues, thereby priming them for subsequent tumor cell accommodation (Figure 12) [58] [59]. Accordingly, VEGFR1-positive bone marrow-derived hematopoietic progenitor cells (HPCs) were shown to localize to premetastatic sites and form clusters before the arrival of tumor cells [84]. Eradication of these cells from the bone marrow prevented the formation of premetastatic clusters and, subsequently, tumor metastasis. Molecular characterization of these recruited hematopoietic cells identified them as expressing VEGFR1, CD133, CD34, and c-KIT [25]. These cells homed to and preconditioned sites of metastasis prior to the dissemination of tumor cells from the primary site. Targeted inhibition of VEGFR1-expressing progenitors using neutralizing antibodies suggested that this preconditioning was necessary for metastatic progression. Furthermore, a subcutaneously inoculated lung carcinoma that induced these bone marrow-derived progenitors to congregate only in the lungs also metastasized only to that site, whereas a melanoma that recruited these progenitors to multiple organ sites where exhibited a widespread metastatic tropism.

In addition to homing of HPCs, preexisting local fibroblasts were noted to increase fibronectin deposition at those sites, which binds to VLA4 integrin (ITGA4), a fibronectin receptor expressed on HPCs, thus facilitating their accumulation at the site of metastases and procuring the “bed” formation, an altered microenvironment furnished with adhesion molecules, proteases, chemoattractants and adecuate growth conditions. Furthermore, activated fibroblasts were shown to induce the remodeling of stroma required for liver metastasis in a murine melanoma model [86]. Thus, in addition to their contribution to the CSC niche at the primary tumor site, fibroblasts are suggested to also have also a critical role in the formation of the premetastatic niche.

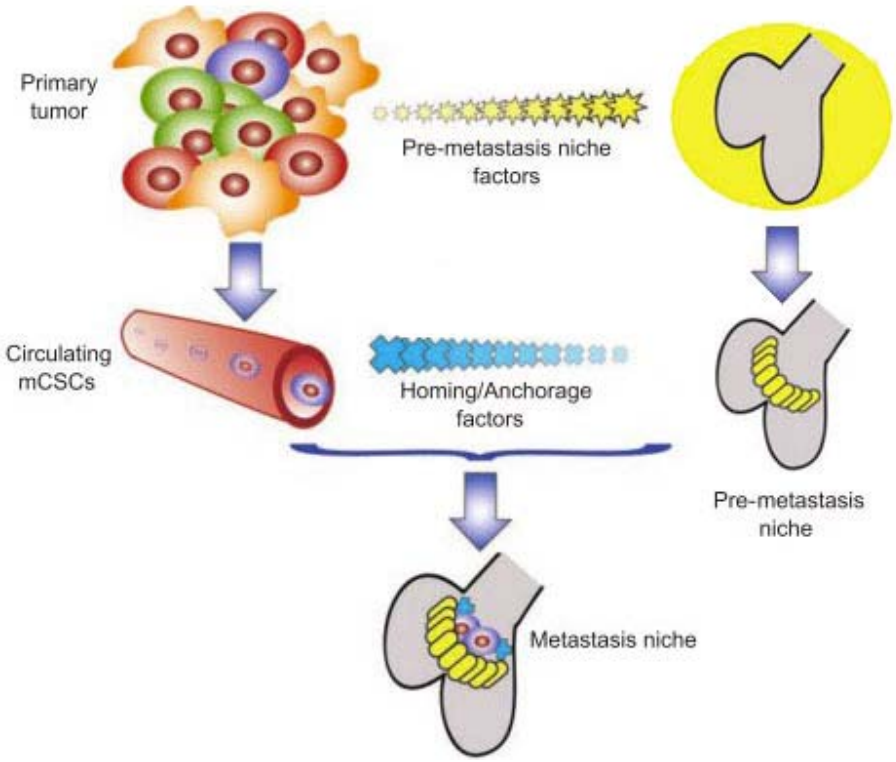


Figure 12. Preparing the bed for metastasis. The heterogeneous CSCs derived from the primary tumor mass, starting the secretion of pre-metastasis niche-forming factors, plays a critical role in determining the tissue tropism of the future metastatic lesion. Once the mCSCs begin to migrate through the blood, they are guided by homing and anchorage factors produced by the pre-established niche. After seeding, the local microenvironment in the niche helps determine if the mCSCs will either proliferate into a metastatic lesion directly, or enter a quiescent period. Li *et al*. 2007.

The importance of the CSCs niche is supported by the fact that the loss of a niche environment leads to the loss of CSCs [80]. Whether premetastatic niches are also capable of imposing a CSC phenotype on more differentiated cells or whether these particular environments are only capable of maintaining the function of metastasized CSCs is still a matter of speculation. Either way, the supporting role of the microenvironment in tumor growth and progression, including metastasis formation, clearly puts the CSC niche and, especially, the mediators of this interaction, in the spotlight as future therapeutic targets. For example, glioblastoma CSCs appear to be maintained by signals from an aberrant vascular niche with potent angiogenic activity. Thus, anti-angiogenic therapy in conjunction with cytotoxic chemotherapy may prove effective in targeting CSCs in glioblastomas [83]. Another example is the importance in colon cancer of HGF produced by surrounding stromal myofibroblasts (or TAFs), capable of enhancing Wnt signaling activity in colon CSCs and thus tumor growth. HGF acts via the tyrosine kinase receptor MET on cancer cells and triggers its downstream targets. Recently, it has been shown that anti-MET antibodies prevent HGF binding to MET and, subsequently, inhibit colon cancer tumor growth [87]. Targeting TAFs also helps to overcome chemotherapy resistance. These examples are only the tip of the iceberg of the extensive studies that explore the modulation of the interaction between cancer cells and niche cells as a therapeutic strategy that could lead to major advances in cancer treatment in years to come.

3.6. CSCs gene transcriptional networks

Modern techniques in stem cell biology in the postgenomic era have led to dramatic advances in our understanding of the molecular underpinnings of both embryonic stem cells (ESCs) and cancer. Detailed gene expression maps have now shown the diversity and distinctiveness in gene expression programs associated with stemness in embryonic and adult stem cells. These maps have further revealed a shared transcriptional program in embryonic stem cells (ESCs) and cancer stem cells (CSCs). The inability to define a consensus stemness signature in a gene-by-gene analysis may suggest that different types of stem cells utilize distinct mechanisms to achieve self-renewal and pluripotency. Alternatively, the failure to identify a robust signature may be due to technical variations in stem cell isolation, degrees of cell purity including contamination from neighbor populations, microarray platforms, or statistical analysis methods.

Indeed, it has been recently shown that embryonic and adult stem cells can be distinguished into two predominant groups based on their gene expression, and surprisingly, cancer stem cells may demonstrate gene expression programs more similar to ESCs than SSCs [88]. In that study, an "ESC-like gene module" was identified based on genes whose promoters are occupied by regulatory proteins conferring pluripotency such as OCT4, NANOG and Polycomb. Based on the motif module map method [89], the c-MYC binding motif was predicted to be the top driver of the ESC module and it was shown to be sufficient to force activation of an ESC-like gene expression program in adult epithelial cells and reprogram them to human epithelial cancer stem cells, thus, achieving pathological self-renewal and tumor initiating capacity [88].

The “ESC-like module” is defined by 335 genes and contains many transcriptional regulators, in particular, several associated with pluripotency, including SOX2, c-MYC, DNMT1, CBX3, HDAC1 and YY1. OCT4 and NANOG are not in the ESC-like module because they are specifically expressed in ESCs. This result implies that some of the transcriptional program mediated by these key ESC regulators, such as OCT4 and NANOG, can be regulated in other stem cells by alternative mechanisms. The ESC-like module is activated in many different human epithelial cancers, including breast, liver, gastric, prostate and lung cancers. The ESC-like transcriptional program is associated with aggressiveness and activated in diverse human epithelial cancers and strongly predicts metastasis and death, particularly in lung and breast cancers, and thus it is clinically relevant [88].

An independent analysis of embryonic stem cell gene signatures by Ben-Porath and colleagues also observed increased expression of ESC signatures in clinically aggressive epithelial cancers, such as glioblastoma and bladder cancer [90]. Histologically poorly differentiated tumors show preferential overexpression of genes normally enriched in ESCs, combined with preferential repression of Polycomb regulated genes. Moreover, activation of targets of NANOG, OCT4, SOX2 and c-MYC are more frequently overexpressed in poorly differentiated aggressive tumors than in well-differentiated less aggressive tumors [90].

Being major epigenetic regulators, embryonic stem cells rely on Polycomb group proteins to reversibly repress genes required for differentiation. In stem cells Polycomb group targets are up to 12-fold more likely to have cancer-specific promoter DNA hypermethylation than non-targets, supporting a stem cell origin of cancer in which reversible gene repression is replaced by permanent silencing, locking the cell into a perpetual state of self-renewal and thereby predisposing to subsequent malignant transformation [91].

In spite of all this new evidence, the precise relationship between the ESC signature and cancer or cancer stem cells it is still not fully clear. A recent study aiming at dissecting a MYC-based network [92] showed that the links between ESCs and cancer are largely explained by MYC. It showed that a MYC-centered protein interaction network interacts with the NuA4 histone acetyltransferase (HAT) complex, suggesting a relevant role for MYC in epigenetic regulation in ESCs. After ChIP analysis of transcriptional targets the study defined a “Myc module” active in ESCs and separable from a “Core module” (composed of the rest of pluripotency factors). This analysis showed that the Myc module is highly expressed and dominant in multiple scenarios: MYC-transformed human epithelial cancers, several mouse myeloid leukemias, some human bladder cancers, and some human breast cancers. Of interest, the “Core” ESC module was not significantly expressed in these situations. The authors of the study point out that, although MYC may affect self-renewal capacity in ESCs and cancer, it may not be a central player in this process because of evidence, discussed above, that MYC is not strictly required for reprogramming [69]. This suggested to the authors that the presence of the Myc module in gene expression signatures from ESC populations and poor prognosis cancers may be more a reflection of the active proliferation occurring in both scenarios, rather than self-renewal [92]. Nevertheless, it is worth reminding, as done by Nakagawa *et al*. [67], that, although the exogenous introduction of MYC may not be required, expression of endogenous MYC is always required for reprogramming to IPCs.

3.7. CSCs and metastasis

There is a growing body of evidence that suggests that metastases develop when distant organs are seeded with CSCs that arise from a primary tumor. This implicates CSCs in the seeding and growth of metastatic lesions. Tissue tropisms associated with cancer metastasis indicate that specific and distinct cellular and molecular mechanisms are involved. The prevailing clonal selection model of metastasis contends that genetic mutations attained late in tumorigenesis provide a selective advantage for cells to metastasize. However, recent studies lend support to the notion that metastasis capacity is pre-determined by genetic changes acquired at early stages of tumor development [93].

It is now appreciated that at least two classes of determinants affect site-specific metastatic outgrowth. Firstly, there must be an initiation of a viable premetastatic niche within the target organ that facilitates the initial survival of

extravasated tumor cells in a non-receptive target organ. Secondly, the invading metastatic cell must display the appropriate functions to effectively colonize the new site.

Several characteristics of CSCs, referred to above, make them likely candidates to occupy and thrive in these foreign sites. It is theoretically possible that only CSCs within tumors have the ability to initiate and sustain cancer growth. It has been known for years that just one cell can initiate a metastatic lesion [94]. The inherent plasticity of stem cells makes them more adept to survive in a foreign environment (primed by the pre-metastasis niche) where growth factors and other signaling molecules are different from those at the primary tumor site. Increased genetic instability in CSCs is also likely to provide a selective advantage in adapting to foreign sites. However, it remains to be seen whether all CSCs are equally capable of forming metastasis at different sites. Tumor-initiating capacity is required at any metastasis site along with an appropriate niche. However, the immediate progeny of the metastatic CSCs (mCSCs) may succumb unless they have an ability to exploit that environment [25]. In other words, CSCs may be necessary to re-initiate the tumor in the strange new environment at metastasis sites, but they will be insufficient to maintain metastatic growth if their progeny cannot survive owing to a lack of organ-specific adaptability.

Since only a few CSCs would need to leave the primary tumor site to initiate metastasis, this could help reconcile the observation that cancer cells can be detected at distant sites long before any detectable dissemination occurs at the primary tumors. At metastatic sites, mCSCs maintain most of the genetic programs acquired at the primary tumor site through self-renewal, which explains the phenotypic similarities between primary and metastatic cancers. However, mCSCs at secondary sites are also able to evolve independently by accumulating additional genetic alterations that render them resistant to treatments that are effective against primary tumors [25].

4. Metastasis: Dissecting the multi-step process

Metastasis is a complex multi-step process. The whole dynamic process includes many steps and phenotypic transformations, starting from the acquisition of an aggressive phenotype enabling distant tissue colonization (Figure 13).

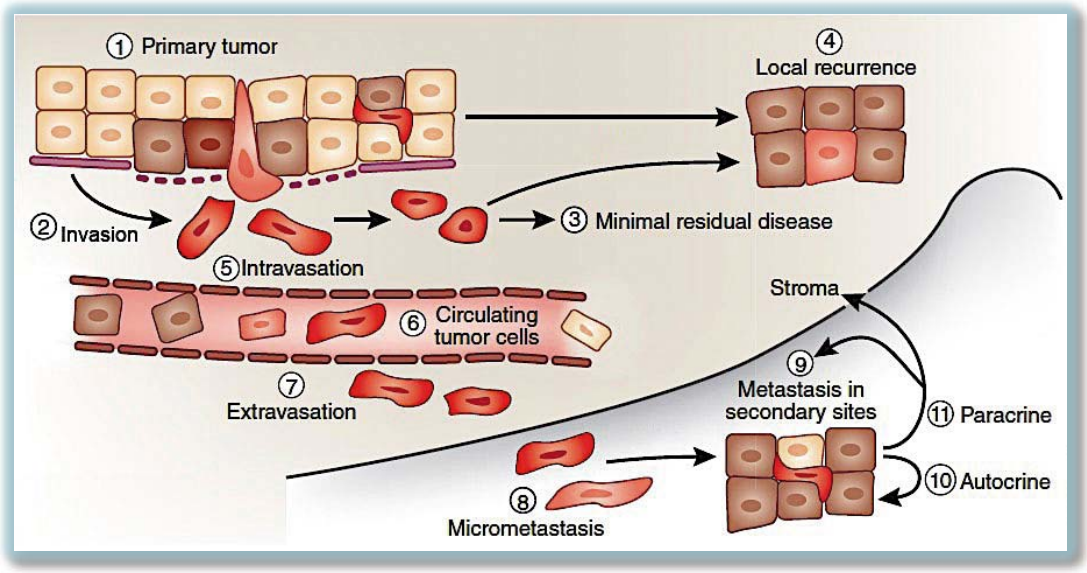


Figure 13. Stages of metastatic progression. 1) Carcinoma *in situ*, acquisition of aggressive phenotype. 2) Loss of adhesion, ECM degradation, Invasive front. 3) Residual cancer cells invading through the surrounding ECM. 4) Local re-seeding and recurrence in the same tissue of origin. 5) Intravasation into the bloodstream. 6) Vascular transport and survival in the bloodstream. 7) Homing attachment and extravasation. 8) Micrometastasis established which can latter form a metastasis. 9) Metastatic colonization of distant tissues. 10 and 11) Autocrine and paracrine interaction with the stroma in order to be adapted at the distant tissue. Thompson, 2010.

4.1. Acquisition of the aggressive phenotype

Gain- and loss-of-function mutations or epigenetic modulations of oncogenes and tumor suppressor genes, respectively, configure the tumor characteristics, capacities and predisposition to evolve to a metastatic behavior. Multiple layers of mechanisms that suppress tumor formation challenge the inappropriate proliferation of cells harboring oncogenic lesions. Several of these barriers are cell intrinsic (such as the genotoxic stress induced by oncogenes, the expression of growth inhibitory, apoptotic and senescence pathways, and telomere attrition). Evasion from these tumor suppressive pathways is a hallmark of primary tumors. Other extrinsic factors in the tumor microenvironment limit tumor progression, including extracellular matrix components, basement membranes, reactive oxygen species, the limited availability of nutrients and oxygen, and attack by the immune system. How tumors cells respond to these external cues influences, sometimes in dramatic fashion, their metastatic potential.

As discussed above, the long-term tumorigenic potential of some tumors may rely on a small proportion of malignant cells endowed with a capacity to indefinitely self-renew, the tumor-initiating or cancer stem cells (CSCs).

4.2. Invasive front

The invasive front is the leading edge of the tumor facing the tumor periphery. Cells need to acquire the capacity to move and invade, breaking the basement membrane, ECM and endothelial layers. Although the invasive front can advance through different modes of cell invasion (as discussed below), a major mechanism in many epithelial tumors entails the loss of E-cadherin-mediated adhesions as they progress toward malignancy in order to migrate and invade as individual cells detached from the bulk of the tumor. Frequently, this is achieved through transcriptional repression of E-cadherin, but can also involve inactivating mutations that predispose to gastric cancer; epigenetic silencing; proteolytic cleavage; and proteosomal degradation [95]. The transcriptional repression of E-cadherin associated with tumor cell migration and invasion can be part of a broader program resembling an epithelial-to-mesenchymal transition (EMT). EMT can occur in cancer cells upon activation of specific transcription factors (such as SNAI1, SNAI2, TWIST), many of which are involved in EMT during embryogenesis (to be described in more detail below). Integrins are also emerging as important mediators of the malignant phenotype during oncogenic transformation and progression [96]. In particular, the α6β4 integrin, which binds to the extracellular matrix protein laminin, forms signaling complexes with oncogenic receptor tyrosine kinases, including MET, EGFR, and HER2 [96]. Thus, diverse alterations in adhesive properties allow cancer cells to disobey the rules of tissue architecture and to advance in their malignant progression.

At its very bare, cell motility implies the movement of cells from one site to a second site. A molecular depiction of cell migration in *in vitro* models has emerged, which involves dynamic cytoskeletal changes, cell-matrix interactions, localized proteolysis, actin-myosin contractions, and focal contact disassembly [97]. Nodes of regulation include small GTPases (such as RHO, CDC42, and RAC), integrin-containing focal adhesion assembly and disassembly, secreted and plasma membrane-tethered proteases, and the actomyosin contractile machinery. Growth-factor signaling, such as that mediated by HGF through the Met receptor, can modulate many of these activities either directly or indirectly [98].

Basement membranes that underlie epithelial and endothelial cell layers are a dense meshwork composed of several glycoproteins and proteoglycans (such as type IV collagen, laminin, perlecan). A well-organized basement membrane is an integral contributor to epithelial structure, but in cancer the basement membrane acts as a barrier to the invasion of transformed cells into the subjacent stroma. Tumor cells that are able to proteolytically disrupt the basement membrane can progress to overt malignancy and metastasis. The activity of extracellular matrix proteases is normally under tight control through specific localization, autoinhibition, and secreted tissue inhibitors [98] [99]. Cancerous cells use diverse mechanisms to disrupt this tight regulation and unleash proteolytic activities on the basement membrane and interstitial extracellular matrices. In addition to facilitating tumor invasion, extracellular proteases may generate a diverse array of bioactive cleaved peptides. These products can modulate migration, cancer-cell proliferation

and survival, and tumor angiogenesis. Adding to the complexity, some of the pleiotropic effects of matrix metalloproteinases (MMPs) may actually antagonize tumor growth [100]. The physiological importance of MMPs is evident by the extensive joint disorders caused by MMP inhibitors in clinical trials, which has so far deterred the effective use of these agents in anticancer therapy [101].

4.3. Intravasation

In order to metastasize, cancer cells must invade tumor-associated vasculature to gain access to distant sites in the body. This is facilitated by the need of developing tumors to establish a neo-vasculature in order to grow beyond the diffusion limit of preexisting blood vessels [102]. The acquisition of this angiogenic phenotype, termed the "angiogenic switch", represents a vital step in the evolution of solid tumors [102]. This event occurs partly through the induced outgrowth of the preexisting vasculature and partly through *de novo* recruitment of vascular cell precursors from the circulation. Lymphangiogenesis is also observed in advanced primary cancers and is another commonly used way to disseminate. This results in a tortuous network of lymphatic vessels designed to collect interstitial fluid effusions and carry them to lymph nodes and subsequently into hematogenous circulation [103]. Perhaps because lymphatic vessels are more leaky in their design than blood vessels, owing to the lack of tight intercellular junctions between lymphatic endothelial cells, the presence of lymph-node metastasis often represents an early prognostic indicator of tumor invasiveness and metastatic dissemination in several types of carcinomas such as melanomas, colorectal cancer and others [103]. However, other metastatic malignancies, such as sarcomas, are notorious for metastasizing to distant sites without any prior evidence of local spread to regional lymph nodes. Anyway, it is thought that access to all organs of the body (lymph nodes excluded) is predominantly through the hematogenous circulation.

The molecular mechanisms controlling intravasation remain to be fully defined. It is possible that once cancer cells become highly motile within primary tumors, they are naturally attracted to blood vessels due to chemoattractive gradients and extracellular matrix tracks emanating from (or terminating) there. Indeed, this was observed in intravital imaging studies of experimental mammary carcinomas [104]. Technological advances that enable the isolation and genomic analysis of circulating cancer cells from patients and in experimental tumor models may yield novel insights into the molecular prerequisites of malignant intravasation.

4.4. Bloodstream circulation

Once malignant cells have invaded this circulatory compartment, they attain ready access to virtually all organs of the body. However, since most disseminating tumor cells die, in order to colonize distant organs they must be able to survive several stresses, including physical damage from hemodynamic shear forces, and immune-mediated killing. Circulating tumor cells may promote their survival by co-opting blood platelets, using them as shields. Clinically observed for well over a century, malignancies have a tendency to induce a hypercoagulable state in their

hosts [105]. Histopathological analysis of early-stage hematogenous metastases in humans frequently reveals the coexistence of thrombosis, with abundant fibrin deposition [106]. Disrupting the interaction between tumor cells and platelets in experimental models has validated this relationship as important for metastasis to multiple target organs [105]. Consequently, tumor emboli are believed to possess greater metastatic potential than naked tumor cells, owing at least in part to their resistance to immune-mediated mechanisms of clearance and to physical hemodynamic forces [105]. Understanding the detailed mechanisms that underlie tumor-cell and platelet adhesion and interaction, as well as selective ways of inhibiting them without disrupting physiological hemostasis, may translate into promising antimetastatic therapies when initiated early in the course of disease progression.

Consideration has also been given to mechanisms that may allow evasion of cell death that is induced by the loss of adhesive supports, called anoikis. The brain-derived neurotrophic factor (BDNF) receptor trkB conferred resistance to anoikis to immortalized cells *in vitro* and increased the metastatic activity of a rat intestinal epithelial cell line [107]. However, the relevance of anoikis in the process of metastasis remains uncertain. In humans, it may take few minutes for a malignant cell departing from a primary tumor to encounter a capillary bed and adhere to the vascular wall. If the time that circulating tumor cells spend devoid of adhesion is so short, anoikis may not be a very significant impediment during the physiological progression of metastasis.

4.5. Homing and extravasation

Rapid mechanical lodging in capillaries and association with platelets are likely a prevalent form of tumor cell entrapment in distant organs. However, it is also possible that the initial homing of disseminated cancer cells to a secondary organ involves adhesive interactions between cell-surface receptors expressed on malignant cells and their cognate ligands expressed in various target sites for metastasis. Integrin receptors have been proposed to participate in such homing interactions. For example, α3β1 integrins expressed on circulating tumor cells have been shown to bind to laminin-5 within exposed regions of the vascular basement membrane during lung metastasis [108]. Metadherin has been identified as a receptor in breast cancer cells to selectively bind on lung-capillary endothelial cells [109]. In addition to adhesion receptors, chemokines and their receptors have also been implicated in metastatic cell homing to target tissues. For example, CXCR4 expression in breast cancer cells was shown to be important for metastasis to CXCL12-rich tissues, such as the lungs [110]. These and other receptor-ligand pairs may account for some of the heterogeneity in target-tissue homing exhibited during metastatic dissemination of different primary tumors. The pre-metastatic niche formation, induced by the primary tumor, can play key role in tissue homing, as discussed above.

Having invaded and endured the circulation, metastatic cells must at some point escape once again, but this time out of the endothelial vasculature and into a

target tissue in a process called extravasation. Exactly when this event occurs along the metastasis cascade may vary from tumor to tumor. In some cases, considerable growth within the intravascular space may occur until the lesion physically bursts through the limiting surrounding vasculature [111]. The cytoskeletal anchoring protein ezrin may facilitate this escape in metastatic osteosarcoma cells. Inhibiting the expression of ezrin in these cells resulted in higher rates of cancer-cell death prior to metastatic extravasation into the lung parenchyma [112].

Are there signals emanating from metastatic cells that actively induce changes in the vascular permeability of blood vessels in target organs? One prime candidate is VEGF, initially identified as a potent vascular permeability factor [113]. The activation of Src family kinases in endothelial cells exposed to VEGF induces disruptions in endothelial cell junctions, which can facilitate metastatic extravasation [114]. Further exploration of molecular players mediating this potentially rate-limiting step of metastatic progression will determine if it occurs within a therapeutically susceptible time frame.

Different tissues oppose different barriers against malignant cells, eventually determining the kind of tumor cell able to reach certain tissues, thus influencing the specific tissue-organotropism of cancer cells. As an example of organ-specific differential architectural barriers to tumor cell penetration, the fenestrated structure of bone marrow sinusoid capillaries is more permissive to cancer cell infiltration than the contiguous structure of lung capillary walls. On the other hand, brain capillaries are more difficult to penetrate, owing to the structure of the haematoencephalic barrier (Figure 14). Infiltration through these barriers selects for tumor cells that express the required extravasation functions. These functions can be provided by genes for which expression in primary tumors independently provides a selective growth advantage (such as vascular remodeling) or by genes for which expression in primary tumors provides no benefit but is a consequence of tumor microenvironment signals [5].

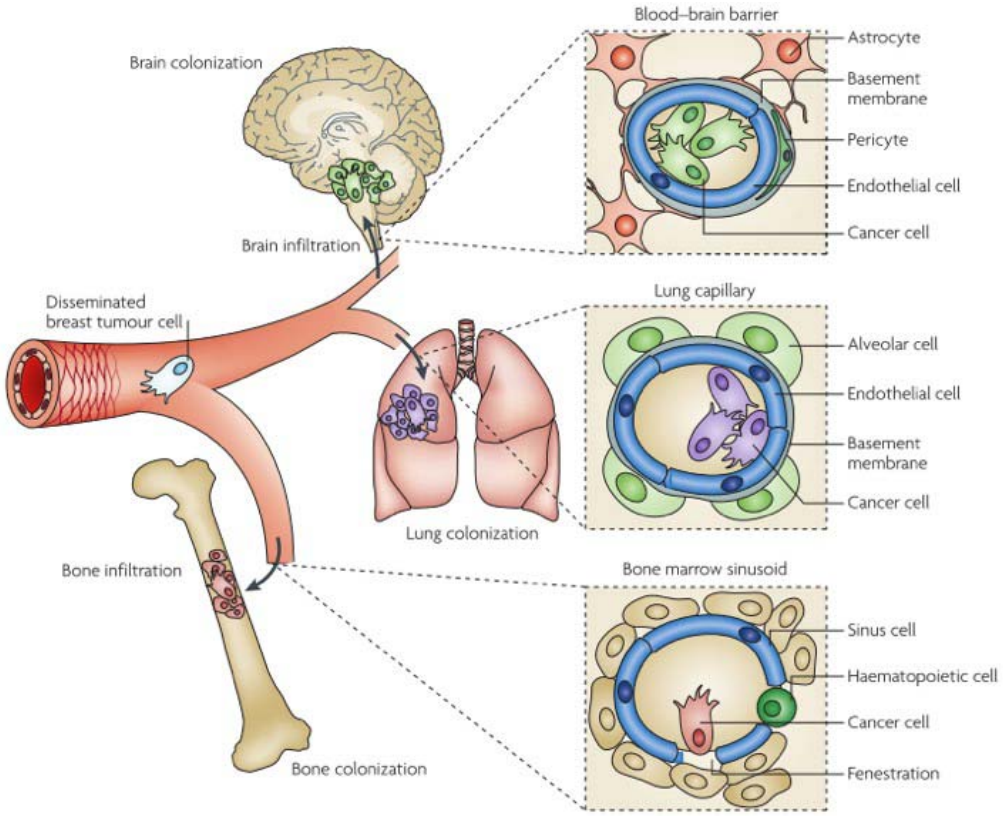


Figure 14. Organ-specific barriers to metastatic infiltration. The blood-brain barrier presents serious difficulties for cancer cells to penetrate. Lung barrier is a contiguous capillary wall. The bone marrow represents a permissive structure for infiltration, since it is an open fenestrated structure of sinusoid capillaries. Nguyen *et al.* 2009.

4.6. Distant tissue colonization

The organ distribution of metastases derived from a primary tumor is not random. After analyzing secondary cancer outgrowths in a series of autopsies for breast-cancer victims, Stephen Paget in 1889 proposed the "seed and soil" hypothesis, where disseminated cancer cells, or "seeds," would only colonize organ microenvironments, or "soils," that were compatible with their growth [115]. Clinical observation of cancer patients supports the notion that circulatory patterns alone provide only a partial explanation for preferred sites of metastasis [116].

General steps of metastasis might be the same in all tumors types but metastasis to different organs might require distinct sets of infiltration and colonization functions, which are acquired over variable periods of time [5]. For example, breast cancer frequently metastasizes to the lungs, bones, liver, and brain, most of which do not have a direct circulatory connection to breast tissue; prostate cancer has a more selective pattern of metastatic recurrence, with bone being the predominant site, whereas visceral organs such as the lungs or liver are much more rarely involved; lung adenocarcinoma metastatizes to brain, bones, adrenal gland and liver; colorectal and pancreatic cancers typically metastasize to liver and lungs; skin melanoma to lungs, brain, skin and liver; uveal melanomas metastasize with high specificity to the liver and sarcomas to the lungs.

One prediction of Paget's seed-and-soil hypothesis is that metastatic cells will only colonize compatible target tissues, even if they are artificially targeted in large numbers to inhospitable sites. A demonstration of this phenomenon in humans was serendipitously obtained in ovarian cancer patients that received palliative peritoneovenous shunting of their ascites fluid into the jugular vein [117]. This procedure inadvertently allowed the release of millions of metastatic cancer cells directly into the venous circulation of cancer patients over the remainder of their lives. Strikingly, the majority of patients did not develop disseminated metastases, sometimes even after two years of continuous vascular shunting. Furthermore, even when metastases were observed after autopsy, they were frequently indolent growths.

Thus, the vast majority of tumor cells that have undergone extravasation still are not able to effectively colonize the new site and may enter in a state of dormancy (latency) or die. For several types of carcinomas, micrometastasis can be detected in the bone marrow years before the development of overt metastasis [84]. The existence of such minimal residual disease represents a predictive factor for disease recurrence and overall survival and is similar to lymph-node metastasis as an indicator of systemic disease [118]. Regardless, most of these cells will fail to convert into macrometastasis. In oestrogen receptor positive breast cancer, prostate cancer and ocular melanoma, metastasis might become manifest decades after removal of even a small primary malignancy. During this period of latency a subset of these disseminated tumor cells can accumulate the full set of functions that are required for overt colonization. However in other types of cancer, metastasis follow a swift course with rapid expansion in multiple organs that leaves little margin for specification of the metastatic cell population, like in lung and pancreatic adenocarcinomas, in which malignant cells might rapidly acquire activities that confer both infiltration and colonization competence [5]. Colorectal carcinoma is a defined paradigm of malignant progression in which most metastatic traits seem to be acquired during local progression at the primary sites.

To escape dormancy or to colonize a new organ outright, disseminated tumor cells must have the capacity to productively interact with the new microenvironment in order to extract growth and survival advantages. In addition to the structural barriers discussed above, different organ microenvironments may impose distinct requirements on cancer cells for full-blown colonization through distinct mediators of organ-specific metastasis. Here the formation of a pre-metastatic niche, discussed above, may play an important role. Mechanistic dissection of secondary organ colonization in model systems has begun to identify sets of mediators that may be necessary to complete this late stage of metastatic progression.

5. Epithelial-mesenchymal transition and its importance in metastasis

Epithelial-mesenchymal transition (EMT) is an indispensable mechanism during morphogenesis, as without mesenchymal cells, tissues and organs will never be formed. It is a highly conserved cellular program that allows polarized, immotile epithelial cells to convert into motile mesenchymal cells. This process was initially recognized during several critical stages of embryonic development and has more recently been implicated in the promotion of carcinoma invasion and metastasis.

During early embryogenesis of most metazoans, mesenchymal cells arise from the primitive epithelium. These cells exhibit a front-back end polarity, in contrast to the apical-basolateral organization and polarization of epithelial cells, and rarely establish direct contacts with neighboring mesenchymal cells [119]. Epithelial cells engage in tight cell-cell contacts with neighboring epithelial cells and do not detach and move away from the epithelial layer under normal conditions. In contrast, mesenchymal cells contact neighboring mesenchymal cells only focally, and are not typically associated with a basal lamina. Moreover, mesenchymal cells, but not epithelial cells, can invade as individual cells through ECM constructed by epithelial sheets and by mesenchymal cells themselves.

EMT was defined as a distinct cellular program in 1980s by Greenburg and Hay. They showed that when epithelial cells from embryonic and adult anterior lens were cultured in 3D collagen gels, these cells elongated, detached from the explants, and migrated as individual cells [120]. Based on the mesenchymal morphology and the pseudopodia and filopodia structures of these migrating cells, they concluded that differentiated epithelial cells could be transformed into mesenchymal cells through a cellular program they named epithelial-mesenchymal transition.

EMT and its resulting phenotype is defined with 4 major functional and molecular changes affecting epithelilal cells [121]:

- Morphological changes from a rounded-like monolayer of epithelial cells with an apical-basolateral polarization to a dispersed, spindle-shaped, fibroblast-like morphology with pseudopodia and filopodia.

- Cytoskeletal rearrangements, as polarized actin cytoskeleton and cytokeratin intermediate filaments of epithelial cells are exchanged for vimentin and fibronectin of mesenchymal cells.

- Changes in the function and/or expression of molecules essential for the establishment and maintenance of specialized cell-cell junctions characteristic of differentiated epithelial cells (E-cadherin, desmoplakin and certain integrins), leading to a loss of such specialized membrane structures (tight junctions, adherens junctions, desmosomes and gap junctions), and a loss of cell-cell and cell-basal lamina contacts [122].

- Functional changes associated with the conversion of stationary cells to motile cells that can invade through ECM. The acquisition of the ability to migrate and invade ECM as single cells is considered a functional hallmark of the EMT program.

During development, a number of extracellular signals can convert epithelial cells into mesenchymal cells by triggering EMT and has been observed to participate in a variety of tissue remodeling events: mesoderm formation, neural crest development, heart valve development, secondary palate formation, and male Müllerian duct regression [123]. Furthermore, similar cell changes are recapitulated during pathological processes, such as fibrosis and cancer. In any case, the EMT program does not necessarily represent an irreversible commitment (Figure 13). Thus, the reverse program, called mesenchymal-epithelial transition (MET), also occurs both during embryonic development and during several pathological processes [121] [124]. The reversibility of this process highlights the plasticity of the epithelial cells of certain embryonic and adult tissues that participate in physiological or pathological processes. Depending on the phenotypic output, epithelial plasticity can be grouped into different types and subtypes [125]. The first division draws a line between the mechanisms of epithelial plasticity used to attain transdifferentation (tissue metaplasia) and the EMT-MET processes (Figure 15). The latter, in turn, can be divided into three subtypes: (1) EMT during development; (2) EMT in fibrosis and wound healing processes; and (3) EMT in cancer progression and metastasis.

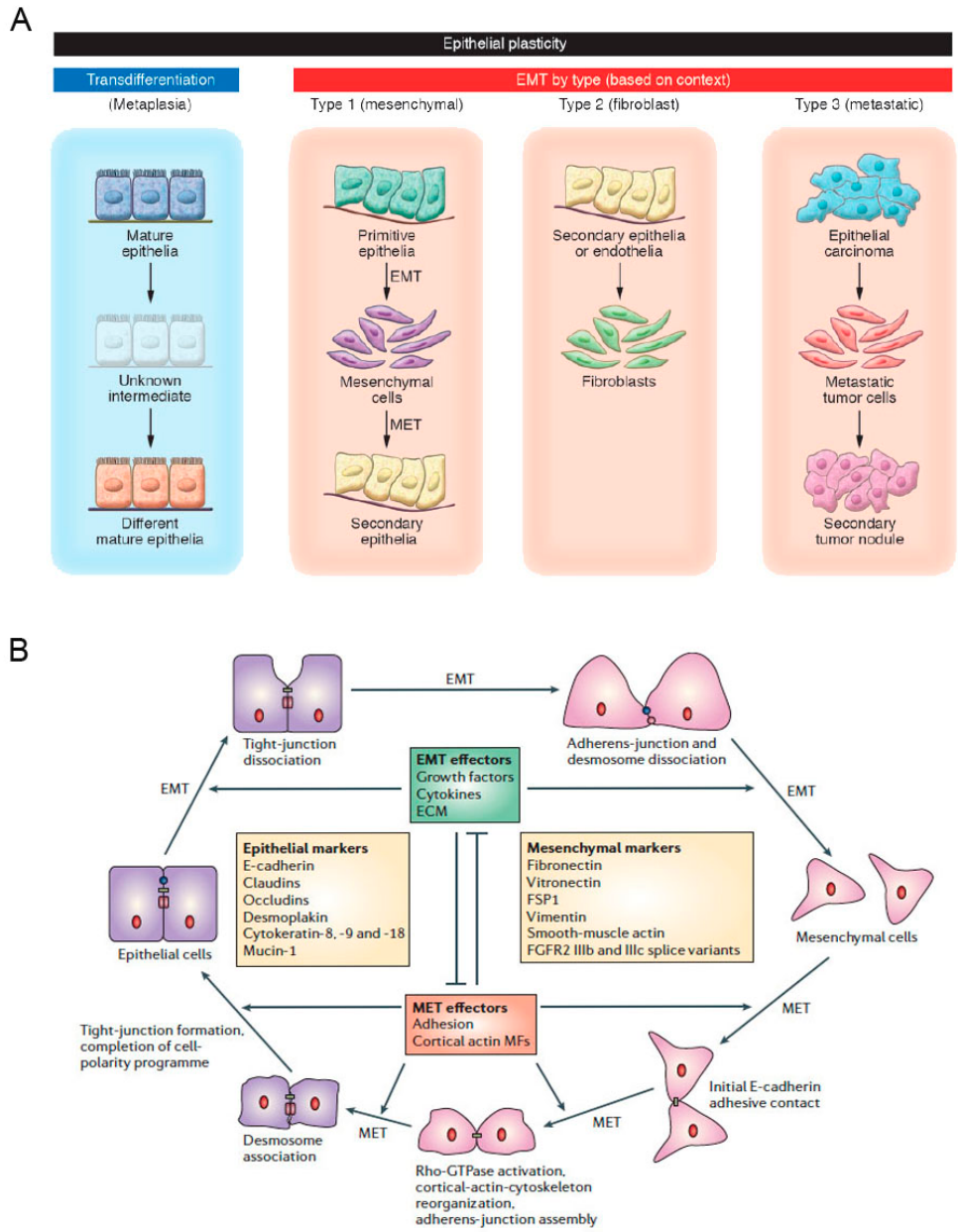


Figure 15. Epithelial cell plasticity as a form of either transdifferentiation or EMT. A) Transdifferentiation generally refers to a process whereby one mature epithelial cell phenotype converts into a different mature epithelial cell, with or without cell division. Three types of EMT are recognized depending on the phenotype of the output cells. Type 1 EMT is seen when primitive epithelial cells transition into mesenchymal cells that form the diaspora of the basic body plan following gastrulation or neural crest migration. Type 2 EMT is seen when secondary epithelial cells or endothelial cells populate interstitial spaces with resident or inflammation-induced fibroblasts in wound healing after injury. Type 3 EMT is part of the metastatic process, whereby epithelial tumor cells leave a primary tumor nodule, migrate to a new tissue site, and reform as a secondary tumor nodule. Zeisberg and Neilson, 2009. B) The cycle of epithelial-cell plasticity between EMT and MET. The different stages during EMT (epithelial-mesenchymal transition) and the reverse process, MET (mesenchymal-epithelial transition) are regulated by effectors of EMT and MET, which influence each other (see text). Thiery and Sleeman, 2006.

EMT events are instigated and regulated by a particular subset of EMT activators and repressors (Figure 16). They proceed through interplays between extracellular signals, components of the extracellular matrix (ECM) (collagen,

hyaluronic acid) and soluble growth factors. Participating molecular pathways in EMT regulation include: TGF-β signaling, Wnt signaling, Notch pathway, Integrin-FAK signaling, FGF, EGF or HGF tyrosine kinase receptors, small GTPases (RAS, RHO and RAC) or Src signaling [126]. For example, transforming growth factor-β2 (TGF-β2) regulates EMT in the atrioventricular canal, whereas TGF-β3 regulates palate-fusion EMT [127]. Furthermore, the effect of a given inducer on EMT is context-dependent. For example, scatter factor/hepatocyte growth factor (SF/HGF) induces EMT during somitogenesis, but it inhibits EMT in other processes [128]. Thus, several mechanisms are involved in the initiation and execution of EMT in development, and the molecular mechanisms that regulate EMT substantially overlap those that control cell adhesion, motility invasion, survival and differentiation. Extensive crosstalk exists between the involved signaling pathways with many common endpoints, including the downregulation of E-cadherin expression or the expression of EMT transcriptional regulators such as SNAIL, ZEB and bHLH factors [129–132], that together, orchestrate the disassembly of junctional complexes and the changes in cytoskeletal organization. Downregulation of E-cadherin has several important consequences that are of direct relevance not only to EMT but to a more general gene program regulation. When E-cadherin levels become limiting, there is a loss of E-cadherin-dependent intercellular epithelial junctional complexes, and E-cadherin-mediated sequestering of β-catenin in the cytoplasm is abolished. As a result, β-catenin localizes to the nucleus and feeds into the Wnt signaling pathway by activating transcriptional regulation through LEF/TCF4.

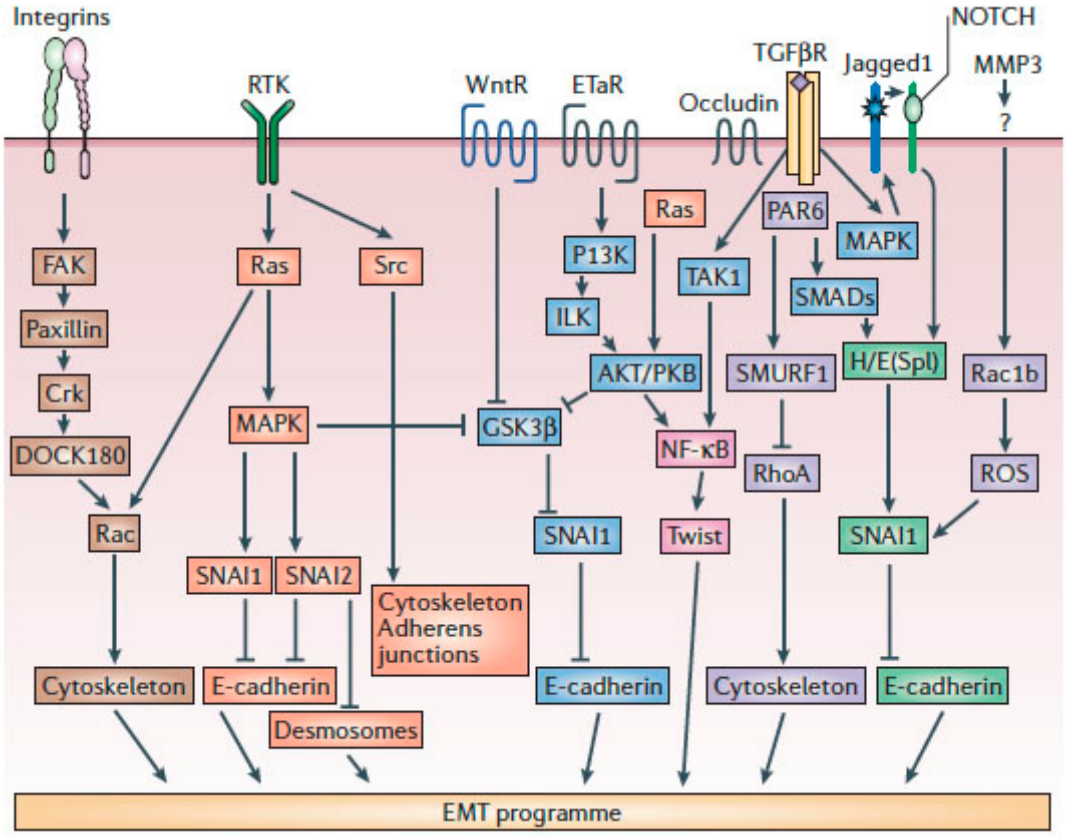


Figure 16. Networks that regulate EMT. A selection of the signaling pathways that are activated by regulators of EMT and a limited representation of their crosstalk. Activation of receptor tyrosine

kinases (RTKs) is known to induce EMT in several epithelial cell types and *in vivo*, but it is now clear that the EMT process often requires co-activation of integrin receptors. The role of transforming growth factor-β (TGF-β) signaling in EMT is established for a limited number of normal and transformed cell lines, whereas *in vivo* data has indicated a mutual regulation of the TGF-β and Notch pathways during EMT. There is now increasing evidence that other signaling pathways could have an important role in EMT, including G-protein-coupled receptors. Matrix metalloproteinases (MMPs) can also trigger EMT through as-yet-undefined receptors. ETaR, endothelin-A receptor; FAK, focal adhesion kinase; GSK3β, glycogen-synthase kinase-3β; H/E(Spl), hairy/ enhancer of split; ILK, integrin-linked kinase; MAPK, mitogen-activated protein kinase; NF-κB, nuclear factor-κB; PAR6, partitioning-defective protein-6; PI3K, phosphatidyl- inositol 3-kinase; PKB, protein kinase-B; ROS, reactive oxygen species; TAK1, TGFβ- activated kinase-1; TGFβR, TGFβ receptor; WntR, Wnt receptor. Thiery and Sleeman, 2006.

5.1. EMT in cancer progression

EMT is considered a key step in cancer progression and metastasis, as a provider of the critical impetus for the dissemination of carcinoma cells from primary epithelial tumors [133]. The initial link associating EMT with cancer originated from studies that suggested the requirement of a downregulation of E-cadherin for tumor progression. This link was firmly established in 2000 when two studies described, in tumors, the repression of E-cadherin by the transcription factor SNAI1 with a consequent induction of EMT and an invasive behavior [129] [131].

As mentioned above, a common spectrum of morphological and gene-expression patterns are associated with the instauration of EMT during morphogenesis, and recent studies have shown a remarkable similarity between these signaling pathways and molecular events with those regulating EMT in pathological processes. The resultant changes in adhesive properties, activation of proteolysis and motility, allow the tumor cells to escape from the primary tumor and reach distant sites (Figure 17).

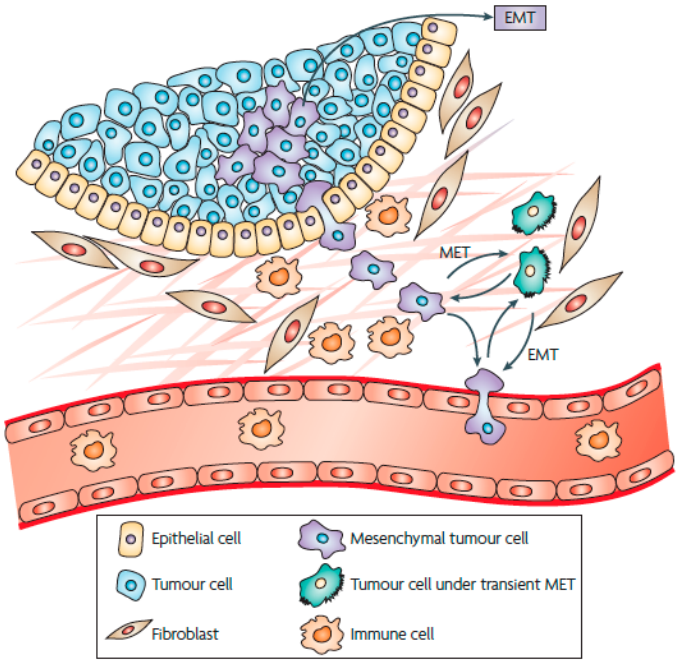


Figure 17. EMT allow the escape. Cells that undergo EMT during tumor invasion are characterized by a loss of cell-cell adhesion and polarity accompanied by cytoskeleton rearrangements and increased cell motility. These invasive cells can break the basal lamina and open paths through the extracellular matrix (ECM), until they reach the bloodstream and intravasate into it. Some cells that had previously undergone EMT could transiently re-acquire an epitheloid phenotype by reverse mesenchymal-epithelial transition (MET) as the result of new interactions with the tumor microenvironment. Peinado *et al.* 2007.

As a consequence of, EMT is increasingly being recognized as a significant therapeutic target in cancer. However, and in spite of the numerous experimental evidences in support of its relevance in tumor progression, there are still controversial views on EMT from pathologists, who often claim insufficient or even opposing evidences stemming from histological analysis of tumors that may counter the notion that tumor cells need to suffer such radical transformations during metastasis. In order to help resolve these issues, a systematic investigation in tumor tissues of the co-expression of specific and robust markers of both EMT and tumor cells would be of great interest. From an experimental perspective, a shortcoming that prevents a better understanding of tumor cells undergoing EMT is the lack of appropriate tools and biomarkers for their identification and *ex vivo* isolation and characterization, similar to those that have allowed the relative purification of CSC (by cell sorting of immunolabeled or ALDH-expressing or autofluorescent cells, or through the formation of tumorsphere cultures or by growth in xenografts) [27].

Feeding into a skeptical view of EMT as a distinct process orchestrated by well-ordered switches in gene programs in the path to cancer progression, some studies suggest that this relationship may be a fallacy [134] arguing that the mesenchymal markers and properties acquired during tumor progression are simply a reflection of genomic instability. However, the vast majority of recent studies do not support this notion, in part, because it is unlikely that this complex coordinated program could be induced by random mutations as a result of genomic instability. On the contrary, it is more likely that genomic instability can induce changes in the expression of some important factors that regulate EMT, for example SNAI1.

These skeptical positions with regards to EMT are also partly fed by the fact that, when metastatic samples are analyzed, they do not always show a mesenchymal pattern of markers, and indeed they frequently resemble primary tumors. As described above, tumor cells need to acquire a series of properties in order to adapt and colonize distant tissues and accommodate to different niche interactions so as to achieve the formation of a secondary tumor. It should thus be plausible that they may also undergo a reversion from the mesenchymal phenotype imposed upon them at their primary sites, or a mesenchymal-epithelial transition (MET) [133], that would afford them to recover their epithelial identity and regain their original proliferative ability and thus their potential to grow tumors at secondary sites. In support of this view, histological analysis has revealed morphological similarities of primary tumors and metastatic lesions [135], and it has been reported that E-cadherin levels are elevated in lymph node metastases relative to matched primary tumor samples, both suggesting that EMT in primary tumors may be followed by MET at distant sites [108] [109]. Despite these correlative clinical

findings, rigorous functional studies linking MET with metastatic colonization ability are scarce. Moreover, cancer cells may, more often than not, pass through a partial EMT program rather than a complete one; such cells may concomitantly express epithelial and mesenchymal markers and this add difficulties to study what exactly is happening.

A separate issue concerns the relationships between EMT and the proliferative and self-renewal states of cancer cells. Although cancer cells are often considered as highly proliferative, less proliferation is often observed at the invasion front of carcinomas [138] [139]. Moreover, cells undergoing EMT during embryonic development, as well as tip cells during angiogenesis, stop dividing when migrating, suggesting that the cytoskeletal changes occurring during EMT may be incompatible with cell division [140]. Purified invasive cells are not proliferative [104]. Moreover, overexpression of the EMT-inducing transcription factor SNAIL induces cell cycle inhibitor P21 and represses cell cycle activator cyclin D [141]. Conversely, downregulation of cyclin D1 alone in breast cancer cells enhances migration and decreases proliferation [142].

5.2. EMT transcription factors and E-cadherin repression

The identification of Snail, ZEB and basic helix-loop-helix (bHLH) factors, as inducers of epithelial-mesenchymal transition (EMT) and potent repressors of E-cadherin expression, afforded a new understanding of one of the hallmarks of EMT. E-cadherin repressors are classified into two groups depending on their effects on the E-cadherin promoter. SNAI1/2, ZEBs, E47, and KLF8 bind to and repress the activity of the E-cadherin promoter [132] [143] whereas factors such as TWIST, Goosecoid, E2.2, and FOXC2 repress E-cadherin transcription indirectly [123] [144] (Figure 18).

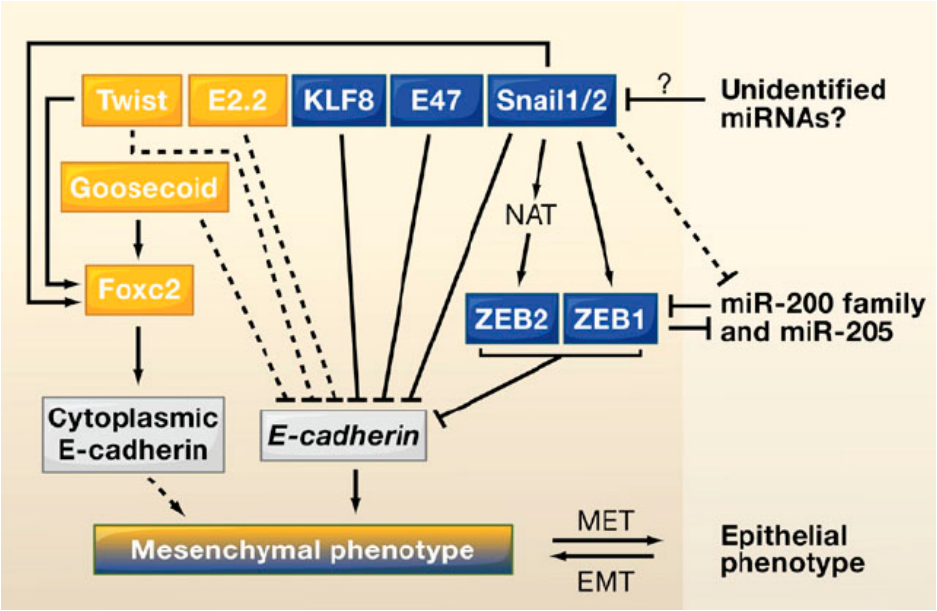


Figure 18. E-cadherin transcriptional repressors. SNAILs, ZEBs, E47, and KLF8 factors directly repress E-cadherin transcription whereas TWIST, Goosecoid, E2.2, and FOXC2 are indirect E-cadherin repressors. SNAI11 activates the expression of the ZEB genes by different mechanisms, including the

induction of a natural antisense transcript for ZEB2 (NAT). The miR-200 family and in some cases also miR-205, represses the transcription of ZEB genes preventing EMT. A loop of miRNAs and ZEB factors crossregulation plus the cooperation of several EMT inducers reinforces the control of the EMT process. Preliminary data indicate that SNAI1 may also repress the expression of the miR-200 family. Whether miRNAs can also control SNAI1 expression awaits further investigation. EMT, epithelial to mesenchymal transition; MET, mesenchymal to epithelial transition. Thiery *et al.* 2009.

A cooperative crosstalk occurs between these transcription factors, with the same final outcome, the downregulation of E-cadherin or the impairment of its functionality to coordinately induce the mesenchymal phenotype. SNAI1 induces SNAI2 during fibrosis, and both cooperate in primary tumor growth and in the focal distant site of metastasis [145]. SNAI1 is an activator of ZEB1 and can induce a natural antisense transcript of ZEB2 promoting its translation, thus upregulating ZEB1 and ZEB2. SNAI1, TWIST1 and Goosecoid are upstream effectors of FOXC2 which mediates mesenchymal differentiation and promotes the cytoplasmic localization of E-cadherin [146]. SNAI2 induction is required for TWIST1-induced epithelial-mesenchymal transition [147]. Moreover, they may share common upstream regulators such as Ppa, an E3 ubiquitin ligase that destabilizes TWIST1, SNAI1, SNAI2 and ZEB2 [148].

Different repressors bind with different affinities to the E-pal repressive element in the promoter regions of E-cadherin. For example, SNAI2 binds with lower affinity than SNAI1 and E47 proteins. These, together with the known expression patterns of these factors in embryonic development and carcinoma cell lines, support the idea that the *in vivo* action of the different factors in E-cadherin repression can be modulated by their relative concentrations as well as by specific cellular or tumor contexts [149].

5.2.1. The Snail family

The Snail family of transcription factors belongs to the so called Snail superfamily, which is divided into the Snail and Scratch families, both being a subgroup of the C_2H_2 zinc finger proteins. The origin of the Snail/Scracth family can be traced back to a proto-Snail gene that underwent tandem duplication in the last common ancestor of Dipoblasts and Bilateralia. The Snail family plays a fundamental role in embryonic development during gastrulation and in neural crest formation and differentiation in vertebrates. Scratch genes are expressed in the nervous system of all species analyzed, but their specific functions remain to be clarified. Snail proteins have been studied in better detail, partly because of their aberrant activation in adult humans leading to several known pathologies, including fibrosis, bone mineralization abnormalities and tumor progression [150]. To date, three members of the Snail family have been described in vertebrates: SNAI1, SNAI2 and SNAI3 [140].

Loss of E-cadherin expression has been considered a crucial step in the progression of papilloma to invasive carcinoma, since 1998 [151]. Importantly, two years later, the transcriptional repression exerted by SNAIL (SNAI1) on E-Cadherin (CDH1) in cancer was discovered, this providing new insights into the molecular

mechanisms of tumor invasion [104] [106]. Subsequently, other transcription factors were described as EMT inducers through the common mechanism of E-Cadherin repression to allow cell invasion. For example SLUG (SNAI2) is also involved in developmental EMT and cancer [140] [149]. Numerous studies have shown that the activity of these two important EMT drivers correlate significantly with disease relapse and survival in patients with breast, colorectal, and ovarian carcinoma, which indicates that EMT leads to poor clinical outcomes. Likewise, many studies have demonstrated that EMT profiles are associated with certain clinicopathological parameters, such as histological grades and tumor subtypes with worse outcomes, such as basal-like and metaplastic breast carcinoma [146].

The Snail proteins (SNAI1, SNAI2, SNAI3) are zinc-finger transcription factors that share a common organization: a highly conserved C-terminal region, containing four to six zinc fingers (C_2H_2) and a divergent N-terminal region (Figure 19). The zinc fingers function as sequence-specific DNA-binding domains that recognize consensus E-box elements C/A(CAGGTG) [152]. Snail factors are currently thought to be transcriptional repressors binding to E-boxes (5'-CACCTG-3') [131]. Their repressor function is dependent on the SNAG domain, constituted by 7-9 aminoacids on the N-terminal end of the proteins. The central region of the Snail proteins includes a serine-proline-rich region that is highly divergent between Snail members. Two different functional domains have been identified in the central region of SNAI1: a regulatory domain containing a nuclear export signal (NES) and a destruction box domain. The phosphorylation of proline/serine residues in both regions and a potential modification of adjacent lysine residues has been implicated in the subcellular localization of SNAI1, protein stability and repressor activity. SNAI2 contains the so-called slug domain in this central region, the function of which remains to be characterized.

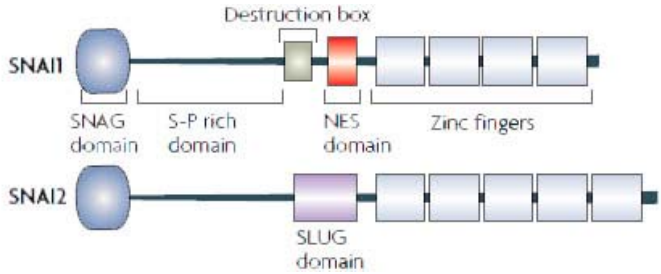


Figure 19. Schematic protein domains structure of SNAI1 and SNAI2. Both proteins contain structure similarities, mainly in the extremes. The C-terminal segment contains the zinc finger domains that confer recognition capacity of specific E-boxes in gene promoters. The central region of the SNAI2 contains a slug domain with no function described, while SNAI1 contains domains for proteasomal degradation and nuclear export. The N-terminal region contains the SNAG domain, which binds to cofactors that impose repressive marks over histones and DNA. Peinado *et al.* 2007.

In a hierarchical scheme of EMT inducers, SNAI1 has been proposed as the initial potent inducer at the onset of EMT and the migratory phenotype, thus being considered as an early marker of EMT that can induce other factors involved in the process. SNAI2, ZEB proteins, E47 and TWIST proteins are subsequently induced and participate in the maintenance of this mesenchymal migratory phenotype during metastasis. Interestingly, Snail genes are expressed in all EMT processes where they

have been studied. As mentioned above, EMT can be triggered by different signaling pathways, such as TGF-β, a potent inducer of Snail factors through AP1 and AP4 responsive elements for SMADS and LEF factors [146]. In agreement with the involvement of Snail in all studied processes of EMT, these signaling molecules have been shown to induce Snail genes in different cellular contexts [153]. This suggests that Snail genes are a convergence point in the induction of EMT in development, fibrosis or cancer (Figure 20).

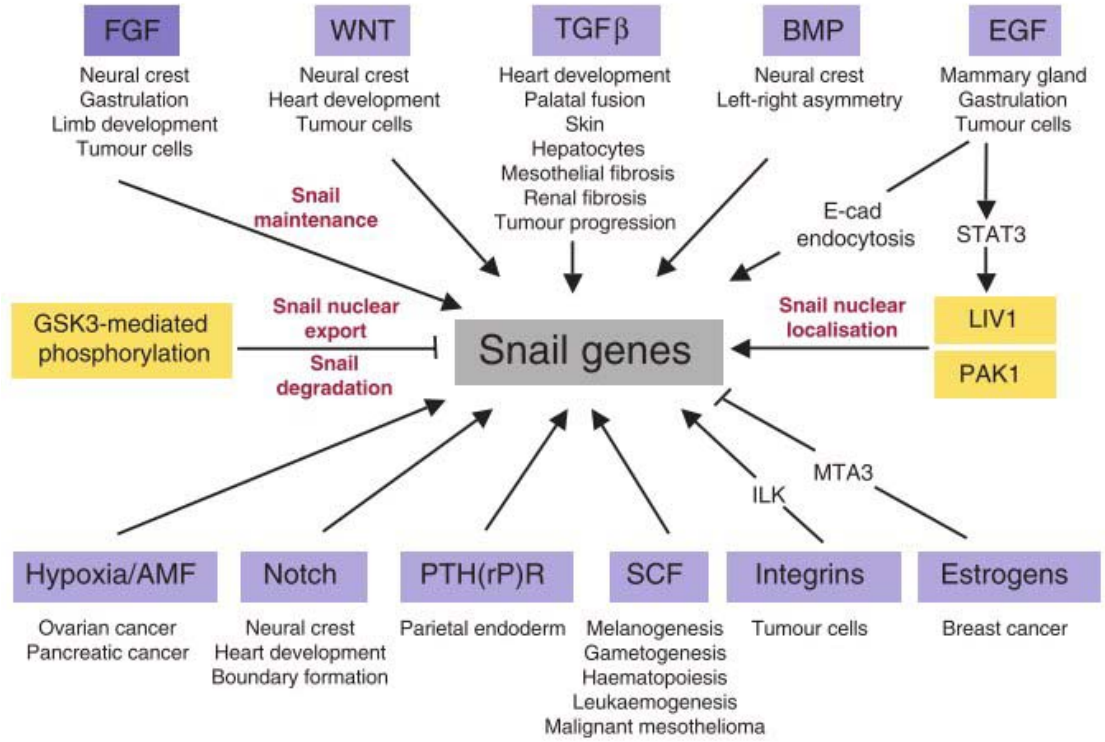


Figure 20. SNAIL genes, a central point of convergence in EMT induction of different biological processes: development, fibrosis and tumor progression. The figure schematically illustrates major molecular pathways and proteins involved in SNAIL induction, maintenance, nuclear transport and degradation. AMF, autocrine motility factor; E-cad, E-cadherin; EGF, epidermal growth factor; FGF, fibroblast growth factor; BMP, bone morphogenetic protein; ILK, integrin-linked kinase; MTA3, metastasis-associated protein 3; PAK1; p21-activated kinase; TGF-β, transforming growth factor β; PTH(rP)R, parathyroid hormone related peptide receptor; SCF, stem cell factor. Barrallo-Gimeno and Nieto, 2005.

Even though a major activity of SNAI1 is the repression exerted over E-cadherin, it binds to promoters other than E-cadherin and other genes are subjected to its repressive activity, including Cytokeratins 17-18, Desmoplakin, MUC1, Vitamin D receptor, Occludin, HNF4α, HNF1β, or EGR1 [129] [132] [154]. Furthermore, SNAI1 directly binds to its own promoter and thus controls its expression in a negative feedback loop [303]. Other epithelial molecules are downregulated by SNAI1 through post-transcriptional mechanisms like claudin-1, or ZO-1, and suggest that SNAI1 may act in translation initiation, thus evidencing its pleiotropic activities [155]. On the other hand, SNAI1 induce the expression of many mesenchymal genes, including Fibronectin, Vimentin, RHOB, LEF1 and importantly ZEB1, another E-cadherin repressor, although not directly [140] [154]. In, general, the action of SNAI1 results in a repression of epithelial genes, cell cycle progression and apoptosis genes.

Reciprocally, it promotes mesenchymal marker expression, cell cycle inhibitors and survival factors (Figure 21).

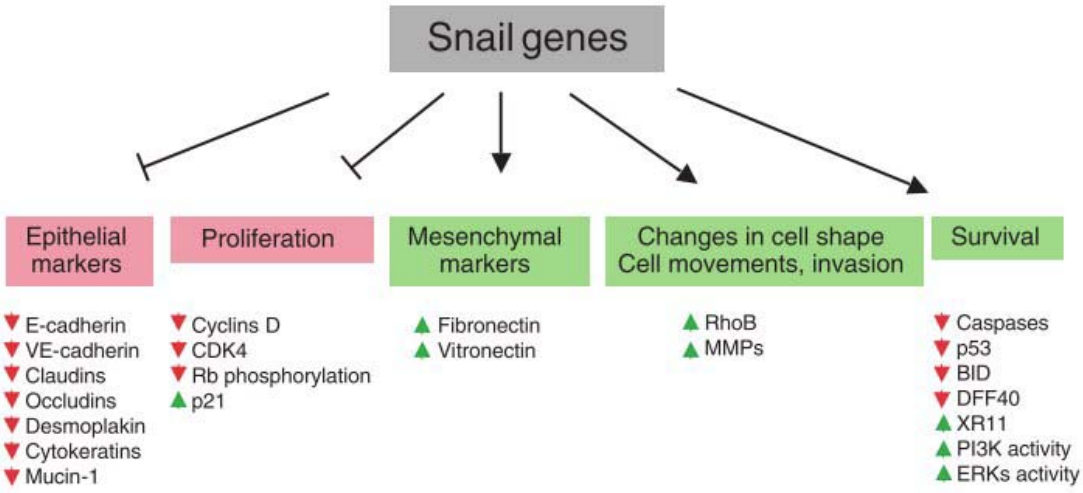


Figure 21. Downstream targets of SNAIL (SNAI1). SNAIL gene expression induces the loss of epithelial markers and the gain of mesenchymal markers, as well as inducing changes in cell shape, and changes related to morphology and to the acquisition of motility and invasive properties. The Snail genes also regulate cell proliferation and cell death. Not all of these targets are directly regulated by Snail genes: because Snail genes function as repressors, from Drosophila to humans, target upregulation might be due to the Snail-mediated repression of a repressor, or through other mechanisms (e.g., ncRNA regulation). Barrallo-Gimeno and Nieto, 2005.

The binding of Snail factors to E-box consensus sequences on the E-cadherin promoter is concominant with the recruitment of chromatin remodelers such the repressor CtBP, the SIN3A co-repressor complex, histone deacetylases HDAC1 and HDAC2, and components of the Polycomb 2 complex (PRC2) [156] [157]. It has been reported that EZH2, a PRC2 protein, mediates the silencing activity of E-cadherin through trimethylation of H3 lysine 27 at the E-cadherin promoter, where it is recruited by Snail [158]. Recently, the N-terminal SNAG domain of SNAI1 has been shown to interact with the amino oxidase domain of histone lysine-specific demethylase 1 (LSD1) resulting in a stabilization of a ternary SNAI1-LSD1-CoREST complex, preventing its sequestration to proteasome, and importantly the repression of specific target genes, such as E-cadherin [159]. LSD1 is found as a component of the HDAC1/2 containing co-repressor complexes CoREST and mi2/NuRD. Notably, the SNAG domain of SNAI1 resembles a histone H3-like structure and serves as a pseudosubstrate to recruit LSD1 to transcriptional target sites where it removes mono- and dimethyl marks on lysine 4 of histone H3 (H3K4) supporting the repression exerted by other SNAI1 cofactors over the E-cadherin promoter [160]. These results directly connect the transcriptional repressor functions of Snail with the epigenetic regulation of gene expression. Thus, when bound to the E-cadherin promoter, SNAI1 achieves transcriptional repression through a multistage sequence of events: Initially, HDAC1 and 2 within the CoREST-SNAI1-LSD1 ternary complex deacetylate histone H3 and H4, thus favoring repression. Second, PRC2 is recruited to promote direct trimethylation of H3K27, a repressive histone modification. Finally, LSD1 removes the activation marks on H3K4 supporting the repression cascade. (Figure 22).

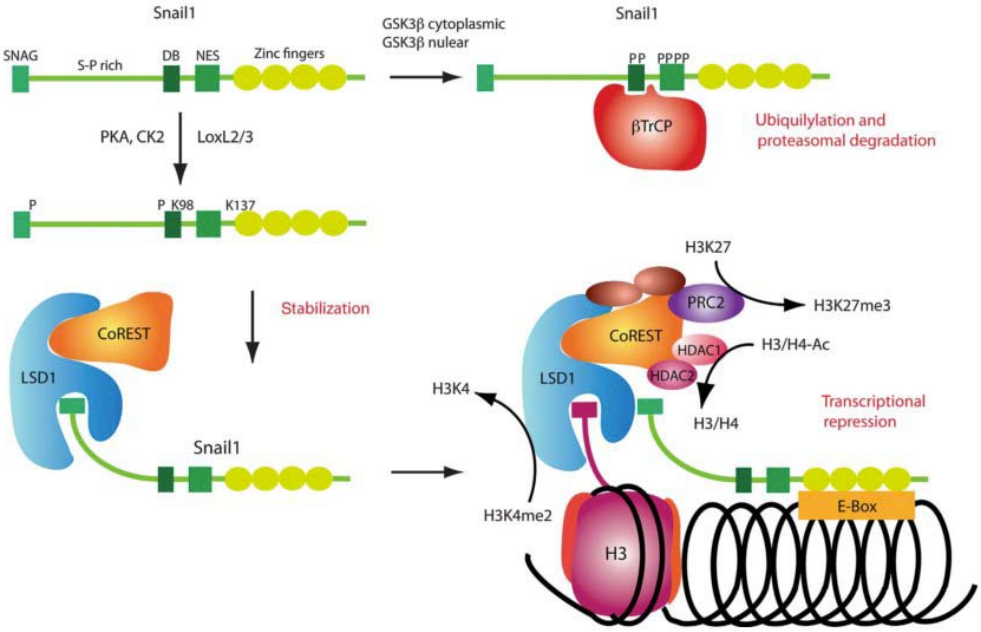


Figure 22. Regulatory mechanisms of SNAI1 and cofactors recruitment. SNAI1 is either phosphorylated by GSK3β in its central region and destined to ubiquitylation by the E3 ubiqutin ligase βTrCP and proteasomal degradation. Alternatively, SNAI1 is stabilized by phosphorylation of the SNAG domain through protein kinase A (PKA) and casein kinase 2 (CK2), and through its SNAG domain binds LSD1. The ternary SNAI1-LSD1-CoREST complex is recruited to E-boxes of SNAI1 target gene promoters, such as E-cadherin, in which LSD1 binds the tail of histone H3 and demethylates lysine 4 of histone H3 (H3K4). Histone deacetylases 1 and 2 (HDAC1/2) deacetylate histones 3 and 4 (H3/H4), and the polycomb repressor complex 2 (PRC2) trimethylates lysine 27 of histone 3 (H3K27). SNAIL-targeted chromatin modifications lead to a repression of the E-cadherin gene. Christofori, 2010.

In addition to being tightly regulated at the transcriptional level, Snail factors undergo posttranslational modifications that control their nuclear localization or degradation (Figure 22). Recent studies show the complex regulation of SNAI1 stability, subcellular localization and function through different phosphorylation events. These modifications include phosphorylation by PAK and GSK3β, dephosphorylation by the small C-terminal domain phosphatase (SCP), and lysine oxidation by LOXL2 [161] [162].

GSK3β (glycogen-synthase kinase-3β) phosphorylates two Ser residues on SNAI1, one of which targets SNAI1 for ubiquitylation and degradation, whereas the other promotes its nuclear export [162]. Mutations in SNAI1 that prevent GSK3β-mediated phosphorylation result in a stabilized form of SNAI1 that localizes in the nucleus and induces EMT. As expected, inhibition of GSK3β activity led to enhanced cellular levels of SNAI1 with concomitant downregulation of E-cadherin. The activity of GSK3β is inhibited by the AKT/PKB (protein kinase B), Wnt and Hedgehog pathways, each of which regulate EMT [163] [164].

PAK1 (p21-activated kinase-1) phosphorylates SNAI1 on a different Ser residue from GSK3β, which results in the accumulation of SNAI1 in the nucleus and subsequent SNAI1-mediated transcriptional repression of target genes. Mutation of the phosphorylation site on SNAI1 or knockdown of PAK1 expression resulted in

cytoplasmic accumulation of SNAI1 and ablation of its transcriptional-repressor activity. The EGF signaling pathway induces SNAI1 accumulation through PAK1 [165].

LOXL2 and LOXL3 increase SNAI1 stability through its interaction and modification of K98 and K137 residues of the SNAI1, which might prevent its degradation and nuclear export, therefore increasing its transcriptional activity and thus cooperating in EMT induction [166]. LOXL2 has been recently suggested as a new poor prognosis marker of cancer progression, specifically in squamous cell carcinomas [167]. SNAI1 protein stabilization is also promoted by NF-κB, which prevents its phosphorylation by GSK3β and subsequent degradation [168], whereas the formation of a ternary complex between wild-type P53, the ubiquitin ligase MDM2, and SNAI2 triggers its degradation [146]. The mechanism of stabilization of SNAI1 upon interaction with LSD1 to form the stable ternary complex SNAI1-LSD1-CoREST [159] has been referred to above.

Other post-transcriptional regulatory mechanisms are the cap-independent translation of SNAI1 mRNA, activated by the transcription/translation regulator YB-1 (Y-box binding protein 1), associated with breast cancer aggressiveness [169]. Another breast cancer-associated protein that increases SNAIL nuclear translocation is the zinc transporter LIV1. LIV1 induces EMT during zebrafish development [170] and promotes invasive properties in tumor cells [171].

5.2.2. Other transcription factors involved in EMT

The **ZEB family members**, ZEB1 (ZFHX1A/TCF8/δEF1) and ZEB2 (ZFHX1B/ SIP1), are E-cadherin transcriptional repressors that have been implicated in EMT, tumorigenesis and metastasis [172] [174]. ZEB1 and ZEB2 are two-handed zinc finger proteins that bind DNA binding sites composed of bipartite E-boxes (CACCT and CACCTG) [175]. The E-cadherin promoter contains such E-box sequences and it has been shown that ectopic expression of ZEB factors in mammary gland epithelial cells or MDCK cells is sufficient to induce dissociation of adherens junctions [176] [177] and enhance invasiveness/motility, respectively.

The ZEB factors can regulate the expression of various EMT- and tumor-related genes. For example, they have been shown to repress the expression of genes encoding proteins critical to maintaining the epithelial phenotype such as E-cadherin, Plakophilin 2 and ZO3 [177]. Conversely, the ZEB factors can also activate the expression of genes promoting migratory/invasive phenotypes, such as the pro-invasion gene MMP2 [178].

Although a great deal of information is available regarding the signaling pathways that regulate the expression of the EMT-promoting Snail superfamily members SNAI1 and SNAI2, little information is available regarding the specific regulation of the ZEB factors during tumor progression [146]. As has been previously mentioned, SNAI1 can activate ZEB factors [146] [154]. Thus, many signaling pathway can finally influence and activate ZEB factors through the action of SNAI1, such as TGF-β. Moreover, in addition to Snail proteins, another way to regulate the

expression of the ZEB factors and influence the epithelial phenotype is through miRNAs. In fact, a review of data from a number of studies suggests a negative correlation between the expression of miR-200s and that of the ZEB factors, suggesting miR-200 mediated targeting of ZEB factors in embryonic tissues [179] [180] [181]. Two reports published in 2007 linked the expression of the miR200 family of microRNAs to the maintenance of an epithelial phenotype in development and cancer through the targeting of ZEBs [182] [183]. ZEB factors contain the so-called "seed" sequence at the 3'-UTR region of its mRNA, specifically recognized by the miR200 family of miRNAs leading to the downregulation of its expression and translation [184–186]. These studies also report the existence of a negative feedback loop between ZEB factors and miR-200 family members, such that increased levels of ZEB induced by SNAI1 may repress the expression of these miRNAs. Furthermore, it has been shown that targeting of ZEB factors by miR200s induces a potent upregulation of E-cadherin and may revert EMT back to an epithelial phenotype, or MET [181] (Figure 23).

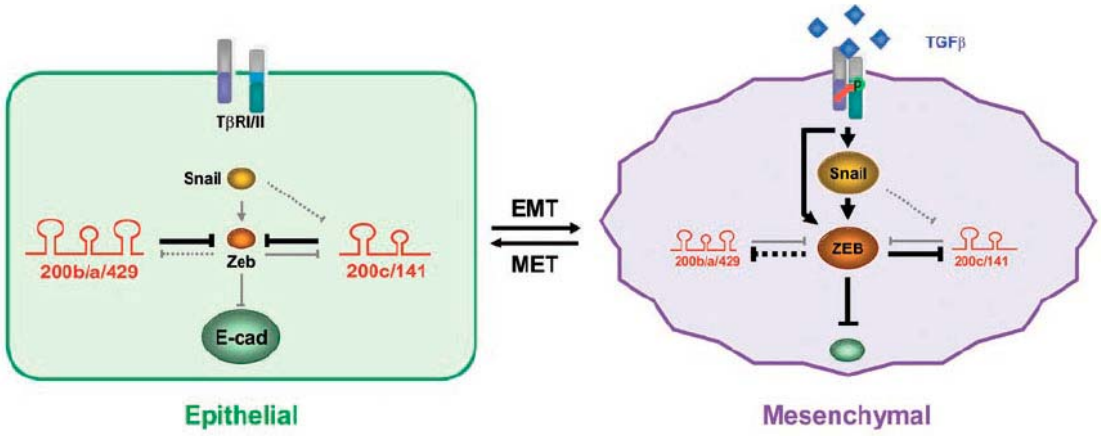


Figure 23. Regulation of ZEB factors and effects on cellular phenotypes. Epithelial cells express low levels of SNAIL and ZEB transcriptional factors and high levels of the miR-200 family miRNAs as two separate clusters. MiR-200 family miRNAs reduce the levels of ZEB factors. Conversely, ZEB1 can repress miR-200c/141 expression. High levels of miR-200 family and hence low levels of ZEB factors result in high expression of E-cadherin. Under the stimulation of certain cytokines, such as TGF**β**, the expression of Snail and ZEB factors is increased. ZEB factors inhibit the expression of miR-200c/141 and directly suppress E-cadherin expression to stimulate a morphological change to the mesenchymal phenotype. Thus, the balance between the levels and activities of miR-200 family and ZEB factors may determine the epithelial or mesenchymal states of tumor cells. Korpal and Kang, 2008.

Krüppel-like Factor 8 (KLF8) induces EMT and cell invasion [187]. It was identified as a transcriptional repressor of the Krüppel-like C_2H_2 zinc-finger transcription factor family of proteins and works as a downstream effector of focal adhesion kinase (FAK). KLF8 expression is promoted by Src and PI3K, but it is independent of SNAIL. In breast cell lines, KLF8 has been described as a potent inducer of EMT through a novel mode of repression of E-cadherin in epithelial cells. It directly binds to the E-cadherin promoter through GT boxes (CACCC). There is evidence that KLF8 may play a critical role in EMT-associated diseases, including breast cancer metastasis. Indeed, there is an inverse correlation between the expression levels of E-cadherin and KLF8 in lymph node-positive breast tumors [136].

Furthermore, KLF8 regulates MMP9 expression by transcriptional activation in human breast cancer cells to promote cell invasion [187].

TCF3 (EA2 or E12/E47) and **TCF4 (E2-2)** are members of the HLH transcriptional regulators, widely represented in most eukaryotic organisms. Most HLH proteins contain basic amino acids constituting the bHLH motif, through which they can bind to E-box sequences (CANNTG) in the promoters of different genes [189]. Both factors belong to the class I bHLH factors, together with TCF12 [188]. They are expressed in many tissues and are capable of forming either homodimers or heterodimers with tissue-specific class II bHLH factors, such as TWIST1 [188]. Class III of bHLH factors are MYC proteins but they do not dimerize with either class I or class II [191].

TCF3 (EA2) encodes the alternative splicing products E12 and E47. E12/E47 acts as a direct repressor of E-cadherin through specific binding to its promoter and triggers EMT in a manner similar to SNAIL or SLUG [189]. SNAI1, SNAI2 and E12/E47 induce common genetic programs but also factor-specific programs, playing different roles in tumor progression and invasion [190].

TCF4 (E2-2) mediates E-cadherin repression in an indirect manner, independent of proximal E-boxes in the E-cadherin promoter and probably involving the participation of additional repressors. Like EA2, E2-2 also encodes two highly related isoforms, E2-2A and E2-2B. E2-2 has been described as a common downstream target of the EMT regulators SNAI1, SNAI2 and E47, but it is dispensable for the EMT action of SNAI1 and E47 [190]. Nevertheless, genetic profiling and expression pattern analyses indicate that E2-2 products induce a genetic program distinct from that of E47, suggesting a non-redundant role of different bHLH factors in EMT and bringing forth an interesting interplay with other E-cadherin repressors in the regulation of EMT [190].

The **TWIST family** of factors belongs to the highly conserved, tissue-specific class II bHLH factors. These factors have the ability to heterodimerize with class I bHLH factors, such as TCF3 or TCF4, to bind to DNA E-boxes and they cannot form homodimers, with few exceptions. These heterodimers are able to bind to canonical and non-canonical E-boxes [192] and a CATATG hexanucleotide sequence known as NDE1 E-box, and activate or repress target genes [191]. In mammals, two Twist-like proteins, Twist1 and Twist2, share high structural homology [193]. Twist proteins play important regulatory functions during embryogenesis. Twist is induced to allow ventral furrow cells to migrate and during this process, these cells lose cell-cell junctions and undergo EMT [194]. Exogenous overexpression of TWIST1 increases the invasive and metastatic abilities of human cancer cells by promoting the downregulation of E-cadherin and the induction of an epithelial-mesenchymal transition (EMT) [195]. TWIST1 and TWIST2 have been shown to be expressed in different kind of cancers, including breast, liver, prostate, pancreatic or gastric cancer [191].

TWIST1 can form homodimers (T/T) or heterodimers with E12 (T/E) and that appear to have distinct activities and regulate the expression of different gene sets. Id (inhibitor of differentiation) proteins, a HLH protein without a basic domain, play a pivotal role between both form of dimers and determine their ratio since Id proteins dimerize with E12 (TCF3) and displace TWIST1, increasing the ratio of T/T homodimers with consequences in donwstream target genes [191]. Furthermore, while TWIST1 is stabilzed through its heterodimerization with TCF3, it is destabilized in heterodimers with Id proteins. Additionaly, it affects the expression of chromatin remodelling enzymes, such as, direct positive transcriptional regulation of the polycomb protein BMI-1, and it favors and promotes the self-renewal of certain stem-cell population by supressing let-7 and cooperatively contributes to E-cadherin downregulation [196]. The TWIST1 N-terminus can interact and inhibit histone acetyltranferases like p300, CBP and PCAF, thus repressing gene expression mediated by these histone remodelling enzymes [191]. TWIST1 C-terminus interacts with RUNX2 binding domain and represses its transcriptional activity for osteoblast differentiation or MMP expression. These findings suggest that TWIST not only regulate the direct transcription of target genes, but also modulates the activity of other transcription factors [191].

Both proteins, TWIST1 and TWIST2, override oncogene-induced premature senescence by abrogating key regulators of the P53- and RB-dependent pathways, and thus acquiring chemoresistance. Interestingly, in epithelial cells, the oncogenic cooperation between Twist proteins and activated mitogenic oncoproteins, such as RAS or ERBB2, leads to complete EMT. These findings suggest an unanticipated direct link between early escape from senescence programs and invasive features by cancer cells to future advance tumor progression [193]. Other studies involved TWIST proteins in the acquisition of stem cell properties by transformed mammary epithelial cells as a consequence of EMT [197].

The EMT inducers that indirectly repress E-cadherin transcription frequently activate some of the direct repressors, and they also have multiple specific targets. The TWIST proteins appear to repress E-cadherin indirectly, but Yang *et al.* [195] and others [198] [199] have reported direct binding of TWIST1 to the E-cadherin promoter. The assumption of an indirect effect TWIST1 as an E-cadherin repressor stems largely from an unreferenced statement in a review of Yang and Weinberg [123] in 2008. A more recent study, authors suggested indirect regulation based on a 7 days delay of E-cadherin downregulation upon TWIST1 activation, but they didin't demostrate it specifically [147]. Be it as it may, TWIST1 appears to require the induction of SNAI2 in order to suppress the epithelial branch of the EMT program [147]. SNAI2 is the specific factor responsible for the repression of E-cadherin in response to TWIST1 activation, and it has also been shown that TWIST1 directly binds to the SNAI2 promoter to induce its expression [147]. Moreover, TWIST1 and SNAI2 are frequently coexpressed in human breast tumors. Therefore, both factors cooperate together to promote EMT and tumor metastasis. In addition, SNAI2 also affects other epithelial genes and TWIST1 can induce other mesenchymal genes, such N-cadherin, through TCF3 heterodimerization [147] [188].

The signaling pathway and gene expression changes induced by TWIST in cancer cells are well studied. The overall data suggest that TWIST1 is a driver of EMT and tumor metastasis as supported by Yang *et al*. [195] in 2004. They suggested that TWIST1 is an inducer of metastasis via its repression of E-cadherin and induction of EMT. However, and intriguingly, the highly metastatic breast cancer cell line used in that study which expresses the highest levels of TWIST1 also expresses very high levels of E-cadherin, with correct membrane localization and displaying an epithelial-clustered morphology [201]. This may raises some questions on the precise role that TWIST1 might be playing in the metastatic behavior of tumor cells, and leads to consider other mechanism that may be regulated by TWIST1, perhaps depending on the balance of its levels among relative to those its parterns TCF3 and Id factors and chromating remodelling factors, independent of full canonical mesenchymal transformation.

Several pathways can induce the expression of TWIST, including HIF-1α signaling under hypoxic conditions [11], Wnt signaling, SRC activation [147], STAT3, MSX2 and NF-κB [191].

Goosecoid (GSC) is another EMT-related factor that indirectly downregulates E-cadherin [146]. This gene encodes a conserved transcription factor that was first identified as the most highly expressed homeobox gene in the Spemann organizer. Elements of the TGF-β superfamily and Wnt/β-catenin signaling pathways, which are known to be involved in tumor invasion, can induce GSC expression in embryonic cells and are required for Spemann organizer formation [200]. Goosecoid is expressed in human cancer cells, allowing such cells to acquire certain characteristics needed to overcome key barriers to tumor metastasis as in breast cancer tumors [178]. GSC induces EMT and cell motility, its expression causing a downregulation of epithelial markers such as E-cadherin, α-catenin, cytokeratins and an upregulation of vimentin and N-cadherin [177].

The **FOX (forkhead box)** proteins, a family of transcription factors that are important in regulating the expression of genes involved in cell growth, proliferation, differentiation and longevity. A feature of FOX proteins is the forkhead box, a sequence of 80-100 amino acids forming a DNA binding motif. The Fox family of transcription factors is expressed in various organs and tissues during development and is involved in a variety of developmental and cellular differentiation processes [202]. FOXC2, belonging to the “C” subfamily of Fox proteins, is required for cardiovascular development [203], early organogenesis of the kidney, podocyte differentiation and glomerular basement membrane maturation [204] [205].

In cancer, FOXC2 can act as an activator of EMT and metastasis in breast cancer. In addition, FOXC2 was found to be overexpressed in highly invasive and metastatic subtypes of breast cancer. In epithelial cells, FOXC2 overexpression resulted in phenotypic EMT with increased migratory and invasive behavior of epithelial cells and increases the metastatic potential of otherwise poorly metastatic breast cancer cells [206]. The fact that FOXC2 expression is induced by a large number of known regulators of the EMT program, notably the TWIST1, SNAI1, and

Goosecoid transcription factors as well as TGF-β1, suggests that FOXC2 is involved in a diverse array of EMT programs. The mechanism of action of FOXC2 to disrupt cell-cell junctions, differ from that of other EMT inducers. FOXC2 does not affect E-cadherin mRNA levels, but instead redirects the E-cadherin that continues to be synthesized from the plasma membrane to the cytoplasm, thus abrogating its functionality. It also has the ability to induce matrix metalloproteinases such, MMP2 and MMP9 to facilitate tumor cell invasion and angiogenesis [207]. Indeed, current evidence suggests that these other EMT regulators suppress the expression of epithelial genes and induce FOXC2 expression, delegating to FOXC2 the task of inducing the mesenchymal component of EMT programs [206].

5.2.3. Other intracellular molecules and external agents: novel EMT inducers

Pez, a tyrosine phosphatase, is induced by TGFβ, and its expression is sufficient to trigger EMT in MDCK cells through the induction of Snail and Zeb. Pez also induces the production of TGF-β, generating an autocrine activation loop [208].

PRL3, another tyrosine phosphatase, induces EMT in a colon carcinoma line by activating PI3K/ AKT. Stimulation of this pathway augments the degradation of PTEN and activates SNAI1 [209]. PRL3 also induced EMT through SRC activation in a kidney cell line [210].

The mucin Podoplanin triggers EMT in MDCK cells by activating RhoA [187]. By contrast, after inhibiting RHOA, podoplanin promotes collective MCF7 epithelial migration through the acquisition of filopodia and the loss of stress fibers. Podoplanin is expressed at the invasive front in an in vivo pancreatic tumor model, augmenting the frequency of high-grade tumors, although these invasive cells still retain E-cadherin expression [211].

The L1 cell adhesion molecule, a member of the immunoglobulin superfamily, induces EMT in epithelial breast carcinoma cell lines by promoting adherens junction breakdown and the nuclear localization of β-catenin [213].

Interleukin-related molecule (ILEI) can induce EMT and metastatic properties in various cell lines; ILEI is overexpressed in tumors, where it correlates with metastasis and poor survival [214]. Likewise, interleukin-6 (IL-6) also promotes EMT in breast cancer cells, and SNAI1 can induce IL-6 expression [215].

ECM proteases. SNAI1, ZEB and others induce the expression of metalloproteases that can degrade the basement membrane. Moreover, some of these proteases are able to induce EMT in a positive feedback manner that stabilizes EMT. Some examples are: MMP3, which triggers EMT by increasing reactive oxygen species, which in turns induces SNAI1 [216]. MMP13 and Eplysin trigger EMT after being induced by FGF1 and TGF-β, respectively [218] [217].

5.3. The mesenchymal phenotype involves more than just invasiveness

Regardless of the mechanism employed to repress E-cadherin, or to blunt its function, a large shift in gene programs is required, with a repression of epithelial genes and induction of mesenchymal genes, to attain a fully transformed mesenchymal phenotype from an epithelial cell (Figure 24). Thus, in addition, EMT inducers repress epithelial cell polarity and cell division while promoting cell survival [132] [146]. The attenuation of cell proliferation favors invasion at the expense of tumor growth. In addition, resistance to cell death confers a selective advantage on embryonic migratory or cancer invasive cells to populate distant organs. Thus, rather than being strictly repressors of E-cadherin expression, *bona-fide* EMT factors are global regulators of epithelial phenotype, cell adhesion and movement.

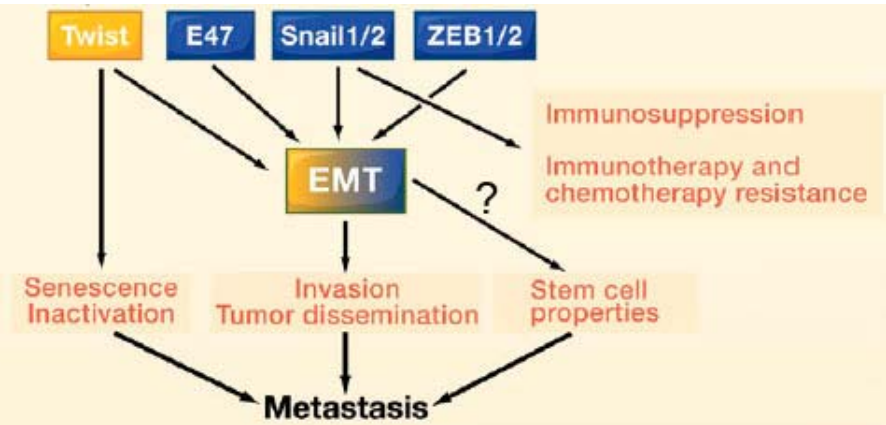


Figure 24. EMT phenotypic effects. . In addition to promoting invasion and tumor dissemination, EMT is involved in other aspects that influence the metastatic potential of tumor cells. Both TWIST and SNAIL may confer stem cells properties, favoring the self-renewal of a small population of cells that can colonize and differentiate into secondary carcinomas. In addition, TWIST also inactivates the cellular safeguard mechanism of cellular senescence triggered by oncogenes and SNAIL induces immunosuppression, immunoresistance, and chemoresistance. Modified from Thiery *et al.* 2009.

Loss of cell polarity in EMT. Epithelial cells have an apico-basal polarity, as orchestrated and maintained by 3 protein complexes: PAR, CRUMBS and SCRIBBLE. All three are regulated by EMT inducers [219]. SNAI1 represses CRUMBS3 and abolishes its localization at the cell junctions, together with PAR. ZEB1 represses CRUMBS3, PATJ and LGL2 (a member of the Scribble complex). Together, SNAI1 and ZEB1 downregulate PAR3 and the PAR6-mediated degradation of RHOA [146].

Evading programmed cell death and early senescence. Major EMT signaling pathways such as TGF-β signaling, can counteract apoptosis [220], depending on the cell context. TGF-β is a potent inducer of SNAI1, known to confer resistance to cell death [193]. TWIST1 and TWIST2 prevent cells from undergoing senescence induced by oncogenes by inhibiting P16/INK4A and P21/CIP [193]. Concomitantly, TWIST1 proteins cooperate with activated RAS to trigger full EMT and promote invasion. Zeb1 also protects mouse embryonic fibroblasts from senescence [221]. This suggests that abrogation of senescence may be a general mechanism associated with EMT. Consequently, those pathways activating Twist, Zeb or Snail will favor cell survival. This feature is a consistent acquisition of cells undergoing EMT, making

good biological sense, since invasive cancer cells travel across hostile environments to metastatize, and enhanced survival properties would at least optimize the likelihood of tumor cells reaching sites of micrometastasis. In turn, if the traveling cells finally succumb and enter senescence, they would acquire a secretory phenotype, which entails a further induction of EMT and the preparation of the secondary microenvironment as a presumptive metastatic niche partly through the release of inflammatory cytokines, immune modulators and growth factors [222].

Resistance to chemotherapy. It has been shown that cells that undergo EMT gain resistance to chemotherapeutic agents. Many antitumoral agents are designed to abrogate the proliferative machinery most active in the highly replicating compartments of tumors, ultimately resulting in the selection of tumor cell subpopulations with a low proliferative index. Since EMT tends to suppress proliferation, cells that undergo EMT will naturally tend to be more resistant to drugs targeting the replication machinery. In addition, EMT endows cells with mechanisms to evade apoptosis triggered by toxic agents or other insults. If these highly invasive and poorly proliferating cell are to be targeted, new drug discovery and development schemes need to be implemented specifically aimed at these tumor cell subpopulations, by targeting specific EMT inducers and mediators in order to eradicate mesenchymal tumor populations and prevent their dissemination. In this regard, mesenchymal cells have shown resistance to a variety of agents crurently in use in cancer therapy, includiing oxaliplatin, paclitaxel, adramicyn or radiation [146].

Immunoresisteance. The induction of EMT has been shown to facilitate the escape of tumor cells from immune surveillance. Tumors evade both natural and pharmacologically induced (e.g., vaccines) immunity by a variety of mechanisms, including induction of tolerance and immunoediting. Immunoediting results in reshaping the immunogenicity of the tumor, which can be accompanied by loss of antigen expression and major histocompatibility complex molecules. Neu-driven tumors undergo EMT to escape from the immune system [223]. SNAI1 expression is correlated with breast tumor recurrence [224] and Snail is associated with the activation of immunosuppressive cytokines, regulatory T cells, cytotoxic T lymphocytes resistance, and the generation of impaired dendritic cells [225]. Thus, SNAI1 and very likely the EMT process in general, can foster cancer metastasis by acting on multiple immunosuppression and immunoresistance mechanisms.

EMT confers stem cell properties. Recent evidences suggests that cells that undergo EMT acquire stem cell-like properties [197] [226–228]. But this intriguing concept is supported by some studies while opposed by other evidences [27] [61] [68] [229] [230]. By fusing together the CSC hypothesis with EMT, one attains an attractive unifying theory in which self-renewal and the capacity to invade are tightly linked into producing a single major cell type driving both primary tumor growth and metastasis. Whether this apparently simple hypothesis explains all relevant theoretical challenges and experimental evidences will be discussed in more detail below.

6. Mechanisms of cancer cell invasion

Cancer invasion is a cell- and tissue-driven process in which the physical, cellular, and molecular determinants adapt and react throughout the progression of the disease. Cancer invasion is initiated and maintained by signaling pathways that control cytoskeletal dynamics in tumor cells and the turnover of cell-matrix and cell-cell junctions, followed by cell migration into the adjacent tissue. Subsequently, the invading tumor cells engage with blood and lymph vessels, penetrate basement membranes and endothelial walls, and disseminate through the vessel lumen to colonize distant organs [231]. This deterministic process depends on the modulation of cell-matrix and cell-cell adhesion, protease, and cytokine systems (Figure 25).

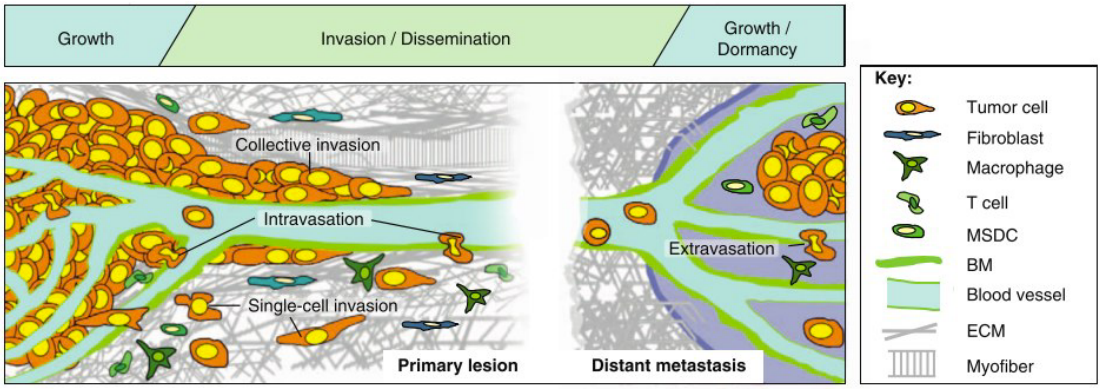


Figure 25. Steps in tumor invasion. At the primary site, tumor growth and single cell or collective invasion leads to the repositioning of tumor cells and intravasation into local vessels. After homing in secondary organs, tumor cells extravasate and, often after a phase of growth arrest and persistence (dormancy), regrow to macroscopic metastases. In both primary tumor and metastasis, reactive stromal cells such as fibroblasts, macrophages, myeloid-derived suppressor cells and regulatory T cells (Treg) contribute to cell growth, survival and invasion by the release of soluble factors and extracellular matrix remodeling. Alexander and Friedl, 2012.

Attempts to define the mechanisms that govern invasive and metastatic cancer cell migration, such as dominant signaling pathways, receptor-ligand interactions, or protease-substrate interactions, have largely failed. Instead, cancer cell invasion is now regarded as a heterogeneous and adaptive process. Indeed, it is this “plasticity” in cell adhesion, cytoskeletal dynamics and mechanotransduction that perpetuates migration and dissemination under diverse structural, molecular, and even adverse microenvironmental conditions [232] [233]. Reciprocal reprogramming of both the tumor cells and the surrounding tissue structures (reactive tumor stroma) not only guides invasion, but also generates diverse modes of dissemination. The resulting ‘‘plasticity’’ contributes to the generation of diverse cancer invasion routes and programs, enhanced tumor heterogeneity, and ultimately sustained metastatic dissemination.

Cancer invasion is a cyclic process in which the cell changes shape, produces morphological asymmetry, and then translocates the cell body. Depending on the cell type and tissue environment, cells can migrate in two major ways: individually (amoeboid or mesenchymal), when cell-cell junctions are absent, or collectively as cohesive multicellular groups, when cell-cell adhesions are retained (Figure 26)

[231]. The underlying process in both types of migration is the dynamics of the cytoskeleton coupling with cell surface receptors that engage with surrounding tissue structures; thus, the cytoskeleton serves as the cell's engine, and the cell surface receptors act as its transmission [234]. Cancer cells recapitulate the types and mechanisms of migration used by normal, non-tumoral cells. They activate the same machineries for changing shape, generating force, and remodeling ECM [235] as normal cells do, but neoplastic cells lack physiological ''stop signals'' immobilizing and anchoring the cells [236], which arguably perpetuates neoplastic cell migration.

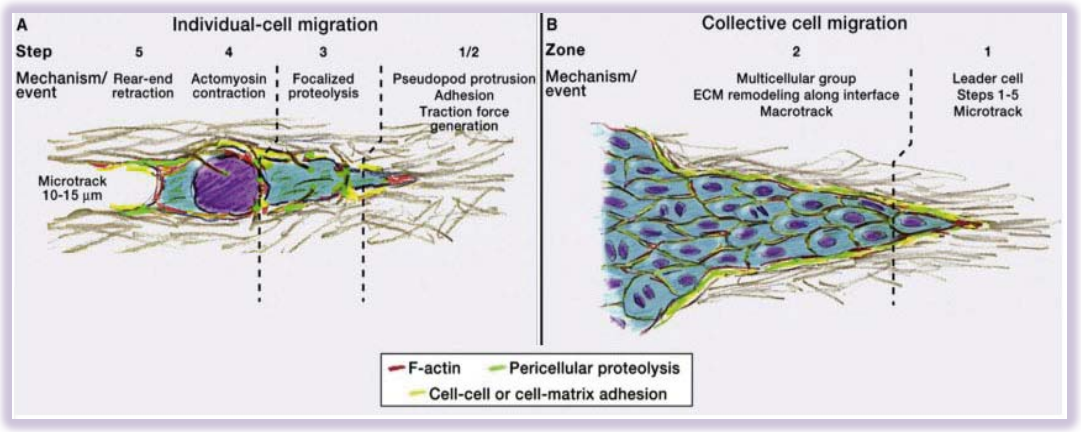
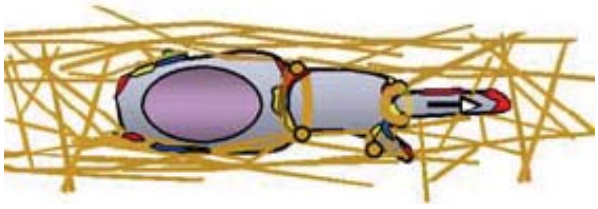


Figure 26. Individual and collective cell migration. A) Single-cell migration involves five molecular steps that change the cell shape, its position, and the tissue structure through which it migrates. After protrusion of an anterior pseudopod (step I) and traction force generation (step II), focal cleavage of individual ECM fibers is executed slightly backward to leading adhesion sites (step III). Transport of loose fiber ends (step IV) results in a small microtrack detected upon forward gliding of the cell rear (step V). B) Collectively migrating cells form two major zones: zone 1, in which a "leader cell" generates a proteolytic microtrack at the front of the migrating group, and zone 2, in which the subsequent cells then widen this microtrack to form a larger macrotrack. ECM macropatterning is executed by multiple cells that collectively fill a small pre-existing tissue gap while remaining connected. By focalizing proteolytic activity toward the cell-ECM interface, a near-continuous ECM layer is generated and further cleaved. Friedl and Wolf, 2011.

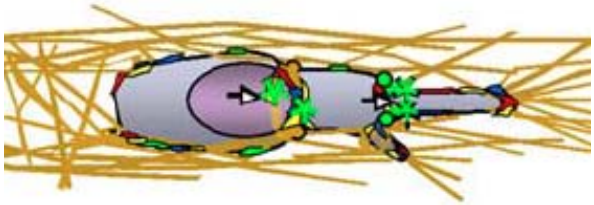
The physicochemical steps in **single-cell migration** are coordinated within the same cell body and executed in a synchronous, often pulsatile manner, which allows the cell body to protrude and generate traction in an oscillatory manner [234]. The invasive single-cell migration results from five interdependent molecular steps that change the cell shape, its position, and the tissue structure through which it migrates [236] (Figure 24A). The five-step model of cell migration is active in many types of cell movement for both normal and neoplastic single cells:

- Step 1, pseudopod elongation: the cytoskeleton polarizes by actin polymerization and forms a leading protrusion at the opposite end of a ''pre-uropod'' region, which marks the constitutive rear end of the cell [237].

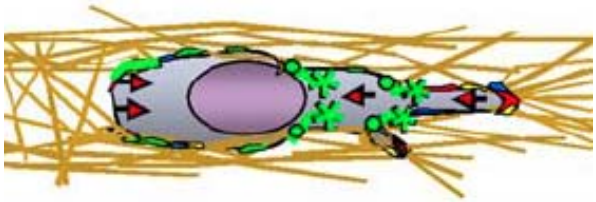


- Step 2, adhesion and force generation: the leading edge protrusion engages with extracellular substrates, followed by recruitment and adhesion of cell surface receptors that form focalized clusters and couple extracellular adhesion to intracellular mechanosignaling and force generation [236]. In most cells, the leading edge protrusion is controlled by the small GTPase Rac or Cdc42, which generate pseudopodia or filopodia that engage with ECM substrate.

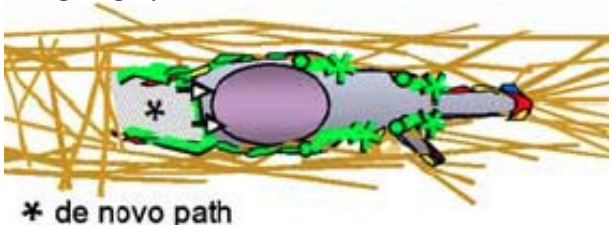
- Step 3, focalized proteolysis: several micrometer rearward of the leading edge, cell surface proteases become engaged with extracellular scaffold proteins and execute locally controlled proteolysis [236]. This proteolysis modifies the molecular and mechanical tissue properties and allows space for the advancing cell body.



- Step 4, acto-myosin contraction: the small GTPase Rho activates myosin II, and contraction mediated by actomyosin generates tension inside the cell.



- Step 5, rear retraction and path release: this contraction is followed by the gradual turnover of adhesion bonds at the trailing edge, which slides forward while the leading edge protrudes further [237].



Regarding **collective cell migration**: If multiple cells originate from the same location, such as a tumor, the ‘‘leader cell’’ forms a proteolytic microtrack of locally removed ECM barriers (Figure 27 B, zone 1). The following cells then widen and/or excavate this microtrack by mechanical force and proteolysis to form a larger macrotrack [238] [244] (Figure 27 B, zone 2). In collective migration, protrusion and retraction are coordinated in a supracellular manner, in which cytoskeletal protrusion and contractility are mechanically mediated through cell-cell junctions

allowing the cell group to behave as ‘‘mega-cell’’ [232] [239]. In this manner, collective invasion involves nodules of cancer cells advancing en masse into adjacent tissues, a mode of invasion characteristic of, for example, squamous cell carcinomas (Figure 27). Interestingly, such cancers are rarely metastatic, suggesting that this form of invasion lacks certain functional attributes that facilitate metastasis [2].

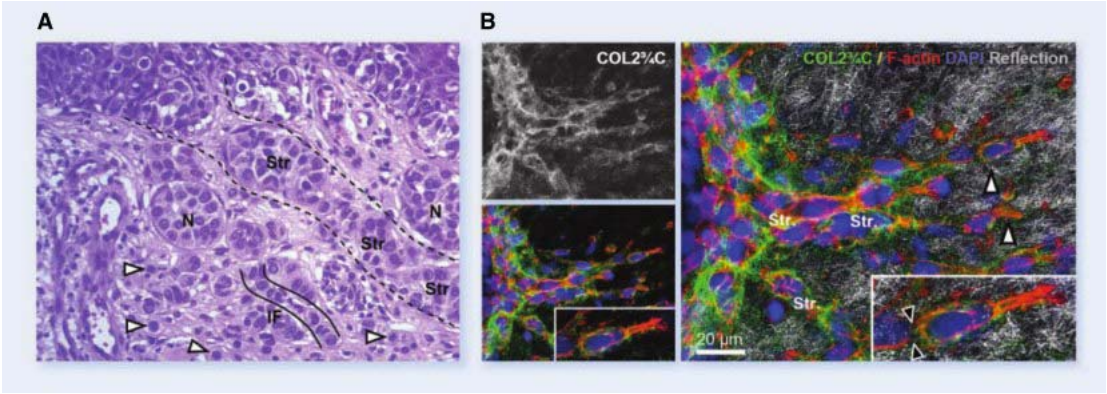


Figure 27. Individual and collective tissue invasions are mediated by two distinct types of pericellular proteolysis. A) Different invasion programs in primary melanoma invading the mid-dermis *in vivo*, including scattered individual cells (arrowheads), multicellular solid stands (Str), nests (N) representing cross-sectioned strands, and single cell chains (IF, “Indian files”). H&E staining. B) Microscopic fluorescence image. Transition from individual to collective invasion from three-dimensional spheroid cultured within a three-dimensional collagen lattice. Single cells (arrowheads) generate small proteolytic tracks (detected by cleavage-site specific COL2 3/4C antibody) that become further remodeled and widened by multicellular strands (Str). Yellow, proteases and MT1-MMP; green, degraded collagen; blue, h1 integrin; red, Filamentous-actin. Modified from Friedl and Wolf, 2008.

6.1. Subtypes of invasion modes, based on individual and collective cell migration/invasion.

The modes of cell migration were originally classified based on the morphology of migration patterns. This terminology was then extended to include molecular parameters, such as cytoskeletal organization, the type of cell–matrix interaction and force generation, and the modification of the tissue structure imposed by migrating cells [232]. Within the two major categories of invasion modes, individual or collective, their subcategories are classified according to morphological and mechanistic features (Figure 28):

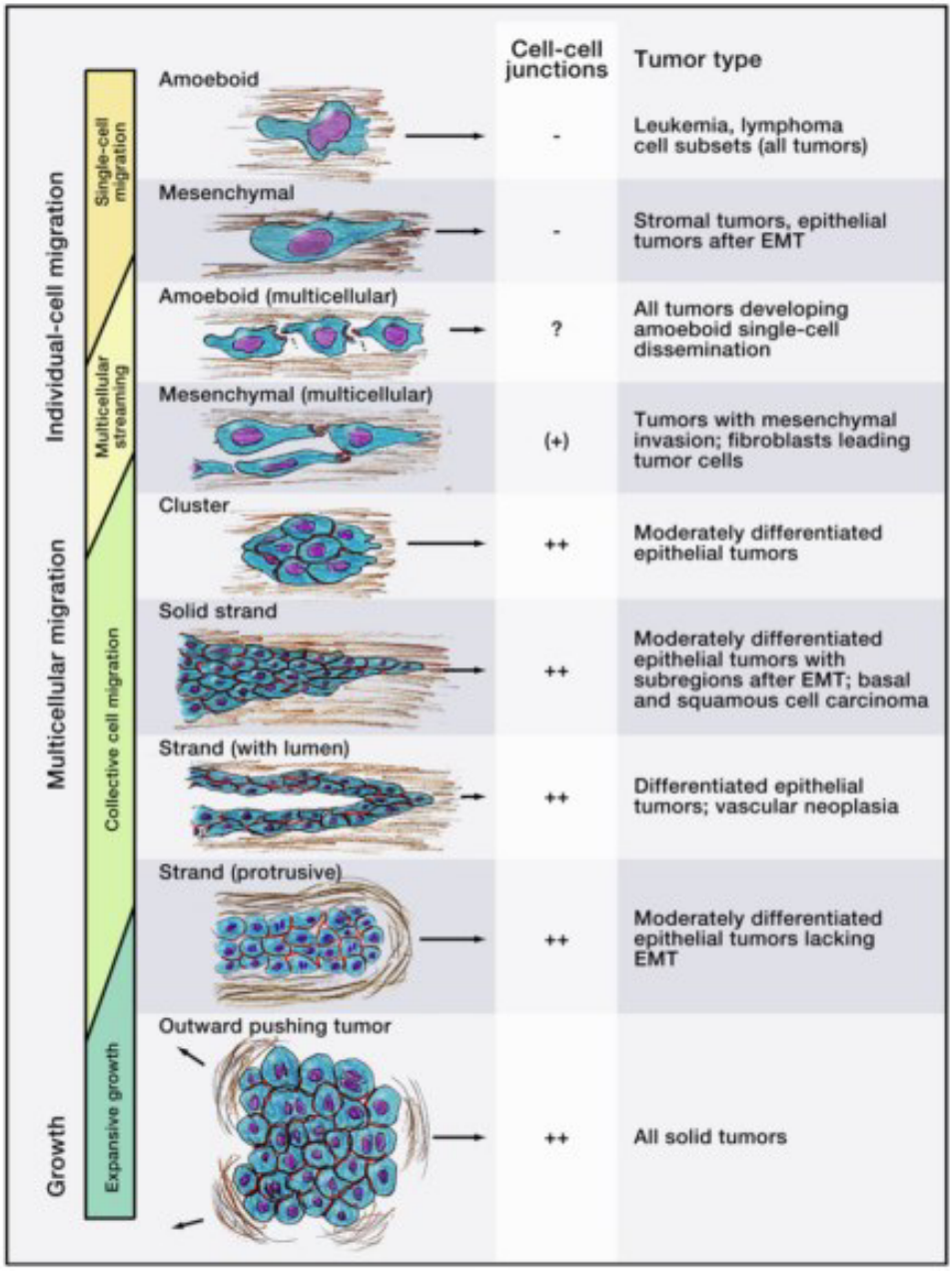


Figure 28. Modes of cell movement involved in cancer invasion and metastasis. Single-cell and collective cell migration can be further partitioned based on the specific cell-cell junctions, the contractility of cytoskeleton, and the turnover of cell attachments to extracellular matrix (ECM). These modes of migration can be further unstable and change upon alterations of cell-cell interactions, cell-ECM adhesion, or cytoskeletal contractility, resulting in intermediate phenotypes. Friedl and Alexander, 2011.

Amoeboid invasion. This commonly refers to the movement of individual rounded or ellipsoid cells that lack mature focal adhesions and stress fibers. There are two subtypes of amoeboid movement. The first is the rounded, blebby migration of cells that do not adhere or pull on substrate but rather use a propulsive, pushing migration mode. The second occurs in slightly more elongated amoeboid cells that generate actin-rich filopodia at the leading edge that engage in poorly defined, weak adhesive interaction with the substrate [240].

Amoeboid cells tend to migrate in the absence of proteolytic ECM breakdown by adapting their shape to and squeezing through tissue gaps and trails [240]. The origin of amoeboid tumors is often hematopoietic or neuroectodermal, including leukemias, lymphomas, and small cell lung carcinoma, but amoeboid movements are also detected as cell subsets in most of tumor types [240]. Amoeboid migration in a multicellular streaming model is also possible, as in peripheral connective tissue.

Mesenchymal invasion. In this invasion submode, individual cells have cytoskeletal protrusions and adhesion capabilities are strongly developed, and invading cells adopt spindle-shaped, elongated morphologies with focalized cell-matrix adhesions containing integrin clusters and proteolytic activity toward ECM substrates. Focalized proteases on the cell's surface generate small microtracks through which subsequent cells can follow (Figures 26A and 28). These microtracks formed by a single leading cell, following a source of chemoattractants, allow the cell to move one after another in a multicellular streaming manner [241] (Figure 27A and 28).

Mesenchymal migrating tumor cells originate from connective tissue tumors, including soft tissue sarcomas. They also originate from all other tumor types after epithelial tumor cells have undergone EMT (described in the previous section) and lose cell-cell junctions [237].

Collective invasion. Collective invasion requires cell-cell adhesion and multicellular coordination to occur simultaneously with migration, which results in multicellular groups and strands originating at the interface between tumor and stroma. Collective invasion may adopt different morphologies, which depend on the cell type, the number of jointly moving cells, and the tissue structure being invaded. For instance, groups of cells can form small clusters, solid strands, or files; if epithelial polarity is retained during migration, these structures can even form an inner lumen [242].

EMT dependence. In most cases of collective cancer invasion, one or several leader cells with mesenchymal characteristics form the tip of multicellular strands and generate forward traction, pulling the rest of cells and pericellular proteolysis toward the tissue structure [139]. Indeed, a variant of this modality is the cooperative action exerted by tumor stromal fibroblasts, also called CAFs (cancer-associated fibroblasts), which lead the invasive tip followed by the group of carcinoma cells in squamous cell carcinoma (SCC) [243] (Figure 29).

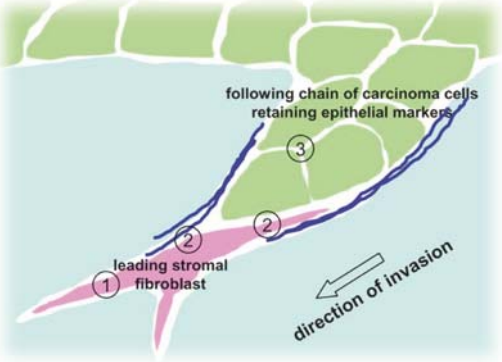


Figure 29. Model of fibroblast-led collective SCC invasion. (1) Initially fibroblasts invade in an MMP-dependent but RHO-ROCK independent manner. (2) A track is generated in the matrix by the leading fibroblast. The generation of this depends on the function of Integrin α 5 > RHO > ROCK > MLC signaling, integrin α 3 and MMP's. (3) SCC cells are unable to remodel the matrix but can utilize tracks in the matrix generated by fibroblasts and as a

result follow in the tracks behind the fibroblast. Gagglioli *et al.* 2007.

In a second type of collective invasion, EMT-independent, a blunt bud-like tip protrudes along tissue spaces consisting of multiple cells that variably change position, lacking a defined leader cell. This type of invasion occurs preferentially in soft tissues and cells of strong epithelial polarity. Collective migration is prevalent in morphogenesis during development and recapitulated in most epithelial and mesenchymal tumor types [139].

Collective cell migration is essential in building, shaping, and remodeling complex tissues and tissue compartments, such as epithelia, ducts, glands, and vessels, but also contributes to cancer progression by local invasion. It is the prevalent invasion mode in morphogenesis during development and recapitulated in most epithelial and mesenchymal tumor types. In contrast, single-cell migration allows cells either to cover local distances and integrate into tissues, such as neural crest cell migration, or to move from one location in the body to another and fulfill effector functions, such as immune cell trafficking or during cancer metastasis to distant sites [231]. Furthermore, solitary, presumably EMT cells (5%), in carcinomas move much more rapidly *in vivo* than the collective clusters, and, induced by TGF-β, intravasate into the blood vessels, whereas clusters rather invade better the lymph nodes [29].

Expansive growth invasion. Operationally, individual cell and multicellular migration follow the paradigm of active cell migration, whereas multicellular growth leads to passive cell movement by pushing. Some surrounding tissues do not impose a strong physical confinement on proliferating tumor cells. When tumor cells grow into these tissues, the increase in volume leads to multicellular outward pushing with intact cell-cell junctions and no signs of active migration. Expansive growth may displace cells by volume expansion and pushing when migration activity is absent. However, it can be coupled with migration tips in the edges, which then importantly contributes to enhanced collective invasion [244].

Cells often display multiple modes of migration in three-dimensional (3D) tissue models. The models described above are informative with regards to the multiple possibilities used by cells to migrate and invade, but these modes are not mutually exclusive. A tumor bulk can execute all these mechanisms at the same time at different localizations or cell populations of the tumor, or different modes can be used by the same cell population at different times. This includes intermediate or transition states in which cells may change their molecular profiles and switch migration modes (e.g., from proteolytic to non-proteolytic migration or single-cell to collective migration) [231]. For example, when individual cells become attracted by the same chemotactic source, they may first undergo multicellular streaming with short-lived cell-cell junctions that briefly form and resolve again; when cell-cell adhesion molecules are then upregulated, the cells may join each other and convert to a collective migration mode [232]. Thus, diverse molecular programs jointly determine the morphology and mechanism by which normal and neoplastic cells move through tissues.

6.2. Cell adhesion molecules and invasion signaling

The molecular mechanisms that induce different types of migration and/or invasion are different depending on the coordination and strength of the particular mode of migration. These variables include: cell-cell adhesion; cell-ECM interactions, cytoskeleton rearrangements; tissue remodeling structures; turnover of cell-substrate adhesions; and intercellular communication and cooperation. Stromal-tumor interactions are also key modulators of invasion, and stromal signals, such as chemokines, cytokines, growth factors and physical interactions, also modulate the signaling status of the invasive cells.

ECM receptors. Integrins are heterodimeric cell-surface receptors that consist of two transmembrane subunits, α and β, which form distinct integrin sub-types linking extracellular matrix (ECM) ligands, such as fibronectin, vitronectin, laminin and collagen, to the intracellular actin cytoskeleton. Importantly, this binding, in turn, induces intracellular signaling cascades that modulate motility and other cell mechanism such as cell proliferation, survival, polarity and differentiation. To detach and migrate, tumor cells depend on changes not only in cell-cell, but also in cell-matrix, interactions. Therefore, this needs to be resolved, whereas transient and weak adhesions are a prerequisite for migration. After associating with ligands, the cytoplasmic tails of integrins connect to cytoskeletal adaptor proteins, including talin, paxillin, and kindlin and the mechanosensing modulators vinculin and p130Cas. This adaptor and modulator proteins engage with the actin cytoskeleton and trigger signaling to protein kinases, including focal adhesion kinase (FAK) and SRC [245]. FAK is a non-receptor tyrosine kinase that plays a central role in cell migration. FAK has been shown to coordinate lamellipodial formation and focal adhesion turnover in fibroblast and interacts with SRC to improve the invasion [246]. Downstream integrin effectors further include the small GTPases RAC and RHO, which reinforce cell protrusion and rear contraction [233]. In addition to contact to ECM substrate, integrin engagement with extracellular ligands is also activated by inside-out signaling through RAC, the RAS-related GTPase RAP1, and Talin [233]. Integrins also contribute to tumor cell invasion by regulating the localization and activity of matrix-degrading proteases, such as matrix metalloprotease 2 (MMP2) and urokinase-type plasminogen activator (uPA), favoring EMT as well as through activation of TGF-β1 by integrin αvβ6 in tumor cells [247].

Many cell surface receptors contribute to invasion upon binding to ECM proteins and cooperative heterodimerization with other growth factors receptors and integrins. CD44, a hyaluronan receptor (also called hyaladherins), binds to hyaluronic acid to induce invasiveness. It also binds, with less affinity, to heparan sulfate, collagen, and fibronectin. CD44 connects to the actin cytoskeleton by the adaptor proteins ezrin, radixin, and moesin (ERM) and ankyrin and mediates intracellular signaling through SRC kinase and small RHO GTPases, including RHOA [248]. To promote invasion it also acts as a coreceptor of growth factor receptors, such as c-MET, FGFR1, EGFR and variants. It also binds to podoplanin to induce RHOA and thus invasion [248]. Interestingly, a splice variant switch from CD44s to CD44v is required to undergo EMT, inducing an entire program to favor invasion

[249]. Mechanistically, the splicing factor epithelial splicing regulatory protein 1 (ESRP1) controls the CD44 isoform switch and is critical for regulating the EMT phenotype. Additionally, the CD44s isoform activates Akt signaling, providing a mechanistic link to a key pathway that drives EMT [334].

Cell surface proteoglycans (syndecans, glyoicans and neurophilin) also can interact with ECM proteins and cooperate with integrins and growth factor receptors to deliver its signaling through PKC and RC. DDR1 and DDR2, discoidin domain receptors, interacts with fibrillar collagen and transmit signals through STAT5, NF-kB and P38 MAPK/ERK or SRC related kinases. They also increase proteolytical functions via MMPs [232].

Integrins are thus the main adhesion and mechanotransduction system for interstitial migration, with modulatory input and crosstalk to alternative cell-ECM and growth factor signaling systems through CD44, cell surface proteoglycans, and DDRs (Figure 30).

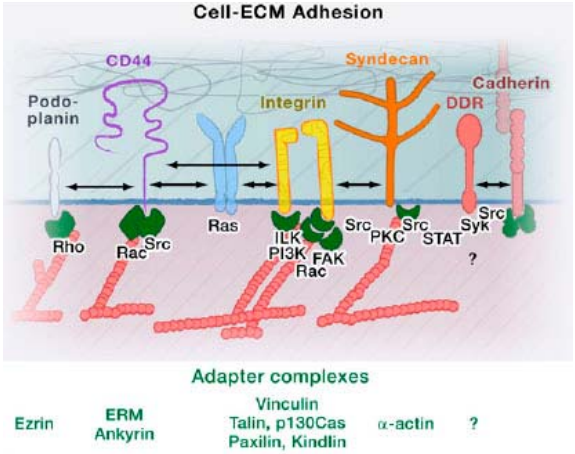


Figure 30. Molecular determinants of cell migration. Cell surface receptors and adaptors that mediate the dynamic interface between the actin cytoskeleton and pro-migratory signaling and the extracellular matrix (ECM). Friedl and Wolf, 2011

Syndecans are a family of transmembrane heparan sulphate proteoglycans with four members, syndecans 1 to 4. Syndecans function mainly as co-receptors by binding to their ECM ligands in conjunction with other receptors, notably integrins. Through their heparan sulphate side chains, syndecans may further engage directly in ligand binding [250].

Dystroglycans are heterodimeric complexes consisting of non-covalently associated α and β subunits with extracellular ligand-binding and transmembrane functions, respectively. Dystroglycans are a part of the larger dystrophin-associated protein (DAP) complex that connects basement membranes to the cytoskeleton, particularly via α2 laminins and perlecan [250].

Cell-cell adhesion receptors. This class of receptors transmit cell-ell adhesion forces toward the actin cytoeskeleton and thus provide cooperation between tumor cells favoring the collective invasion [139]. They also support single cell and collective movement over other cell surfaces. These receptors include cadherins,

CAMs (Immunoglobulin cell adhesion molecules), ephrin receptors and gap junctions (Figure 31).

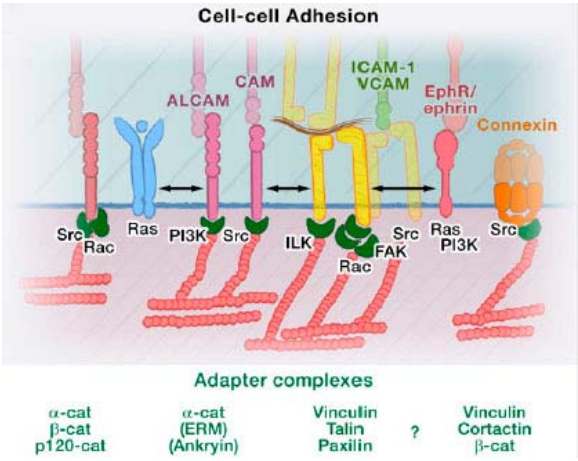


Figure 31. Molecular determinants of cell migration. Cell surface proteins that mediate and regulate interactions between cells. Similar adhesion mechanisms may mediate homotypic cell-cell cohesion during collective invasion and transient and more dynamic heterophilic interaction to resident tissue cells encountered during tissue invasion. Friedl and Wolf, 2011.

The cadherin family of adhesion receptors mediate homotypic interactions between cells of the same type and heterotypic interactions between different cell types. These interactions include stable cell-cell adhesion through adherens junctions, dynamic adhesion via the transient co-engagement of the small GTPases RAC1 and RHOA, and dynamic junctional remodeling by cytoskeletal dynamics. In both stable and dynamic cell-cell adhesion, cadherins engage with cytoskeletal adaptor and signaling proteins, including α-catenin, β-catenin, and p120-catenin, which connect to the actin and microtubule cytoskeleton [251]. Depending on the type of tumor, different sets of cadherins are expressed and involved in cell-cell interaction.

In polarized resting epithelium, E-cadherin suppresses migration signaling by inhibiting RAC1 and further maintains cell-cell cohesion, polarity between the basal and luminal layer of an epithelium, and epithelial stability [232]. Thanks to its capacity to stabilize epithelial tissues, E-cadherin has been described as a tumor progression suppressor, presumably by preventing cell invasion and thus metastasis. Indeed, the loss of E-cadherin (Figure 32 A), is a major hallmarks of EMT occurring during invasion processes. First, loss of cell surface E-cadherin disrupts adhesion junctions between neighboring cells and thereby supports detachment of malignant cells from the epithelial-cell layer. Second, loss of E-cadherin has direct effects on signaling pathways involved in tumor-cell migration and tumor growth, including the canonical Wnt signaling pathway and Rho family GTPase-mediated modulation of the actin cytoskeleton [252]. In EMT, the loss of E-cadherin is frequently a switch for the expression of mesenchymal cadherins, such as N-cadherin, which enhance tumor-cell motility and migration. This cadherin switch may make a critical contribution to tumor invasion and metastatic dissemination, because compared to E-cadherin, N-cadherin mediates much weaker interactions, and also through co-engagement of growth factor receptors, including FGFR or PDGFR, which enhances downstream signaling through MAPK and PI3K, modulating various signaling pathways and transcriptional responses (Figure 32 B). Thus, cadherins show duality

in delivering both migration-inhibiting and migration-promoting signaling in a context-dependent manner [232].

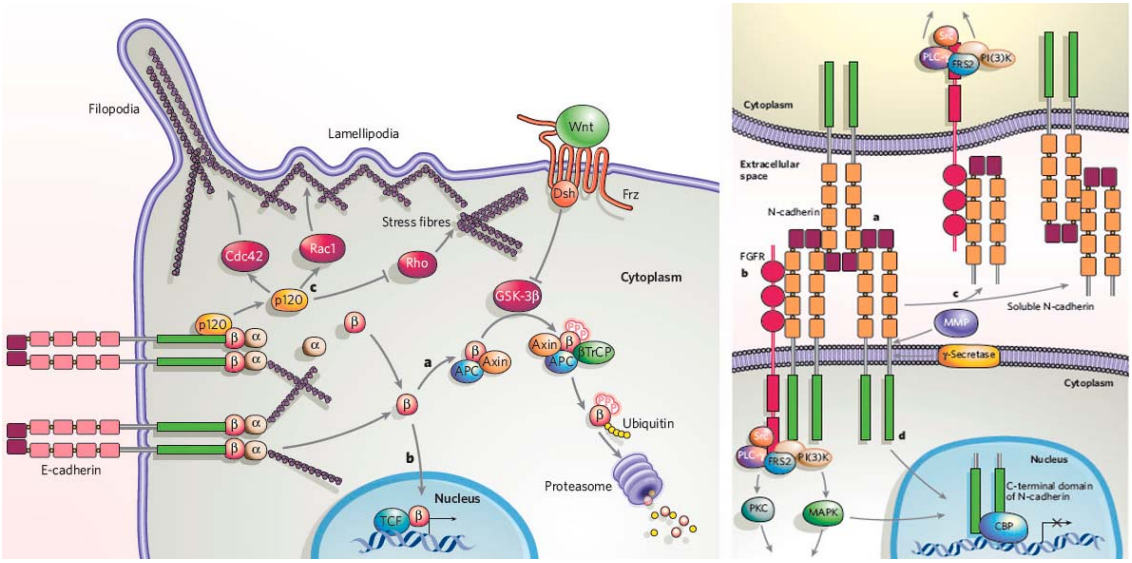


Figure 32. Potential signaling pathways downstream of the loss of E-cadherin function and N-Cadherin mode of action. Left panel: (a) After loss of E-cadherin function, β-catenin (β) is sequestered by the adenomatous polyposis coli (APC)–axin–GSK-3β complex and phosphorylated by GSK3β. This phosphorylated β-catenin is specifically bound and ubiquitylated by βTrCP, a subunit of the E3 ubiquitin ligase complex. Ubiquitylation marks β-catenin for proteasomal degradation. (b) When the Wnt signaling pathway is activated, GSK-3β is repressed and instead of being phosphorylated, β-catenin translocates to the nucleus. Together with TCF/LEF-1 transcription factors, it modulates the expression of several target genes involved in cell proliferation and tumor progression. (c) On disassembly of the E-cadherin adhesion complex, displaced p120 represses the small G protein RHOA and activates RAC1 and CDC42, which together modulate the actin cytoskeleton and the migratory behavior of tumor cells. Filopodia are induced by CDC42, lamellipodia are induced by RAC1, and stress fibers are induced by RHOA. Right panel: N-cadherin has several functions, all of which may contribute to tumor invasion. (a) Cell-cell adhesion to N-cadherin-expressing cells of the stroma. (b) Binding and activation of FGFRs, which results in the assembly of a classical FGFR signaling complex and activation of downstream phospholipase Cγ (PLC-γ), PI3K and MAPK signaling pathways, thereby promoting cell survival, migration and invasion. (c) Cleavage and shedding of the extracellular domain of N-cadherin by MMPs. Shedded N-cadherin may neutralize N-cadherin-mediated cell-cell adhesion and/or stimulate FGFR signaling on neighboring cells. (d) Cleavage of N-cadherin by a γ-secretase-like protease results in translocation of the carboxy-terminal fragment of N-cadherin to the nucleus, where it represses CREB-binding protein (CBP)-mediated gene expression. Modified from Chistofori 2006.

In contrast, as recently suggested by Friedl and Gilmour [139], in activated and neoplastic epithelium, E-cadherin and other cadherins jointly coordinate collective movements. Such scenario of collective invasion may change the current view of E-cadherin in tumor progression and, as with many other oncogenic genes, it should be re-identified both as a tumor suppressor and a tumor progressor, depending on tumor type, time frame and context. If there is no requirement to repress E-cadherin for tumor cells to disseminate, it follows that the widely accepted notion that tumor cells require EMT for metastatic spread may need to be revised.

In activated epithelial cells, co-signaling of E-cadherin and integrins, together with downstream SRC activation, enhances actin dynamics and acto-myosin

contractility, leading to both single-cell and collective migration. When co-engaged with DDR1, E-cadherin signaling limits actomyosin contractility along cell-cell junctions, which stabilizes cell-cell junctions and supports collective invasion. In contrast, during single cell invasion, as EMT, there is a crosstalk between integrins and E-cadherin that coordinates the switch from cadherin- to integrin-mediated adhesions during EMT [121]. Downregulation of E-cadherin has been shown to signal to integrins, because endocytosis of E-cadherin results in the activation of the small GTPase RAP1, a protein that regulates the cytoplasmic activation of integrins and that is required for focal-adhesion formation. On the other hand, integrins can cause the downregulation of E-cadherin in a number of ways. Integrin-linked kinase (ILK) interacts with the cytoplasmic domain of β1 and β3 integrins, and it is activated through cellular interactions with the ECM and growth factors. ILK downregulates E-cadherin expression and is required for TGFβ-induced EMT. β1 integrins also and activate RHOA and RAC1, which leads to the disruption of cadherin-mediated adhesions. Integrin-mediated focal adhesion kinase (FAK) activation can transiently downregulate RAC1 in epithelial cells at sites of formation of N-cadherin-mediated cell-cell contacts. These findings are in agreement with the observation that constitutive RAC1 activation does not allow the establishment of cell-cell contacts. In addition, constitutively active SRC induces EMT and internalization of E-cadherin that is dependent on signaling by αvβ1 integrin as well as on SRC-dependent FAK phosphorylation.

Moreover, multiple transcriptional pathways are activated on E-cadherin loss inducing wide range of functional changes and favoring the expression of EMT inducers such TWIST1 and ZEB1. This reveal a complex transcriptional network controlled by E-cadherin due to its intracytoplasmic domain [253].

The Immunoglobulin superfamily of cell adhesion molecules (CAM), mediates homophilic cell-cell interactions in neoplastic cells through the direct or indirect coupling to actin cytoskeleton via actin binding adaptor proteins α-actinin, ankyrin and ezrin [254]. This family of molecules includes L1CAM, EpCAM, NCAMs and ALCM [232]. L1CAM is upregulated in the leading front of collectively invading epithelial tumors that display a stabilized mesenchymal phenotype with high invasion capability. This is consistent with a role for L1CAM in leader-cell function and partial EMT during collective invasion. EpCAM is a marker of epithelial cells and in stemness signaling [255]. ALCAM is upregulated in cell-cell junctions of collectively invading epithelial cancer associated with increased metastasis. Besides mechanocoupling between CAM and integrins, CAMs enhance the signaling of integrins and growth factor receptors (e.g., EGFR and FGFR) through ERK, ILK, or SRC. Their contributions to homotypic interaction between tumor cells and heterotypic interactions between tumor and stromal cells make CAMs versatile mechanotransduction and signaling devices in both single-cell and collective invasion.

Protease systems. Proteases have the ability to degrade ECM components and thus remodel it during physiological processes and upon aberrant induction in malignant diseases co-opted in angiogenesis, invasion and metastasis. Many different proteases are alterated during tumor progression by tumoral cells and

stromal cells. Recent evidence reveals that tumor-promoting proteases function as part of an extensive multidirectional network of proteolytic interactions, in contrast to the unidirectional caspase cascade. Furthermore, the proteolytic network interacts with other important signaling pathways in tumor biology, involving chemokines, cytokines, and kinases.

There are five human protease classes categorized by their catalytic mechanism (aspartic, cysteine, metallo, serine and threonine). Likewise, endogenous inhibitors, that control their activities, exhibit specificity in their targets. Thus, cystatins predominantly inhibit cysteine proteases, serpins are most effective against serine proteases and tissue inhibitors of metalloproteinases (TIMPs) target metalloproteinases [257]. Key proteases participating in tumor progression are: MMPs, ADAMs, cathepsins, uPA-uPAR [256] (Figure 33).

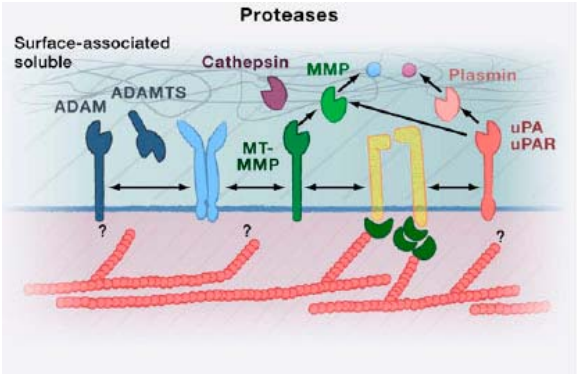


Figure 33. Molecular determinants of cell migration. Protease systems upregulated in cancer progression, angiogenesis, invasion, and metastasis. Friedl and Wolf, 2011.

Extracellular proteolytic enzymes may follow different mechanisms of action during migration and invasion [232]:

- Cell surface proteases, such as MT-MMPs and ADAMs, degrade ECM structural proteins, including collagens (fibrilar and non-fibrilar), fibronectin and laminins, and matricellular proteins such tenascin or glyplican. This ECM degradation has two consequences: (a) It generates biologically active epitopes of ECM components with adhesion or migration-promoting effects. And (b) it remodel the structure of the surrounding tissue opening new gaps and trails to go, bordered by multifiber ECM bundles.

- Secreted proteases, most of them MMPs and ADAMs, process other proteases and cell surface receptors, including adhesion receptors and also growth factor receptors. This implies the existence of a turnover for all these receptors and thus condition receptor availability on tumoral and stromal cells. Cathepsins may favor invasion by cleaving the extracellular domain of E-cadherin [257].

- Regulation of the repertoire of growth factors available in the ECM by enzymatic activation or inactivation or degradation. Secreted proteases such as MMPs and plasmin mediate this control.

Several hallmarks of aggressive cancer are a direct result of proteolytic activity, including, but not limited to, tumor cell invasion into the stroma, angiogenesis and metastasis. A separate group of cascading proteolytic interactions controls apoptosis, which cancer cells must escape in their progression to malignancy.

Further interactions in the network come through indirect interactions, often as a result of one protease cleaving and inactivating the inhibitor of another protease, or by proteases modulating the activity and availability of a signaling factor that can affect the abundance of another protease. These interactions enable proteases to indirectly increase the activity of other proteases without physically interacting with them. In normal physiology, this can serve to control and prevent the over-activation of proteolytic pathways; the extent to which this occurs in the tumor microenvironment is still unclear but it could potentially serve to inactivate tumor-suppressing proteases.

Understanding how inhibition of one protease can affect the overall proteolytic balance of the tumor microenvironment is crucial to enable researchers to design agents that target proteases with maximal impact and minimal toxicity [256].

Chemokines, growth factors and their receptors. The transition from a fixed, tissue-anchored state to a mobile state is often induced by extracellular chemokines, cytokines, and growth factors released by tumor cells themselves or by activated stromal cells. These factors engage redundant and non-redundant intracellular signaling networks in both tumor and stromal cells. Because of their pleiotropic effects, pro-migratory conditioning of the tumor-associated tissue increases: (1) the invasion and dissemination of tumor cells; (2) the motility and activity of stromal cells, including fibroblasts and macrophages; (3) the recruitment and transendothelial migration of circulating leukocytes and precursor cells into the tumor stroma; and (4) the mobilization of bone marrow-derived cells into the circulation through systemic growth factor effects in other organs, including the bone marrow [232]. This is mediated by invasion promoting chemokines such as CXCL12, CXCL10, CCL21 or CCL25, and receptors such as, CXCR4, CXCR3 and CCR9. Growth factors, which influence to undergo EMT (described above) and are upstream to RAC and CDC42 signaling and other cytoskeletal influencing factors (Figure 34).

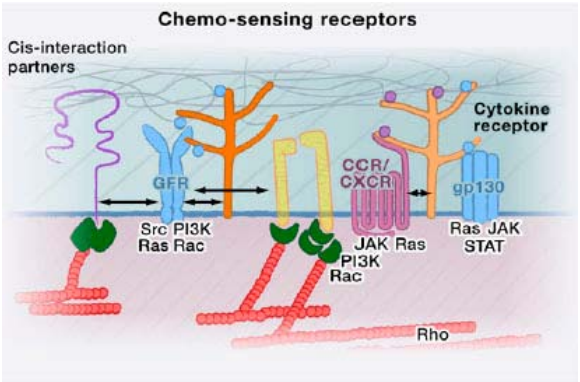


Figure 34. Molecular determinants of cell migration. Receptors for chemokines, cytokines, and growth factors, which sense soluble, ECM-, or proteoglycan-bound factors and interaction partners. Friedl and Wolf, 2011

6.3. Invasion routes

Invasiveness is a the key capacity that tumor cells acquire in order to metastasize and although often associated to EMT, tumor cells can move and migrate in the absence of EMT, as described above, depending on the tumor type, context and stage, by virtue of different types of collective invasion and amoeboid migration. Intravital microscopy and histopathological analysis of human tumors, strongly suggest that the simplest strategy for cancer cells to disseminate is the movement along pre-existing tissue structures, or interstitial guidance, which can be coupled to more active invasive mechanisms. In other words, invasion into healthy tissues develops preferentially along pre-existing tracks of least resistance, followed by secondary tissue remodeling and destruction. The tissue scaffolds supporting or preventing guidance of invasion vary in structure and molecular composition between organs [250].

Anatomic structures guiding the movement of cancer cells include: epithelial and endothelial surfaces devoid of ECM, as it happens in the lumen of small vessels in peripheral tissue, liver sinusoids, and peritoneum; basement membranes interfacing with the ECM between cells and tissues, as in perivascular space, perineural space or fat tissue; fibrilar interstitial tissue, as in bundled or random fibrilar ECM; and complex interfaces composed by cell surfaces and ECM scaffolds, such as bone cavities or brain vessels [232]. As examples, two principal types of interstitial tissues are transmigrated by tumor cells [250]: (i) collagen-rich interstitial connective tissue present in most parenchymatous organs of the body, such as breast, and (ii) the nervous interstitium, such as the brain. In the brain, the guidance is provided by myelinated axons, astrocyte processes, and blood vessels, which are used as invasion routes by glioma cells. In the human breast, containing interstitial collagen-rich connective tissue, disseminating breast cancer cells preferentially invade along bundled collagen fibrils and the surface of adipocytes. In both invasion types, physical guidance prompted by interfaces and space is complemented by molecular guidance (Figure 35). Generic mechanisms shared by most, if not all, tissues include (i) guidance by integrins towards fibrillar interstitial collagen and/or laminins and type IV collagen in basement membranes decorating vessels and adipocytes, and, likely, CD44 engaging with hyaluronan; (ii) haptotactic guidance by chemokines and growth factors; and likely (iii) physical pushing mechanisms. Tissue-specific, restricted guidance cues include ECM proteins with restricted expression (tenascins, lecticans), cell-cell interfaces, and newly secreted matrix molecules decorating ECM fibers (laminin-332, thrombospondin-1, osteopontin, periostin).

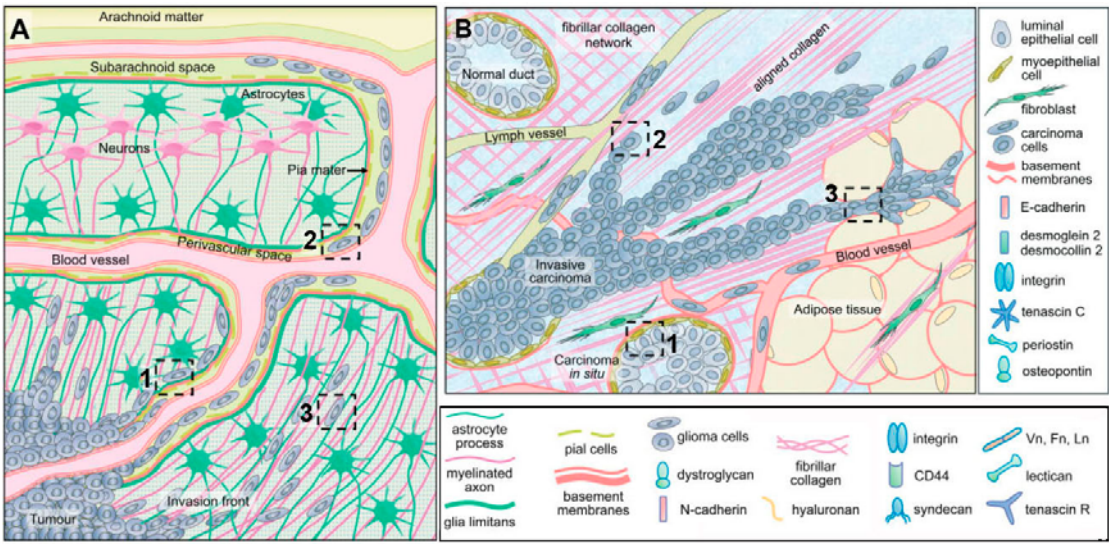


Figure 35. Anatomic and molecular guidance of glioma cell invasion and breast cancer invasion. (A) Guidance along the glia limitans and perivascular space, as well as by neuronal and astrocyte tracks. 1) Extravascular guidance along the vessel–stroma interface. 2) Guidance by the perivascular space. 3) Glioma cell migration along white matter tracks. (B) Overview of guidance structures in the mammary gland. 1) Intraductal guidance in breast carcinoma in situ. 2) Guidance by fibrous tissue. 3) Carcinoma cell invasion in adipose tissue. Modified from Gritsenko and Friedl, 2011.

Not only pre-existing structures serve as guidance. Molecules help in guidance by establishing gradients of promigratory factors, soluble factors or ECM-bound factors; ECM guidance contributes to cell adhesion, migrations and dissemination; cell membranes provide guidance over cell-rich tissue scaffolds through E-cadherin and RHO-mediated actomyosin contraction for migration; and guidance by secreted molecules, many chemokines and growth factors containing ECM-binding domains, which immobilize the factors in tissues, thereby forming a stable pro-migratory scaffold, as exemplified by TGF-β, which is immobilized via fibronectin and fibrilin [232].

6.4. Invasive phenotype plasticity

The executive mechanotransducing mechanisms of cell migration are plastic and allow the rapid adaptation to environmental changes and challenges; these adaptations often result in transitions between different modes of migration. Such plasticity likely originates in response to tissue microregions and responses to therapeutic challenge. The different modes of cancer cell invasion, the receptors and cytoskeletal regulators available for cell-cell and cell-matrix adhesion, the divergent degree of ECM remodeling capability, and the range of invasion-guiding molecular and physical tissue environments provide a multiscale framework of combinatorial possibilities or states that allow cancer invasion to be a plastic and adaptive process. Consequently, with altered tissue composition and conditioning by released factors, tumor cells undergo changes in signaling and function that lead to secondary effects in the invaded tissue and, in turn, the tumor cells themselves [231] [234]. Interstitial cancer cell invasion occurs in different phases that can be labeled operationally as an initial, nondestructive guidance phase, followed by a phase of tissue remodeling. In a stepwise manner, invasive migration leads to the production of pores, tunnels, and

lagunae, which guide and can be populated by mobile tumor cells [258] [104]. De novo generated tracks guide tumor cells and, with pressure exerted by the invading cells, become gradually widened until the tissue space consumed by invading cell masses matches the regression of the ECM [244].

EMT-dependent invasion and metastasis programs are strongly responsive to microenvironmental changes and adaptive in their signaling program and associated invasion dynamics. Because EMT can be displayed at varying degrees of completion, the modulation of EMT along the dissemination process can be a continuous dynamic modulation without requiring that the whole group of migrating cells undergo the same extent of EMT. Depending on their individual spatial positioning, cells can be more or less affected by microenvironmental factors such as WNT, TGF-β, FGF and EGF, all of which induce EMT through the upregulation of E-cadherin repressors and trigger a cascade of additional signaling events (Figure 36). In epithelial cancer lesions, EMT can be detected in a few cells, often forming cohesive clusters, located at the leading edge of migrating tumors, as well as in small cohesive groups and individual cells scattered and moving independently without connection to the main tumor [259]. Thus, besides representing a program for complete loss of cell-cell junctions, EMT may further contribute to collective cell functions, including collective invasion. This is consistent with the prominent collective invasion observed in primary mesenchymal tumors and melanoma.

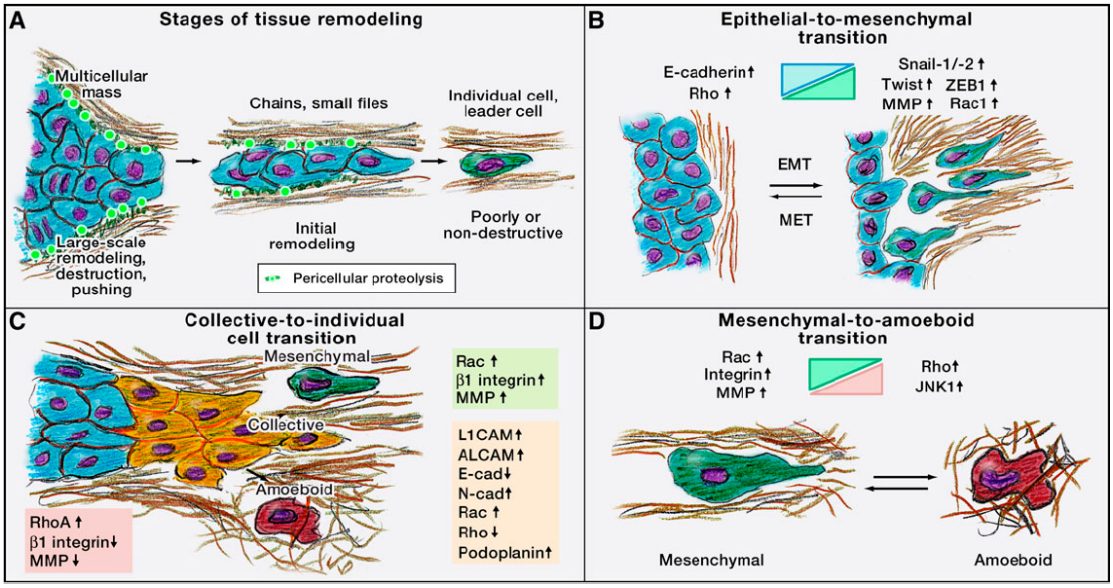


Figure 36. Plasticity of invasion mechanisms. A) Migrating cells transition from an initial nondestructive dissemination to migration that involves small- and large-scale tissue remodeling. The pre-existing space available to invading cells governs the caliber of individual and multicellular invasion and becomes widened by pericellular proteolysis. B) Epithelial-to-mesenchymal transition of a stable epithelium facilitated single-cell detachment. C and D) Invasion programs display plasticity, or adaptability, including transition from collective cell migration to individual cell migration and mesenchymal- to-amoeboid transition. Friedl and Alexander, 2011.

Environmental stimuli can favor single-cell detachment from tumors, partly through EMT and partly in the absence of EMT. Thus, alternatively, amoeboid dissemination may originate from collective invasion when cell-cell junctions are

abandoned and release the cells toward a single-cell migration program of low integrin-mediated adhesion and high RHO-ROCK-mediated cortical actomyosin contractility (Figure 34). This results in a non-proteolytic amoeboid cell deformation to bypass narrow ECM barriers or move along obstacle-free guidance paths [260]. Signaling networks govern the interconvertibility among single ameboid contractile movement *vs.* mesenchymal movement, and ROCK-RHO and JAK-STAT3 favors the actomyosion contractibility of ameboid movement, which affords 10 to 100 faster migration than the mesenchymal mode, while RAC-WAVE2 favors the conversion to the slower mesenchymal proteolytical movement [261] [262]. How invasion plasticity is connected with or distinct from EMT programs remains to be shown in vitro and in vivo. This will identify EMT-dependent and -independent routes and niches of natural and therapy-induced plasticity of invasion and their contribution to metastatic dissemination.

This dynamic tumor cell plasticity follows a spiral spatiotemporal circuit of, cause and consequence, plasticity and reciprocity (Figure 37). Tumor cells, through mediators, affect stroma and stroma affects tumor cells, and these interactions, evolve and drift as cells advance and move from one environment to another [232]. Such, spatiotemporal spiral circuits also occur during morphogenesis, although under strict control and end points. In cancer, those spirals do not have a clear end point.

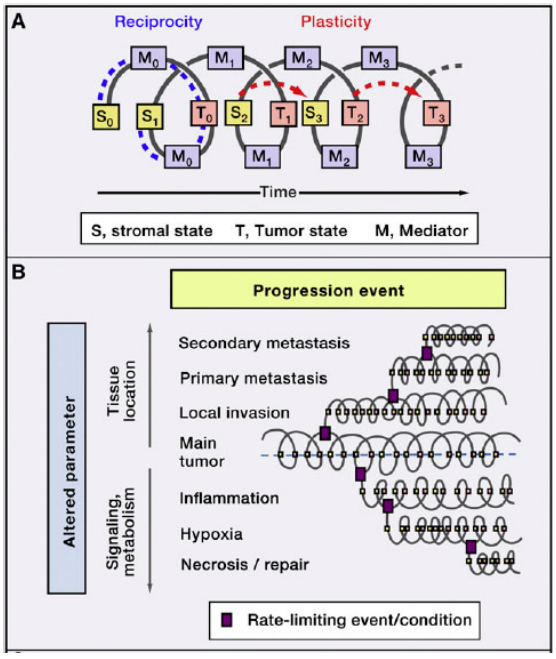


Figure 37. Spatiotemporal spiral of plasticity and reciprocity during cell invasion. A) Reciprocal crosstalk between tumor cells and the stroma (i.e., stromal cells together with ECM and released factors) results in evolutionary plasticity of both tumor cells and the tissue environment. ‘‘Reciprocity’’ results from the bidirectional communication between stromal (S) and tumor (T) compartments, which is transmitted by mediators (M) released by both compartments in a reciprocal manner. Stromal alterations include cell-derived physico-chemical changes of the microenvironment, such as deposited ECM components, ECM degradation and remodeling, change of ECM stiffness and porosity, and released cytokines and growth factors. Consequently, with each cycle of interactive engagement with the stroma, the cell state diverges from its origin, leading to progression of the tumor or the metastasis (indicated as a spiral). B) Branching and altered direction of reciprocal plasticity in the course of cancer progression. Direction-changing dichotomy is reached by a change in the position of the tumor cell, which results in a different tissue location and change in environmental input; likewise, step-wise bifurcation of reciprocal evolution may be induced by changes of the local tissue conditions, including altered composition of infiltrate cells during inflammatory and metabolic stress, insufficient perfusion resulting in hypoxia, and tissue repair programs induced by spontaneous or therapy-induced (tumor) necrosis. Friedl and Alexander, 2011.

Taken together, these considerations lead to a view of cancer cell invasion as a complex network of physico-chemical interactive mechanisms, employing a

diversity of invasion modes and following a dynamic spatiotemporal evolution. Tumor cells use different strategies and explore possibilities to escape from local sites and overcome the barriers that attemp to stop the invasive process. The block of one of these mechanistic pathways may result in the activation of alternative mechanisms of invasion and dissemination. The plasticity and thus adaptability of these cells, renders tumor cell invasion as a robust perpetuating process, the targeting of which will require the understanding of the hierarchy of stringent control points, including the application of mathematical modeling approaches to classify rate-limiting nodes and modifiers of molecular mechanotransduction.

7. The tumor microenvironment

Over the past few years, tumors have increasingly been recognized as organs whose complexity may even exceed that of normal healthy tissues. When viewed from this perspective, the biology of a tumor can only be understood by studying all the interactions between the individual specialized cell types within it, including the non-neoplastic cells that configure the ‘‘tumor microenvironment’’. This introduction has already commented several aspects of tumor microenvironment and its relationship with the evolution of the disease. A more detailed description is provided below of the relationship between neoplastic epithelial cells as a compartment (the parenchyma) and the non-neoplastic component that forms the tumor-associated stroma (Figure 38).

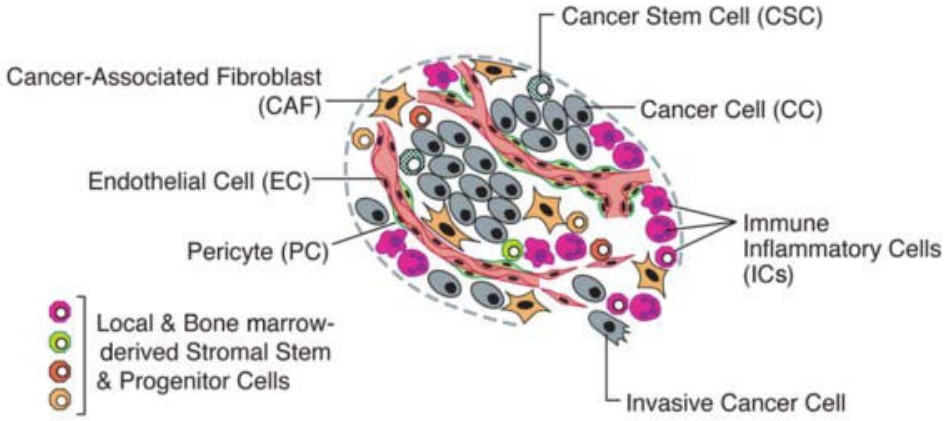


Figure 38. Neoplastic and tumor microenvironmental cells. Both the parenchyma and stroma of tumors contain distinct cell types and subtypes that collectively enable tumor growth and progression. Notably, the immune inflammatory cells present in tumors can include both tumor-promoting as well as tumor-killing subclasses. Hanahan and Weinberg, 2011.

7.1. Cell types composing tumors

Cancer cells are the initiators and drivers of tumor progression. It is important though, to recognize that they display a great degree of heterogeneity, with different subpopulations arising from the original neoplastic cells. Morphologically, this results in histopathological heterogeneity, with regions showing various degrees of apparent differentiation, proliferation, vascularity, inflammation, and/or invasiveness. In addition to this form of heterogeneity, it is still debated how heterogeneous are the subpopulations of CSCs that are at the origin and maintenance of the tumor and the rest of derived cancer cells. As discussed above, the dynamic phenotypic plasticity under which they are governed may produce a bidirectional interconversion between CSC and non-CSC states, resulting in a dynamic variation in the relative abundance of CSCs. Such dynamic equilibrium may be subordinated and orchestrated by different environmental cues. The phenotypic plasticity of CSC is an important feature that modulates the progression of tumors, as they can give rise to distinct and functionally diverse tumor populations, convert epithelial tumor cells into mesenchymal-like invasive cells or transdifferentiate into non-tumoral-collaborative cells, such as endothelial-like cells to form tumor-associated neovasculature, as has been shown in glioblastoma [263].

These observations indicate that certain tumors may acquire stromal support by inducing some of their own cancer cells to undergo various types of metamorphosis to produce stromal cell types rather than relying only on recruited host cells to provide their functions. Therefore, this genetic diversification may enable functional specialization, producing subpopulations of cancer cells that contribute distinct, complementary capabilities, which then accrue to the common benefit of overall tumor growth.

Other cells participating in the progression of the tumor and composing the surrounding stroma are endothelial cells, pericyte cells, immune inflammatory cells, cancer-associated fibroblast and stem/progenitor cells of the stroma (Figure 44).

Endothelial quiescent cells can be activated by an “angiogenic switch”, causing them to enter into a cell biological program that allows them to construct new blood vessels. Networks of interconnected signals modulate endothelial tumor-associated phenotype, including NOTCH, Neuropilin, ROBO, EPH, VEGF, angiopoetin and FGF signals [2]. Understanding the conversion of normal to tumor-associated endothelial cells will be important for novel therapies. While within the tumors the lymphatic vessels are collapsed, tumor-associated endothelial cells engage lymphangiogenesis at the periphery of the tumors, facilitating the channeling out of tumor cells, leading to cancer cell dissemination.

Pericytes, another specialized type of mesenchymal cells with finger-like projections, wrap around blood vessels and collaborate to synthesize the vascular membrane that anchors and helps vessels walls to withstand the hydrostatic pressure of blood flow. As a consequence, their destabilization affects vascular integrity and function. An intriguing hypothesis, still to be fully substantiated, is that tumors with poor pericyte coverage of their vasculature may be more prone to permit cancer cell intravasation into the circulatory system, enabling subsequent hematogenous dissemination [264].

Immune inflammatory cells were evidenced in the tumors in the 90s and over the years they have been shown to play diverse and critical roles in fostering tumorigenesis. Influx of pro-inflammatory cytokines, including TNF-α and TGF-β, as well as cytotoxic mediators, proteases, MMPs, interleukins, and interferons, produce potent lymphangiogenic and angiogenic growth factors allowing tumor growth and metastatic spread to the lymph nodes. Tumor cells themselves produce cytokines which attract neutrophils, macrophages, lymphocytes, and dendritic cells all attributing to tumorigenic growth and metastatic potential [265].

Inflammatory cells involved in tumor microenvironment and the modulation of tumor growth and progression include BMDCs (bone marrow derived-cells), such as macrophages, neutrophils, mast cells, myeloid cell-derived suppressor cells (MDSCs) and mesenchymal stem cells (MSCs), as well as T and B lymphocytes, and partially differentiated myeloid progenitors [266]. This compendium of cells releases a wide list of inflammatory signaling molecules that serves as effectors of tumor

promoting activities, including EGF as a tumor growth factor, VEGF as an angiogenic factor and other pro-angiogenic molecules like FGF2, chemokines and cytokines. In addition, these cells may produce proangiogenic and/or proinvasive matrix-degrading enzymes, including MMP-9 and other matrix metalloproteinases, cysteine cathepsin proteases and heparanase. These effectors help to sustain tumor angiogenesis, stimulate cancer cell proliferation, facilitate, via their presence at the margins of tumors, tissue invasion, and support the metastatic dissemination and seeding of cancer cells. Furthermore, a class of tumor-infiltrating myeloid cells has been shown to suppress CTL and NK activity [2] [267]. Myeloid cells may play a double role by directly promoting angiogenesis and tumor progression while at the same time affording a means to evade immune system. The recruitment of tumor-promoting macropahages and neutrophils and, on the other hand, the innate immune system with tumor killing response is subjected to a balance which will demark the progression of the tumor.

Tumor-associated macrophages (TAMs) are the prototypical BMDC capable of modifying the behavior of cancer cells behavior. At the tumor periphery, they can foster local invasion by supplying matrix-degrading enzymes such as metalloproteinases and cysteine cathepsin proteases in one model system, the invasion-promoting macrophages are activated by IL-4 produced by the cancer cells. And in an experimental model of metastatic breast cancer, tumor-associated macrophages (TAMs) supply epidermal growth factor (EGF) to breast cancer cells, while the cancer cells reciprocally stimulate the macrophages with CSF-1; their concerted interactions facilitate intravasation into the circulatory system and metastatic dissemination of the cancer cells [268] [269].

Cancer-associated fibroblasts (CAFs) are a subpopulation of activated fibroblasts or reactive stromal fibroblasts. Numerous growth factors such as TGF-β, chemokines such as monocyte chemotactic protein 1 (MCP1), and ECM-degrading proteases have been shown to mediate the activation of fibroblasts [270]. And, reciprocally, their production of growth factors, chemokines and extracellular matrix facilitates the angiogenic recruitment of endothelial cells and pericytes, and influence tumor stromal cells (Figure 39). Fibroblasts are therefore a key determinant in the malignant progression of cancer and represent an important target for cancer therapies. CAFs include at least two distinct cell types: cells with similarities to the fibroblasts that create the structural foundation supporting most normal epithelial tissues and the myofibroblasts, whose biological roles and properties differ markedly from those of tissue-derived fibroblast. Myofibroblasts are identifiable by their expression of α-smooth muscle actin (SMA). They are rare in most healthy epithelial tissues and transiently increase in abundance in wounds or chronic inflammation sites. Recruited myofibroblasts and reprogrammed variants of normal tissue-derived fibroblastic cells have been demonstrated to enhance tumor phenotypes, notably cancer cell proliferation, angiogenesis, and invasion and metastasis. They secrete a variety of extracellular matrix components, and are implicated in the formation of the desmoplastic stroma that characterizes many advanced carcinomas [271], and they may induce EMT on tumor cells, whereas and

in other situations they may break up ECM opening path for tumoral populations to reach the bloodstream [272].

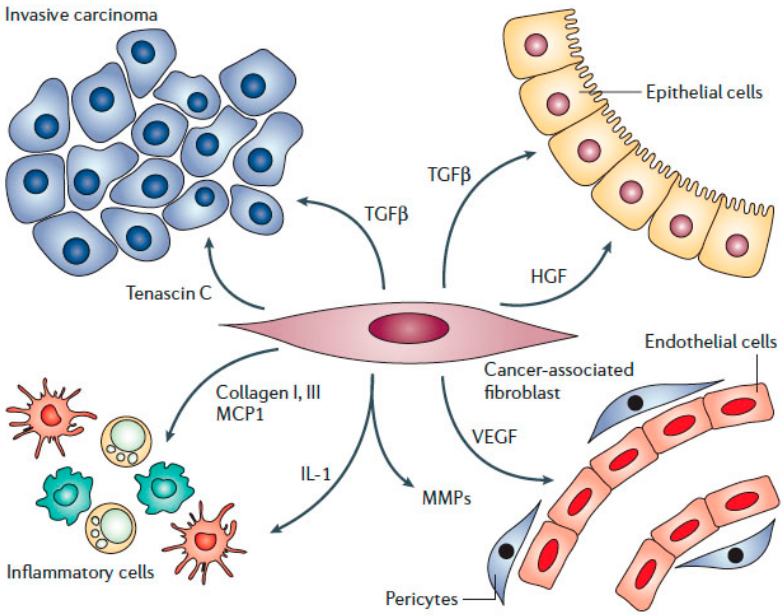


Figure 39. Functions of activated fibroblasts in tumor stroma. Fibroblasts communicate with cancer cells, resident epithelial cells, endothelial cells, pericytes and inflammatory cells through the secretion of growth factors and chemokines. Through the increased deposition of collagen types I and III and de novo expression of tenascin C, they induce an altered extracellular-matrix microenvironment, which potentially provides additional oncogenic signals, probably leading to accelerated cancer progression. Fibroblasts mediate the inflammatory response by secreting chemokines such as monocyte chemotactic protein 1 (MCP1) and interleukins such as IL1. Fibroblasts interact with the microvasculature by secreting matrix metalloproteinases (MMPs) and vascular endothelial growth factor (VEGF). Fibroblasts also provide potentially oncogenic signals such as transforming growth factor-β (TGF-β) and hepatocyte growth factor (HGF) to resident epithelia, and directly stimulate cancer-cell proliferation and invasion by secreting growth factors such as TGF-β and stromal-cell-derived factor 1 (SDF1). Kalluri and Zeisberg, 2006.

Mesenchymal stem cells (MSCs) are multipotent cells that differentiate into osteoblasts, chondrocytes, adipocytes and other cells of mesenchymal origin [325]. They originate in the bone marrow, where they predominantly reside, though they are not of hematopoietic origin. Mesenchymal stem and progenitor cells have been found to transit into tumors from the marrow, where they may differentiate into the various well-characterized stromal cell types. Some of these recent arrivals may also persist in an undifferentiated or partially differentiated state, exhibiting functions that their more differentiated progeny lack. Co-mingling of human MSCs with weakly metastatic breast cancer cells significantly increased the dissemination of the cancer cells to the lung from a subcutaneous xenograft, an effect that was not observed with other mesenchymal cells, such as normal fibroblasts [273]. Interestingly, the inductive effects of MSCs on cancer cells were mediated exclusively at the primary site, apparently priming them for dissemination to the lung. This is controlled, at least in part, through a paracrine loop involving MSC-supplied CCL5 and its receptor, CCR5, on the breast cancer cells. Various lines of evidence indicate that tumor-associated stromal cells may be supplied to growing tumors by proliferation of preexisting stromal cells, by differentiation in situ of local stem/progenitor cells

originating in the neighboring normal tissue, or via recruitment of bone marrow-derived stem/progenitor cells [2].

Interestingly, it has recently been reported [273] that MSCs in the breast tumor microenvironment can stimulate the epithelial mesenchymal transition (EMT) of cancer cells. In that study, molecular characterization revealed that upon cellular co-culture, the cancer cells gained in mesenchymal markers with the reciprocal downregulation of E-cadherin and loss of proliferation-associated genes.

7.2. Signaling network in the tumor microenvironment

Complex interactions take place between the neoplastic and stromal cells within a tumor and the dynamic extracellular matrix that they collectively erect and remodel [274] (Figure 40). Both neoplastic cells and the stromal cells surrounding them change progressively during the multistep transformation of normal tissues into high-grade malignancies. Such stepwise progression is likely to depend on back-and-forth reciprocal interactions between the neoplastic cells and the supporting stromal cells. This dynamic course of multi-stage tumor development clearly complicates the goal of fully elucidating the mechanisms of cancer pathogenesis but the understanding of these dynamic variations will become crucial to the development of novel therapies designed to successfully target both primary and metastatic tumors [2].

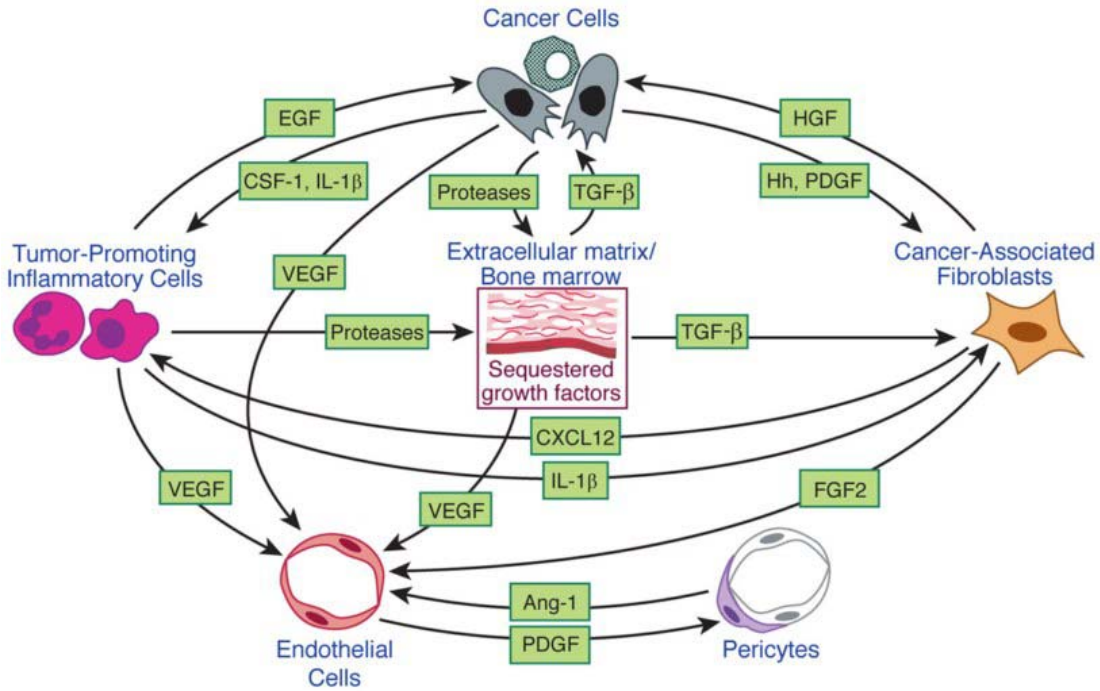


Figure 40. Interactions in the tumor microenvironment. The assembly and collective contributions of the assorted cell types constituting the tumor microenvironment are orchestrated and maintained by reciprocal heterotypic signaling interactions. Only few of them are here represented. Hanahan and Weinberg, 2011.

7.3. Proteases in the microenvironment

The importance of chemoattractant signaling in cancer cell invasion is evident. However, to physically invade into blood vessels, proteolytic degradation is required, as discussed in previous sections [274]. Proteases are often produced by invasive cancer cells, but in many cases bone marrow-derived cells (BMDCs), including macrophages, have been shown to be the major cell types that supply crucial proteases to the tumor microenvironment. In this context, it is important to note that such proteases, that play essential roles in ECM degradation, also engage in more specialized roles that are important for cell signaling, such as in the restricted cleavage of pro-domains and subsequent activation of growth factors and cytokines, which may significantly increase chemoattraction, cell migration and metastasis. Indeed, all these different modes of protease-enhanced invasion are not mutually exclusive; rather, it is likely that they act in concert to promote cancer cell spread. All of these functions are tightly regulated in a cascade of protease interactions, allowing for control and amplification of proteolysis in invasion and metastasis.

7.4. Hypoxia in tumors and metastasis

A critical hallmark of solid tumors is low oxygen tension, or hypoxia. Because the vasculature cannot sustain the demands of the cancer cells after a certain tumor mass is attained, solid tumors are invariably less well oxygenated compared to the normal tissue of origin. In an attempt to overcome the deficiencies of the existing vasculature and stimulate their own growth, tumors can change their metabolism and/or induce the expression of pro-angiogenic factors to increase their blood supply. These newly formed blood vessels are highly twisted, tortuous with blind ends and arterio-venous shunts. The malformed vasculature results in sluggish blood flow as well as transient changes in oxygenation due to the unpredictable dilation and constriction of newly constructed vessels. Furthermore, solid tumors are subjected to chronic changes in oxygenation that are directly related to their distance from a blood vessel. Both the proportion and rate of actively dividing tumor cells are a function of their distance from the vasculature. Additionally, this neovasculature provides accessible new ways for cancer cell dissemination [275].

At the cellular level, low oxygen concentration is known to induce genes involved in the regulation of cell proliferation, extracellular matrix production, cell adhesion, and other hallmarks of tumorigenesis. The mechanism behind these effects is frequently accomplished through induction of the hypoxia-inducible factor (HIF) family of transcription factors (Figure 41). This family consists of three members, HIF-1,-2, and -3, which act to regulate cellular processes including glucose metabolism, angiogenesis, cell proliferation, erythropoiesis, and tissue remodeling in response to low oxygen levels [275] [279]. Additional transcription factors also respond to an hypoxic environment, namely NF-κB, P53, AP-1, C/EBPB, EGR-1 and SP-1, and, together, contribute to the adaptation of normal and tumoral cells to hypoxia.

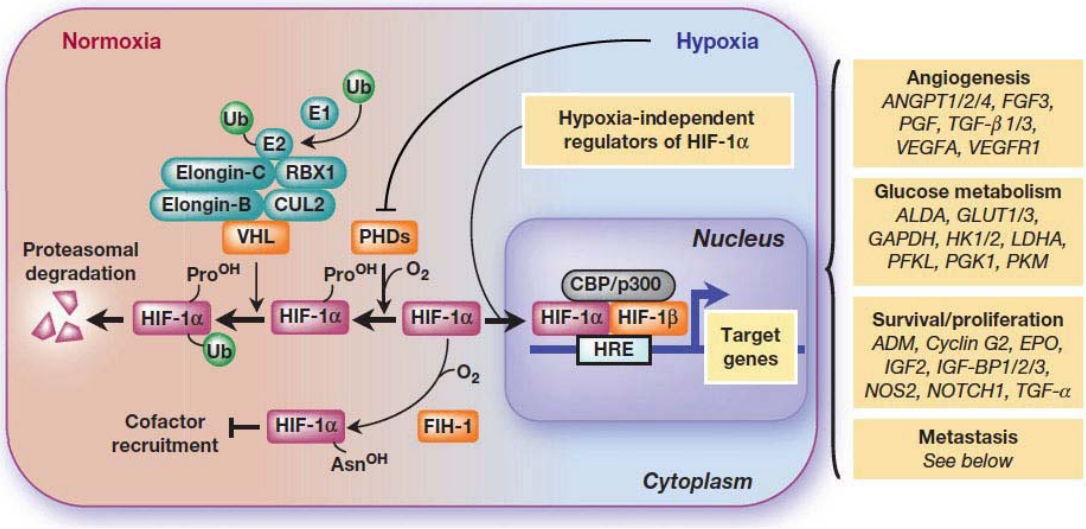


Figure 41. Regulation of HIF-1α under normoxia and hypoxia conditions. In the presence of O2, prolyl-4-hydroxylases (PHDs) hydroxylate two conserved proline residues, 402 and 564, of HIF-1α, a reaction that can be inhibited by iron chelation. Hydroxylated HIF-1α is recognized and interacts with VHL, which is part of an E3 ubiquitin ligase complex that mediates, in concert with elongins, the ubiquitylation (Ub) of HIF-1α and targets HIF-1α for degradation in the proteasome. Under hypoxia, HIF-1α is not hydroxylated because PHDs requires oxygen and 2-oxyglutarate as substrates, as well as iron and ascorbate as cofactors, thus preventing its ubiquitylation by VHL. Stabilized HIF-1α moves into the nucleus, dimerizes with HIF-1β, and binds to the HRE, to a core of 5'-RCGTG-3' sequence. By interacting with cofactors such as CBP/p300 and DNA polymerase II, HIF-1 activates the transcription of target genes. Binding to CBP/p300 is blocked when HIF-1α is hydroxylated by FIH-1. HIF-1α is additionally regulated by oncogenic pathways (e.g., ERBB2, SRC, ET-1, RAS/MARK pathway, PI3K-Akt-mTOR pathway), mutations of tumor suppressor genes (e.g.,PTEN, VHL, SDH, and FH), and reactive oxygen species (ROS). HIF-1 target genes related to cancer are categorized into four groups, with representative genes in each group, angiogenesis, glucose metabolism, survival-proliferation and metastasis. Lu and Kang, 2010.

In fact hypoxia may play a dual role in tumorigenesis: insufficient oxygen limits tumor cell division while at the same time selecting for more malignant cells and inducing cell adaptations allowing for more invasive behavior. Regarding tumor growth, cancer cells, similar to normal cells, need oxygen to generate energy as well as acting as a substrate for many fundamental cellular processes, including generating macromolecules. Yet, hypoxia is also strongly associated with tumor progression and metastatic disease. This is likely because low oxygen tension is able to increase cell invasiveness, cause cells to switch to anaerobic metabolism, increase genetic instability, and promote angiogenesis.

Of note, regulators of SNAI1, such as lysyl oxidase like 2 (LOXL2) and related proteins, such as lysyl oxidase (LOX), have also been described as HIF-1 targets that are essential for EMT. Notch is reported to recruit HIF-1α to the LOXL2 promoter [276]. However, the molecular mechanisms by which hypoxia induces EMT are still not fully understood. HIF-1 can control the expression specific EMT inducers (SNAI1, TCF3, ZEB1, and ZEB2) likely by direct binding of HIF-1α to the HRE on the SNAI1 and ZEB2 promoters or indirectly inducing TGF-β or β1-integrin [277] [278]. It has been shown that TWIST1 is a direct target of HIF-1α, which induces EMT in its own right, and also mediates the activation of other pathways, including the induction of miR-

10b, which in turn regulates HOXD1, affecting RHOC and finally controlling migration. SNAI1 and TWIST1 collaborate with other EMT transcription factors induced by hypoxia and also with LOXL2, which clinically correlates with hypoxia, playing an important role in EMT induction [281]. Other EMT related pathways induced by HIF-1α are the activation of WNT/β-catenin activation, the activation of uPA, uPAR, MMP2, MMP9 and the CTGF [282]. Thus, hypoxia is able to engage an EMT program to eventually transforms cells into malignancy, invasion and dissemination of the tumor.

From a therapeutic perspective, the fact that hypoxia triggers all these adaptive mechanisms adds to the complexities of targeted treatment of tumors. Many promising anti-angiogenic therapies were designed in order to prevent the growth of the neovasculature that supports tumor growth. However, these therapies have generally failed partly owed precisely to the consequences of the hypoxic response in tumor cell evolution: deprivation of blood supply causes tumor hypoxia, which kills the more sensitive tumor cells, but also promote the emergence of better fit cell populations through the activation of HIF-1α and the rest of the hypoxia response networks. Downstream mechanisms are subsequently engaged and tumors eventually become more aggressive upon anti-angiogenic therapies, leading to therapeutic failure [283]. Many drugs inhibit HIF-1 via different molecular mechanisms. For example, Acriflavin inhibits the heterodimerization of HIF-1α and HIF-1β, resulting in suppression of prostate and breast cancer metastasis in mouse models. Therefore, a promising strategy would be the simultaneous target of both mechanisms, angiogenesis and hypoxia, or its downstream consequences [281] to succeed in the inhibition of tumor progression.

7.5. Microenvironment and pH

Another feature that characterizes the pathophysiological tumor microenvironment which arises, at least partially, as a consequence of the abnormal tumor vasculature, is low extracellular pH. This is closely linked to hypoxia, since lack of oxygen from inadequate vasculature triggers cell glycolysis resulting in the accumulation of lactate. But hypoxia is not the only, or major, mechanism fostering glycolysis in tumor cells, as shown by the Nobel Prize winner Otto Warburg more than 90 years ago. The ‘‘Warburg effect” consists of an increase in glycolysis, observed in tumor cells, that is maintained even under conditions of high oxygen tension [280]. A major end product of the sustained glycolysis observed in tumor cells is lactate, which is secreted to the extracellular space, reducing the pH in the tumor milieu.

It has been shown that acidity induces the up-regulation of the proteolytic enzymes MMP-2 and MMP-9 and the angiogenic factors VEGF and IL8 in vitro, all of which are known to be involved in the metastatic process [282]. In addition to favoring tumor cell invasion, the acidity in the environment suppresses anticancer immune effectors.

7.6. Metabolism of tumor cells and stromal cells

The general enhancement of the glycolytic machinery in various cancer cell lines is well described and recent studies provide a better view of the changes in mitochondrial oxidative phosphorylation during oncogenesis. While some studies demonstrate a reduction of oxidative phosphorylation (OXPHOS) capacity in different types of cancer cells, other investigations revealed contradictory modifications with the upregulation of OXPHOS components and a larger dependency of cancer cells on oxidative energy substrates for anabolism and energy production. This apparent conflicting picture is explained by differences in tumor size, hypoxia, the sequence of oncogenes activated and stromal cell populations [285].

The continuous consumption of glucose and release of lactate by tumor cells produces many effects and adaptive responses in the rest of the cellular compartments of tumor microenvironment. The excess lactate produced by tumor cells can be taken up by stromal cells and used to generate pyruvate that either can be secreted and then refueled to the cancer cell (in a bidirectional feeding) or used for OXPHOS by other stromal cells, such as endothelial cells. This arrangement generates an ecosystem in which anaerobic components (cancer cells) and aerobic components (nontransformed stromal cells) engage in complementary metabolic pathways, thus buffering and recycling products of anaerobic metabolism to sustain cancer cell survival and growth (Figure 42). It is suggested that the newly formed stroma and vasculature express complementary metabolic pathways, buffering and recycling products of anaerobic metabolism to sustain cancer cell survival. Tumors survive and grow because they are capable of organizing the regional fibroblasts and endothelial cells into a harmoniously collaborating metabolic domain [285].

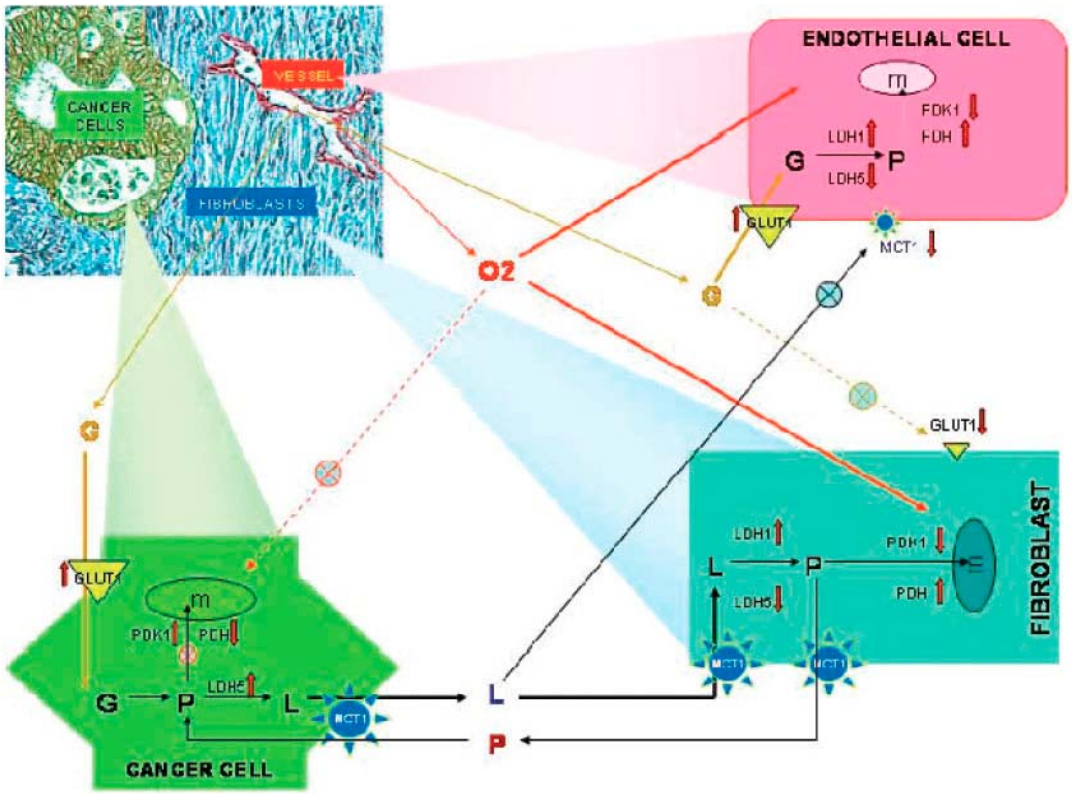


Figure 42. A self-feeding cancer microecosystem. Cancer cells share common enzyme/transporter activities suggestive of an anaerobic metabolism [high lactate dehydrogenase 5 (LDH5)/hypoxia-inducible factor-α (HIF-1α)] with high ability for glucose absorption and lactate extrusion [high glucose transporter 1 (GLUT1)/monocarboxylate transporter (MCT1)]. Tumor-associated fibroblasts expressed proteins involved in lactate absorption (high MCT1/MCT2), lactate oxidation (high LDH1 and low HIF-1α/LDH5), and reduced glucose absorption (low GLUT1). Expression profiles of the tumor-associated endothelium indicated aerobic metabolism (high LDH1 and low HIF-1α/LDH5), high glucose absorption (high GLUT1), and resistance to lactate intake (lack of MCT1). Koukourakis et al. 2006.

There are several reasons why enhanced glucose uptake for glycolytic ATP generation or anabolic reactions constitutes an advantage for tumor growth: with glycolysis, cells can live independently of oxygen fluctuations; bicarbonate and lactate favor invasion and immune evasion; tumors can metabolize glucose via pentose phosphate pathway (PPP) to generate NADPH against chemotherapeutic agents; they can use glycolytic intermediates as a source of carbon for anabolic reactions and thus proliferate and grow faster and the generation of microecosystems to self-feeding and survive [280]. Moreover, lactate contributes to cancer progression by stimulating signaling pathways such as HIF-1, NF-κB, TGF-β, Hyaluronic acid and CD44 [286], as well as angiogenesis due to the acidification of the microenvironment, as mentioned above.

7.7. Interstitial fluid pressure

The highly irregular tumor neovasculature generated by angiogenesis is characterized by increased permeability and microvascular resistance, which results in the movement of fluid into the ECM. Because tumors are either lacking in lymphatic vessels, or the intratumoral vessels are non-functional, excess fluid

accumulates in the interstitium, extending the elastic ECM and elevating interstitial fluid pressure (IFP). It has been suggested that the tumor vasculature is the driving force in increasing tumor IFP. However, the tumor stroma is also thought to play an active role due to anomalies within the ECM composition. Collagen fibers and their binding to integrins by fibroblasts modulate ECM elasticity. Furthermore, there is evidence that the process of collagen fiber contractions is actively regulated in response to specific cytokines such as PDGF, the receptors of which are expressed in the stroma of multiple tumors. Several studies have demonstrated that inhibition of PDGF signaling can reduce tumor IFP levels, consistent with a dynamic role in increasing tumor IFP. Moreover, IFP plays a role as a barrier against the delivery of therapeutic agents, thereby reducing their efficacy and multiple studies having shown improved uptake of chemotherapeutic drugs following a reduction in tumor IFP [283].

7.8. The power of the microenvironment in cell plasticity

Studies of cell fate developmentally imposed upon stem cells show that cell fate may depend on the tissue of destination of uncommitted or committed cells. For example, the mammary gland microenvironment can reprogram both embryonic and adult stem neuronal cells (NSCs) so that they become incorporated into the branching mammary tree to make chimeric glands and express the milk protein β-casein, progesterone receptor, and estrogen receptor alpha [287] In a different example of the dominant nature of microenvironment-directed fate determination, human embryonic stem cells (hESCs) that do not form tumors during blastocyst implantation form teratomas when transplanted into immunodeficient mice. In the context of carcinoma cells, which are expected to be more differentiated than adult stem cells, an interesting study impinges upon the idea of microenvironmental reprogramming of melanoma cells [288]. In a very interesting study [288] melanoma cells exposed only to 3D matrices preconditioned by hESCs were induced epigenetically into reprogramming, forming spheroids similar to the colonies formed by hESCs, along with significant reductions in their invasive capacity in vitro and in tumorigenic potential in vivo. Similar ongoing studies show similar suppressive effects of the hESC microenvironment on metastatic breast and prostate cancer cells.

Those are clear examples of how microenvironment can reprogram cells, including cancer cells into earlier multipotent stages, clearing part of their acquired malignant properties. Thus, the identification of the epigenetic mechanisms driving such regression will reveal interesting targets for therapy.

8. Epigenetics of cancer progression: DNA and chromatin modifications, and non-coding RNAis

Despite having identical genomic sequences, different cell types exhibit substantially different gene expression profiles, and their cellular identity must be conserved during somatic cell divisions. How then are cell-specific gene-expression patterns specified and maintained? The answer is epigenentics. Thus, cancer development depends not only on genetic alterations but also on an abnormal cellular memory, or epigenetic changes, which convey heritable gene expression patterns critical for neoplastic initiation and progression. These aberrant epigenetic mechanisms are manifest in both global changes in chromatin packaging and structure and in localized gene promoter changes that influence the transcription of genes important to the cancer process, since epigenetic changes might drive cancer progression. An exciting emerging theme is that an understanding of stem cell chromatin control of gene expression, including relationships between DNA methylation, histone post-translational modifications, histone variants and non-conding RNAs, which may hold a key to understanding the origins of epigenetic changes in cancer. This possibility, coupled with the reversible nature of epigenetics, has enormous significance for the prevention and control of cancer [289].

8.1. Epigenetic mechanisms

Pathological epigenetic changes, or non-sequence-based alterations that are inherited through cell division, are important alternatives to mutations and chromosomal alterations in disrupting gene functions resulting in cancer. Epigenetic mechanisms for such alterations include: alterations in DNA methylation patterns, chromatin alteration and loss of imprinting [8].

DNA methylation. It consist in the dynamic regulation of DNA cytosine methylation at CpG sites. Most CG dinucleotides are methylated on cytosine residues in vertebrate genomes. CG methylation is heritable, because after DNA replication the DNA methyltransferase 1, DNMT1, methylates unmethylated CG on the base-paired strand, providing memory activity. CG dinucleotides within promoters tend to be protected from methylation. Although individual genes vary in hypomethylation, all tumors examined so far, both benign and malignant, have shown global reduction of DNA methylation. Methylation marks over DNA promoters act as repressive marks inhibiting the transcription of nearby genes. Therefore, both hypomethylation and hypermethylation are relevant in cancer since they can alter gene expression. Hypomethylation generally arises earlier in the evolution of tumors and is strongly linked to chromosomal instability and loss of imprinting, allowing the expression of previously repressed oncogenes. Hypermethylation is associated with aberrant silecing of tumor suppressor genes, developmental transcription factors, tissue remodeling genes, DNA repair genes, cell cycle control genes, antiapoptotic genes and other genes important for initiation and progression of the tumors (Figure 43). In fact, any sinlge cancer may simultaneously have all such genes epigenetically silenced [8] [289]. Elevated levels of DNMT

proteins are elevated in various cancer types, including gastric, bladder, leukemia, colon and lung [289].

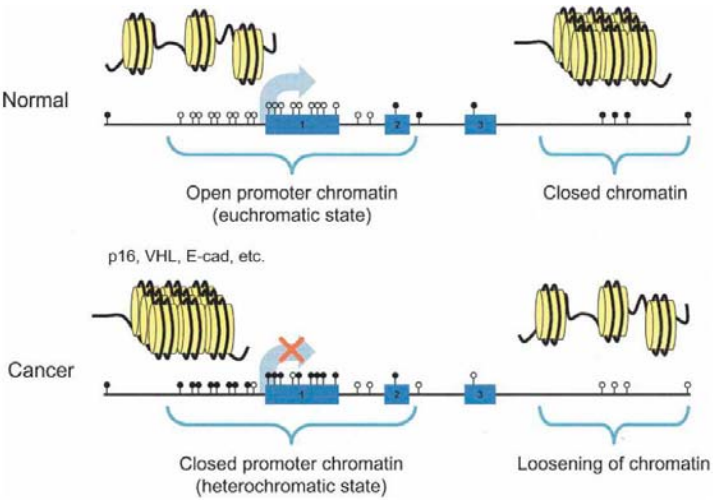


Figure 43. Normal versus cancer epigenome. In normal mammalian cells, CpG islands in proximal gene promoter regions (a three-exon gene) are largely protected from DNA methylation (cytosines are shown as open lollipops) and reside in restricted regions of open chromatin with nucleosomes with wide spacing, or euchromatic states, favorable for gene transcription (large blue arrow). In contrast, for most regions of the genome, such as in the bodies of many genes and areas outside genes, particularly including repeat elements and pericentromeric regions, the cytosines in CpG dinucleotides are methylated (black lollipops). This DNA methylation is characteristic of the bulk of the human genome, which is packaged as closed chromatin with nucleosomes with higher-order and tight compaction, unfavorable for transcription. In cancer cells, there tends to be a reversal of this pattern. Proximal promoter CpG islands for many abnormally silenced genes (representing tumor supressor genes) become DNA hypermethylated and reside in a closed chromatin, or more heterochromatic-type state, which is not favorable for transcription. In contrast, cytosines in CpG dinuleotides in other regions of the genome display hypomethylation and are associated with states of aberrantly loosened chromatin. The overall result is abnormal chromatin packaging with the potential for underpinning an abnormal cellular memory for gene expression and for conveying abnormal structural function for chromosomes. Ting et al. 2006.

Chromatin alteration. Chromatin is not simply a packaging tool, it is also a dynamically adjusted entity that reflects the regulatory cues necessary to program appropriate cellular pathways. Chromatin structure imposes significant obstacles on all aspects of transcription that are mediated by RNA polymerase II, transcription initiation and elongation. The dynamics of chromatin structure are tightly regulated through multiple mechanisms including histone post-translational modification, chromatin remodeling, histone variant incorporation, and histone displacement (eviction). Chromatin structure imposes profound and ubiquitous effects on almost all DNA-related metabolic processes including transcription, recombination, DNA repair, replication, kinetochore and centromere formation, and so forth [290].

The histone amino-terminal tails that protudes from the nucleosome and globular domains, are subject to a vast array of post-translational modifications: methylation of arginine residues; methylation, acetylation, ubiquitination, ADP-ribosylation, and sumolation of lysines residues; and phosphorylation of serines and threonines. Modifications that are associated with active transcription, such as acetylation of histone 3 and histone 4 (H3 and H4) or di- or tri-methylation (me3) of H3K4, are commonly referred to as euchromatin modifications (Figure 38). Modifications that are localized to inactive genes or regions, such as H3K9me and H3K27me, are often termed heterochromatin modifications. Most modifications are distributed in distinct localized patterns within the upstream region, the core

promoter, the TSS (transcription start site), the 5' end of the open reading frame (ORF) and the 3' end of the ORF. Indeed, the location of a modification is tightly regulated and is crucial for its effect on transcription (Figure 44). Histone acetylation occurs at multiple lysine residues and is usually carried out by a variety of histone acetyltransferase complexes and it has an accumulative effect. In contrast, histone methylation, phosphorylation, ubiquitination, etc., are often catalyzed by a specific enzyme at a specific site and result in unique functions [290].

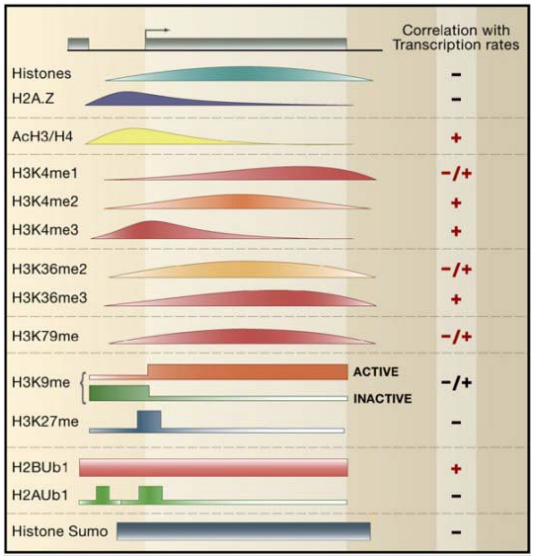


Figure 44. Genome-wide distribution pattern of histone modifications from a transcription perspective. The distribution of histones and their modifications are mapped on an arbitrary gene relative to its promoter and ORF. The curves represent the patterns that are determined via genome-wide approaches. The squares indicate that the data are based on only a few case studies. Li et al. 2007.

The packaging of the template into nucleosomes appears to affect all stages of transcription from activator binding and preinitiation complex formation to elongation. Once activators bind to the promoter, they trigger a cascade of recruitment of coactivator complexes. These coactivators include chromatin-remodeling complexes, histone-modification enzymes, and mediators.

A further aspect deserving consideration is the heritable epigenetic information in somatic cells transmitted from one cell generation to the next. Different models of histone segregation may explain these highly complex mechanisms still remainning unclear. Evidence of such inheritance are the transcription memory and position-effect of variegation. Understanding histone deposition following replication is crucial to elucidate the mechanism since the disruption of chromatin is inherent to replication and replication means that newly synthesized histones will need to be incorporated, as double the amount of DNA needs to be packaged into nucleosomes and somehow acquire the same epigenetic information [291]. This may be explained by different models of histone segregation and propagation aided by histone chaperones. Histones are lost to some extent during elongation of replication, at least partially and/or temporarily. However, with the help of histone chaperones, histones evicted in front of an elongating Pol II complex appear to be rapidly redeposited onto DNA behind Pol II.

All theses classes of proteins, involved in histone modifications and the transcriptional machinery, are shared as downstream effectors of many cell fate regulators and cell signaling pathways. Thus, as master regulators, they require to be

tighly regulated because their deregulation may cause a pleiotropic effects alterating many gene programs and the expression of individual genes.

Loss of imprinting. Abnormalities in gene imprinting may be linked to the origins of some neoplasias. Genomic imprinting refers to parent-of- origin-specific gene silencing. It results from a germ-line mark that causes reduced or absent expression of a specific allele of a gene in somatic cells of the offspring. Loss of imprinting (LOI) refers to activation of the normally silenced allele, or silencing of the normally active allele, of an imprinted gene. An example of this is the LOI of the insulin-like growth factor 2 gene (IGF2), resulting in a biallelic IGF2 expression which accounts for half of Wilms tumours in children, due to an abnormal progenitor cell expansion [8]. In these cases, LOI represents a switch in heritable gene expression patterns, in the absence of mutations, that may lead to abnormal expansion of stem/progenitor cells and, thus establish a risk that subsequent events will promote full transformation and evolution of cancer [289].

8.2. Oncogenic pathway addiction

Although epigenetic changes may occur at any time during tumor progression, it occurs most frequently during the early stages of the neoplastic process, such as the pre-cancerous stages of tumor development. Recent evidence indicates that epigenetic changes might "addict" cancer cells to altered signal-transduction pathways during the early stages of tumor development. Dependence on these pathways for cell proliferation or survival allows them to acquire genetic mutations in the same pathways, providing the cell with selective advantages that promote tumor progression.

The concept was proposed by Weinstein [292] as applicable to cancer cells that become addicted to mutated oncogene products or hypersensitivity to the loss of function of mutated tumor-suppressor genes. A variant of this concept has been applied to epigenetic alterations by Baylin and Ohm, who designated it as “epigenetic sensitization” [293]. An example is provided by colon cancer, which is believed to be caused only by genetic mutations activating the Wnt pathway. Studies of early-stage, pre-invasive colon lesions that are at risk of progression to colon cancer, called aberrant crypt foci (ACF), indicate that epigenetic mechanisms might induce abnormal Wnt-pathway activation even before the appearance of mutations in the pathway. These foci contain pre-adenomatous, pre-malignant hyperplastic cells that are derived from individual colon epithelial villi. Most ACF cells do not contain mutations in genes that would abnormally activate the Wnt pathway, although they might acquire these genetic changes during tumor progression. Studies of ACF cells have revealed abnormal methylation in the promoter regions of members of the SFRP gene family. SFRP proteins have homology to the frizzled proteins, which are receptors for the WNT secreted signalling proteins. SFRPs function at the cell membrane to antagonize Wnt-pathway activation. Methylation of SFRP-gene promoter regions was not only present in all ACFs that were examined, but persisted for one or more of the genes in almost all primary colon cancer cells that were examined. Furthermore, re-expression of SFRPs in colon cancer cells that

have silenced these genes blocks Wnt signaling and results in apoptosis. Importantly [293], these events occurred in colon cancer cells that harbor the key inactivating or activating mutations in downstream factors of the Wnt pathway, and are also believed to contribute to the earliest stages of colon tumorigenesis. These Wnt-pathway genes have been called "gatekeeper", factors. Inactivation of these factors is thought to be required for these cells to begin the process of colon cancer tumorigenesis. This activation would contribute to abnormal expansion of colon epithelial stem- or progenitor-cells that rely on the Wnt pathway for proliferation, as opposed to differentiation. These cells are then effectively addicted to the overactivity of the Wnt pathway, and can later acquire further mutations in other factors that lie downstream in this pathway, such as APC and others [293] (Figure 45). This combination of epigentic and genetic events fully activates the Wnt pathway to promote tumor progression. Without the epigenetic events that silence the SFRP genes, mutations that disrupt the APC complex might not be sufficient to promote tumorigenesis or tumor progression.

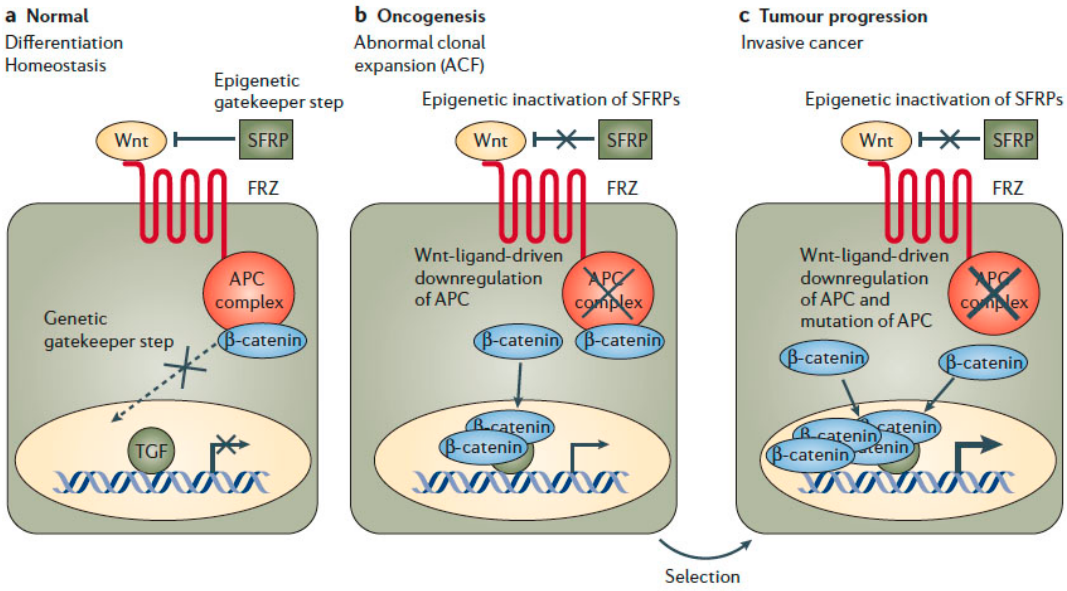


Figure 45. Addiction to the Wnt signalling pathway. A) In normal colon epithelial cells, secreted frizzled-related proteins (SFRPs) function as antagonists of Wnt signaling. When Wnt signalling is inactive, the adenomatosis polyposis coli (APC) complex phosphorylates β-catenin, leading to its degradation. This prevents the accumulation of nuclear β-catenin and therefore its ability to engage its transcription factor partners (TGF), which results in the differentiation and homeostasis of colon epithelial cells. Expression of APC is therefore a genetic gatekeeper step. B) When SFRP expression is lost, through epigenetic silencing of the gene that encodes it, Wnt signaling becomes activated through the receptor FRZ. This Wnt signaling potentially inactivates the APC complex, allowing β-catenin to accumulate in the cytoplasm and eventually in the nucleus. In the nucleus, β-catenin activates transcription of genes such as MYC, cyclin D and other genes whose products promote cell proliferation and survival rather than differentiation. This results in the expansion of colon epithelial stem and progenitor cells and formation of atypical crypt foci (ACF). C) Persistent activation of the Wnt pathway allows mutations to occur in other pathway components, such as those that permanently disable the APC complex and promote nuclear accumulation of β-catenin. These cells are selected for because of their survival and proliferative advantages. Baylin and Ohm, 2006.

8.3. Epigenetic reprograming during EMT

Epithelial-to-mesenchymal transition (EMT) is a good example of cell plasticity that is also likely to be regulated through epigenetic mechanisms, although this is still poorly understood. Genome scale epigenetic analysis of modifications during EMT mediated by TGF-β [294], revealed a global reduction in the heterochromatin mark H3 Lys9 dimethylation (H3K9Me2), an increase in the euchromatin mark H3 Lys4 trimethylation (H3K4Me3) and an increase in the transcriptional mark H3 Lys36 trimethylation (H3K36Me3). These changes depended largely on lysine-specific demethylase-1 (LSD1), and loss of LSD1 function had marked effects on EMT-driven cell migration and chemoresistance, in part, through stabilizing SNAI1. Genome-scale mapping showed that chromatin changes were mainly specific to large organized heterochromatin K9 modifications (LOCKs), which suggests that EMT is characterized by reprogramming of specific chromatin domains across the genome.

Being the downregulation of E-cadherin one of the hallmarks of EMT, it is worth mentioning that different epigenetic factors play a key role in the repression of E-cadherin. The Polycomb complex 2 (PRC2) component EZH2 mediates transcriptional silencing of the tumor suppressor gene E-cadherin by trimethylation of H3 K7. Histone deacetylase inhibitors can prevent EZH2 mediated repression of E-cadherin and attenuate cell invasion, suggesting a possible mechanism that may be useful for the development of therapeutic treatments [157]. Moreover, it has been shown that SNAI1 is required PRC2 for E-cadherin repression. SNAI1 promotes the recruitment of SUZ12 (another PRC2 component) to the E-cadherin promoter and the subsequent trymethylation of H3K27, with an enrichment of the repressive mark H3K27m3 around the E-cadherin promoter [158]. A prolonged transcriptional repressive status favors the DNA methylation of the E-cadherin promoter, which stabilizes it repressive state [295].

8.4. Epigenetics in stem cells and cancer stem cells

Epigenetic processes, thus regulate the balance between stem, progenitor and mature cells in adult and developing tissues as well as foster, abnormal imbalances in cell population states, such as in cancer. It is increasingly apparent that genes that are important in development overlap with those that govern adult renewal systems and whose expression patterns are altered in cancer. Most of these genes are subject to a stringent epigenetic control of their transcription [296].

Molecular events that lock in a degree of “stemness” in neoplastic cells can drive tumor progression. Normal ES and progenitor cells use chromatin organization to maintain their status. In both murine and human ES cells, transcription factors specifying for cell stemness, including OCT4, NANOG, and SOX2, are localized to promoter regions of a restricted group of some 1,000 genes for each factor, and some 350 target genes for all three. Importantly, the target genes are largely maintained at low expression states in ES cells by having a zone spanning their proximal promoter regions characterized by PcG proteins and the pivotal repressive

mark, H3K27me3. Also, such genes usually have proximal promoter CpG islands and sequences outside gene coding regions that are conserved between mouse and man. The PcG occupancy of the above promoters appears to function for preventing their full expression until ES and other precursor cells are signaled to undergo commitment steps toward cell lineages [297]. In the subsequent maturation steps, or in adult cell differentiation, the role of PcG complexes may come back into play in that their maintenance of long-term gene silencing is seen in mature stem/precursor cells as well.

A defining chromatin feature of the above PcG target genes in ES cells is the ocurrence of bivalent marks, within the broad zone of PcG localization surrounding their promoters, marked by the distribution of H3K27me3, and a narrower zone distinctly marked by the presence of the activating mark, H3K4me3. This has been termed as "bivalent chromatin," which is hypothesized as essential to maintain certain ES genes at a low expression level, but prepared for rapid eventual upregulation as needed for cell lineage commitment [297]. Therefore, a carefully orchestrated, plastic state of gene expression is maintained by PcG proteins and their induced chromatin marks to govern the balance between maintenance of stem cell phenotype and cell differentiation during embryonic development. The chromatin of DNA hypermethylated and silenced DNA in cancer may be remarkably similar to the above bivalent chromatin. These genes start with a heritable, silenced state associated with chromatin consisting of highly repressive marks. After the induction of DNA demethylation, the chromatin of the re-expressed genes does not fully return to the fully activated state but, rather, active marks are restored while repressive marks, including the PcG related mark, H3K27me, are prominently retained [289]. In cancer-predisposing events, abnormal pressure for stem/progenitor cell proliferation with retained bivalent chromatin may allow polycomb proteins and/or marks to recruit other silencing marks such as H3K9me2 and H3K9me3 and DNMTs. The promoter evolves abnormal DNA methylation and a tight heritable gene silencing, which results in loss of function for genes. Tumors may arise in such clones with subsequent progression steps [289].

8.5. The epigenetic origin of cancer

The fact that epigenetic changes are found so early in tumorigenesis, and even in apparently normal tissues before tumors arise, indicates to us that early epigenetic changes in stem cells might provide a unifying view of cancer aetiology. The epigenetic disruption of progenitor cells is a key determinant not only of cancer risk, but of tumor progression and heterogeneity late in the course of the tumors that arise from these cells. Epigenetic changes can provide mechanistic unity to understanding cancer, since they can occur earlier than genetic mutations and set the stage for genetic alterations, and have been linked to the pluripotent precursor cells from which cancers arise. Importantly, early epigenetic changes could explain many of the heterogeneous properties that are commonly associated with tumor cell-growth, invasion, metastasis and resistance to therapy. To integrate the idea of these early epigenetic events, Feinberg et al. 2006, proposed that cancer arises in three steps: an epigenetic disruption of progenitor cells, an initiating mutation, and

genetic and epigenetic plasticity [8] (Figure 46).

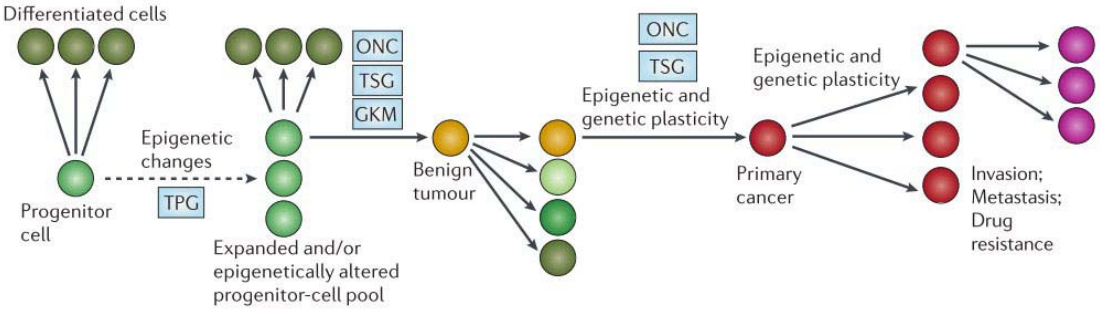


Figure 46. The epigenetic progenitor model of cancer. First is an epigenetic alteration of stem/progenitor cells takes place within a given tissue, which is mediated by aberrant regulation of tumour-progenitor genes (TPG). This alteration can be due to events within the stem cells themselves, the influence of the stromal compartment, or environmental damage or injury. Second is a gatekeeper mutation (GKM) occurs (tumour-suppressor gene (TSG) in solid tumours, and rearrangement of oncogene (ONC) in leukaemia and lymphoma). Although these GKMs are themselves monoclonal, the expanded or altered progenitor compartment increases the risk of cancer when such a mutation occurs and the frequency of subsequent primary tumors (shown as separately arising tumors). Third, genetic and epigenetic instability develope, which leads to increased tumor evolution. Note that many of the properties of advanced tumors (invasion, EMT, metastasis and drug resistance) are inherent properties of the progenitor cells that give rise to the primary tumor and do not require other mutations (highlighting the importance of epigenetic factors in tumor progression). Feinberg et al. 2006.

8.6. MicroRNAs in tumorigenesis and metastasis

MicroRNAs were originally discovered because of their roles in controlling the timing of Caenorhabditis elegans larval development. Less than a decade later they were identified in plant and mammalian cells [298]. miRNAs are initially synthesized by RNA polymerase II as long primary transcripts, which are subsequently capped and polyadenylated. These transcripts are processed into ~70 nucleotide stem-loop pre-miRNAs by Drosha RNase III endonuclease and are transported out of the nucleus by Ran-GTP/exportin5. Pre-miRNAs are further processed by Dicer in the cytoplasm to yield a ~21-23 nucleotide duplex. One strand of the duplex is incorporated into the RNA-induced silencing complex (RISC) and is used to regulate the expression of target genes. Binding of miRNAs to the 3' UTR of mRNAs with perfect or near-perfect complementary sequences induces mRNA degradation, whereas imperfect complementarity often induces translational repression. The seed sequence of miRNAs, representing 7-8 nucleotides in the 5' end, is critical for efficient targeting and miRNAs harboring similar seed sequences can theoretically regulate the expression of a similar subset of genes [181].

miRNAs have also been documented as contributors to cancer pathology. More than 50% of human miRNAs are located in fragile chromosomal regions that are prone to mutations during tumor progression. Importantly, functional characterization has revealed a role for miRNAs as oncogenes (miR-21, miR-155, miR-17-92 cluster) or tumor-suppressor genes (miR-34a, let-7, miR-15a, miR-16) and other miRNAs with both abilities dependent of the cell context (miR-146 and miR-29)

[299] through the silencing of target tumorsuppressor or oncogenic protein-coding genes, respectively. Interfering with miRNA processing also has been shown to enhance experimental tumorigenesis, further confirming the role of miRNAs in cancer. Although there is a clear role of miRNAs in oncogenesis, the contribution of miRNAs to malignant progression of human tumors has only recently been investigated and characterized.

miRNAs also play a key role along the metastasis cascade, as promoters or suppressors Tavazoie et al. [300] performed array-based miRNA profiling of the MDAMB-231 human breast cancer cell parental population as well as its highly bone- or lung-metastatic derivatives to uncover the role of miR-335 and miR-126 as metastasis suppressors in human breast cancer. They found that miR-126 reduces the overall tumor growth and proliferation and miR-335 inhibits cells invasion metastasis through targeting SOX4 and the ECM component tenascin C. Others laboratories have identified pro-metastatic miRNAs in breast cancer, such as miR-10b. Other miRNAs, like miR-373 or miR-520c, also promote the metastatic potential of the miRNAs non-metastatic MCF-7 cell line. Thus, many miRNAs may be involved in different steps of the metastasis cascade [298] (Table 2).

	Primary growth	EMT	Adhesion migration invasion	Apoptosis	Angiogenesis	Colonization
Pro metastatic	Multiple tumor-specific	miR-10b	miR-10b miR-21 miR-143 miR-182 miR-373 miR-520c miR-183	miR-29a/b/c miR-182	miR-27a/b miR-19a/b miR-221 miR-222 let7f	
Anti metastatic	Multiple tumor-specific	miR-141 miR-200a/b/c miR-205 miR-429	miR-146a/b miR-206 miR-335 miR-31	miR-31	miR-15b miR-16 miR-20a/b miR-126	miR-31

Table 2. Critical steps of metastasis and influence of Metastamirs. Modified from Hurts et al. 2009.

miRNAs in EMT regulation. The cytokine TGF-β is known to be involved in many processes that promote EMT, several of which are regulated by miRNAs, among which is the inhibitory activity of the miR-200s family (Cluster I: 200a, 200b and 429; Cluster II: 200c and 141), one of the best studied miRNAs families and widely affected in different types of cancer. Experimental inhibition of miR-200 induces a mesenchymal-like spindle cell morphology and cell migration, accompanied with an increase in the expression of ZEB1. Furthermore, loss of miR-200s correlates with a lack of E-cadherin expression in invasive breast cancer cell lines and in breast tumor specimens. Indeed, overexpression of the individual members or separate clusters of miR-200s represses EMT by directly targeting and downregulating ZEB1 and ZEB2 via miR-200-binding sites located within their 3'UTRs, and thus resulting in enhanced E-cadherin expression. miR-200s also regulate E-cadherin through their targeting of SUZ12 (a PRC2 component) which in turn mediates the repression of E-cadherin by accumulating H3K27me3 marks at the E-cadherin promoter [301]. Moreover, TGF-β is inhibited by a member of the cluster II, miR-141. Thus, the expression of the miR-200 family of miRNAs in mesenchymal cells promote the reversion of a mesenchymal phenotype into an epithelial

phenotype, thus being a positive mediator of mesenchymal-epithelial transition (MET) [302]. miR-200s also directly target β-catenin, suppressing β-catenin/Wnt signaling. The expression of miR-200s inversely correlates with one of the β-catenin downstream targets, cyclin D1, thereby promoting cell proliferation [302].

Other miRNAs like miR-10b have been related as EMT inducers and pro-metastatic miRNAs. miR10b is upreglated by TWIST1 and targets HOXD10. miR-155 favors epithelial cell plasticity in EMT by targeting RHOA. miR-21, upregulated by ZEB1 and others, targets many EMT suppressor genes inhibiting them. miR-373 and miR-520c, upregulated by BRMS1, promote migration and targets CD44 [298]. Other miRNAs counteract EMT, such as: mir-205, mir-30a and other miRNAs, by targeting EMT transcriptional factors like SNAI1, ZEB1, ZEB2 and others [304]. miR-193 inhibiting uPA; miR-222 inhibiting MMP1 and SOD2; miR-31, by inhibiting FZD3, ITGA5, MMP16, and RHOA. The list of miRNAs favoring or counteracting EMT is increasing fast [305].

miRNAs in stemness regulation. A subset of miRNAs is preferentially expressed in undifferentiated stem cells [305] and plays essential roles in proliferation, pluripotency, and differentiation. Genetic disruption of the essential miRNA biogenesis factors Dicer and DGCR8 in the mouse results in numerous defects in ES cell behavior, including impaired self-renewal, diminished expression of pluripotency markers, and a failure to execute induced programs of differentiation [307]. The best characterized miRNA network involves the ES-specific miR-290 cluster (in mouse), which regulates Oct4-methylation in differentiating ES cells [306]. miRNAs play an indirect role in the control of de novo DNA CpG methylation in differentiating ES cells. The miR-371 cluster (expressed in humans) and miR-302 cluster (expressed in mouse and humans) are direct regulators of cell cylce of ESCs, i.e. suppressing cyclin D1 [309], a common target of miR-200, mentioned above. However, others suggest miR-200c as mammosphere forming suppressor [331]. MYC regulated microRNAs attenuate ESC differentiation [332]. The promoter regions of miRNAs that function in ESCs, miR-302-367 and mir-371 are typically occupied by key members of the stem cell transcription factor network, including NANOG, OCT4, and SOX2 [308]. Interestingly, NANOG, OCT4, and SOX2 are themselves subject to miRNA regulation including, miR-134-296-470 [310], and also to PcG modulation, during ES cell differentiation [322].

miR-145 is low in self-renewing human embryonic stem cells (hESCs) but strongly upregulated during differentiation. Interestingly, human tumors and cancer cell lines from colorectal, breast, prostate, and cervical carcinomas frequently exhibit reduced expression of miR-145 [307]. A recent study identified the pluripotency factors OCT4, SOX2, and KLF4 as direct targets of miR-145 and showed that endogenous miR-145 represses the 3'-UTR regions of them [312]. Thus, increased miR-145 expression inhibits hESC self-renewal, represses expression of pluripotency genes, and induces mesoderm and ectoderm differentiation determined by the upregulation of mesenchymal markers such as Vimentin and MIXL1 [312]. Furthermore, that study found that the miR-145 promoter is bound and repressed by OCT4 in hESCs, revealing a direct link between OCT4, SOX2, KLF4, and miR-145 as

a double-negative feedback loop involving miR-145 and this core of reprogramming factors, which in turn induce miR-302 and miR-371 [307].

Let-7 family members are post-transcriptionally inhibited by LIN28, a promoter of pluripotency and used to reprogram human fibroblast cells into iPSC. LIN28 is dispensable in cell reprograming because can be substituted by MYC, which in turn, is known to negatively regulate let-7 family members. Thus let-7 may repress self-renewal and induce diferentiation leading to a loss of stemness, in both normal development and cancer [311]. Indeed, while miR-205 (above mentioned in its capacity to repress ZEB factors) and miR-22 are highly expressed in self-renewing mammary progenitor cells able to reconstitue the mammary gland, it has been shown that let-7 and miR-93 are depleted in such self-renewal progenitor populations and recover its expression in differentiated compartments. Forced expression of let-7 induced a loss of self-renewing cells [313] and several imporant oncogenes have been identified as the targets of let-7, including RAS, MYC and HMGA2 [305].

miR-34a has been reported to directly repress CD44, which is considered a major CSC biomarkers in many different type of cancers like breast, pancreas, head and neck, liver, stomach bladder ovary and prostate. The repression od miR-34a over CD44 affects clonogenic expansion, tumor regenerations and metastasis [314]. Other studies already reported how miR-34a repress important genes for proliferation like cyclin D1 and genes involved in self-renewal and pluripotency like Myc and Notch among others. However, whether donwregulation of CD44 is relevant is still controversial, other miRNAs targeting CD44, such as miR-373 and miR-520, are metastasis-promoting miRNAs, and CD44 has been found lost in breast cancer metastasis and acting as metastatic suppressor in prostate and colon cancer [305]. Other reports involve p53-induced miR-34a as the mediator of p53 suppressor effects on somatic cell reprogramming to a more pluripotentn state, since miR34 can repress genes like NANOG, SOX2 and MYC [315].

Collectively, these reports highlight the emerging view that miRNAs regulate, and are themselves regulated by, distinct transcriptional repressive mechanisms that directly impact on the invasive phenotpe and on stemness.

9. Cancer stem cells and epithelial mesenchymal transition in tumor cell populations and non-tumoral cell populations

A new paradigm was formulated in 2008 proposing that epithelial-mesenchymal transition itself can generate cells with stem cell-like properties [197]. Although many other studies support this hypothesis [227] [228], we will argue here that the emergence of opposing evidences warrant further studies that address possible contradictions, addressing epithelial-mesenchymal transition as a possible metastasis supressor and not linked with self-renewal in some cellular systems [27] [61] [229] [230] [351].

The studies by Mani et al. and Morel et al. showed that $CD44^{high}/CD24^{low}$ populations isolated from transformed normal human mammary epithelial cells $CD44^{high}/CD24^{low}$ populations display stronger mesenchymal phenotype and markers than $CD44^{high}/CD24^{high}$ populations, and that such $CD24^{low}$ cells show a better ability to form mammospheres and are more tumorigenic than their $CD24^{high}$ counterparts. These studies assume that $CD44^{high}/CD24^{low}$ breast cancer cells represent bona fide stem cells, an assumption that is supported by prior reports. However, there is no definitive evidence that all breast stem cells are detected by the combination of these two cell surface markers, or that stem cells in other epithelial tissues share this feature. In addition, it is not clear that all cancer stem cells in breast or other epithelial carcinomas are $CD44^{high}/CD24^{low}$. Thus, the definition of stemness assumed by Mani et al. and Morel et al., although most likely adequate for one particular type of normal or cancer stem cells, does not necessarily exclude the existence of other cells harboring cancer stem cell features and not fitting that definition. In fact, stem cell features have been attributed by others to $CD24^{high}$ populations in some epithelial tissues [27] [28] or carcinomas such pancreatic and ovarian [284]. Importantly, CD24 has been found to form part of embryonic stem cell-like gene expression signatures in poorly differentiated aggressive human tumors, including breast tumors [90].

The major distinctions that characterize different studies that address this issue are the type of tumor studied and the particular subpopulations analyzed, defined by the markers used to define such subpopulations. In pancreatic carcinoma, there is evidence of the existence of two different stem cell-like populations: one that maintains the growth of the primary tumor and another that produces metastatic growth [324]. It will be interesting to determine whether the latter have undergone EMT and whether these cells originate from resident cancer stem cells or other carcinoma cells. In the cited Hermann et al. paper [324] the authors chose CD133 as the marker molecule to enrich for TIC populations from pancreatic cancer samples and an established cell line (L3.6pl). The $CD133^{+}$ subpopulation injected into mice in their experiments were initially cytokeratin negative, but the tumors that these cells formed in the animals and were found to be epithelial cytokeratin positive. Sequential passages showed that secondary xenografting resulted in faster growth of tumors, which Hermann et al. interpret this as evidence of heterogeneity within the $CD133^{+}$ population and the imposition of processes in vivo that favor the

selection of more aggressive subpopulations. Hermann et al. [324] do not provide further characterization as to the epithelial or mesenchymal nature of these secondary tumors.

The Pang et al. [326] describe the use of 3 markers, CD26, CD133 and CD44, to isolate and characterize tumor cells with different tumorigenic (subcutaneous implantation) and metastatic (cecal wall implantation) potentials from colorectal cancer samples. They found that several subpopulations expressing different marker combinations have tumor initiating capacity, but only those expressing CD26 can metastatize from their orthotopic implantation sites. These $CD26^+$ cells with metastatic capacity are still heterogeneous for CD133 and CD44, and they could also be heterogeneous with respect to their expression of E-cadherin (as a measure of how epithelial these populations are). Pang et al. show that $CD26^+CD133^+$ cells express less E-cadherin than $CD26^-CD133^+CD44^+$ cells. Unfortunately, in their paper data are not shown for comparative E-cadherin levels in different $CD26^+CD133^+$ subpopulations ($CD26^+CD133^+CD44^+$ vs. $CD26^+CD133^+CD44^-$), and therefore we do not have information as to whether the most strongly metastatic subpopulation ($CD26^+CD133^+CD44^+$) is more epithelial. Also, they did not perform tumor heterogeneity reconstitution experiments, in which subpopulations separated on the basis of surface markers and with distinct properties are later combined to study effects on tumor initiation or metastatic capacities. Nor did they did perform studies in which sorted $CD26^+CD133^+$ cells were mixed with human intestinal fibroblasts, to determine if that combination enhanced metastatic capacity from orthotopic implantation sites.

A recent study [249] shows that one needs to consider specific isoforms when attempting to use CD44 as a marker to identify and/or isolate cancer stem cells. The study reveals splice isoforms of CD44 produced by alternative splicing determine the EMT or non-EMT status of cells. Using both *in vitro* and *in vivo* systems, the authors demonstrated a shift in CD44 expression from variant isoforms (CD44v) to the standard isoform (CD44s) during EMT. This isoform switch to CD44s was essential for cells to undergo EMT and was required for the formation of breast tumors that display EMT characteristics in mice. This implies that in order to isolate appropiate subpopulations, antibodies recognizing specific isoforms of CD44 should be used.

9.1. EMT in non-tumoral cells

Thus, further studies in normal tissues appear to be needed to corroborate the hypothesis that EMT itself induces a CSC state. On the other hand, mesenchymal tissues contain mesenchymal stem cells (MSCs), as adult stem cell pools sustaining mesodermal tissues. It has been shown that these cells derive from bone marrow cells and are pluripotent progenitor cells that contribute to the maintenance and regeneration of a variety of connective tissues, including bone, adipose, cartilage and muscle [273]. The contributions of MSCs to tissue formation become apparent only in cases of tissue remodeling after injury or chronic inflammation. Since carcinomas can be viewed as "wounds that never heal", it has been suggested that a

subpopulation of tumor-associated fibroblasts akin to MSCs are present in the stroma of tumors and instruct cancer cells in the primary tumor to enhance their metastatic ability by promoting migration and extravasation [15]. Interestingly, these stromal MSCs do not induce EMT, indicating that this process occurs independently in carcinoma cells. Another study showed that skin fibroblasts can be differentiated to hepatocytes upon different conditions [327]. Those studies support the notion that pool of mesenchymal cells can be maintained in undifferentiated states.

The existence of MCSs in different normal and pathological tissues has been demonstrated beyond reasonable doubt. It is important that they are not confused with cancer cells that suffer an EMT. MSCs derived from plurpotent embryonic cells to give rise to mesechymal tissues differentiating a variety of cell types. Cancer cells that undergo EMT, are cells which were initially malignant epitehlial cells with variable degrees of differentiation that have transformed into malignant mesenchymal cells through a drastic gene program and phenotypic switch in oder to acquire motility. This convertion into a mesenchymal phenotype, does not necessarly imply the reprogramming to a more pluripotent state. Nevertheless, recent studies have posed a new conundrum in which the maintenance of mesenchymal states would be incompatible with cell reprogramming and acquisition of pluripotency. In 2010, two independent studies by Li et al. [229] and Samavarchi-Tehrani et al. [328] showed that a mesenchymal-to-epithelial transition is a critical initiating event during the generation of induced pluripotent stem cells (iPSCs) from fibroblasts (Figure 47). And, in 2011, another interesting study reported that the expression of E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming [68]. Furthermore, other studies, reported in 2009 and 2010 found that TGF-β inhibitors accelerate or improve the reprograming, replacing two of the 4 transcription factors (OKSM) required to transform fibroblasts into iPSCs [329]. It follows from these recent evidences that, during cellular reprogramming, the acquisition of an epithelial fate is tightly linked to the emergence of a pluripotent state and may be a reflection of requirements for cell-cell interactions that initiate and sustain pluripotency [328].

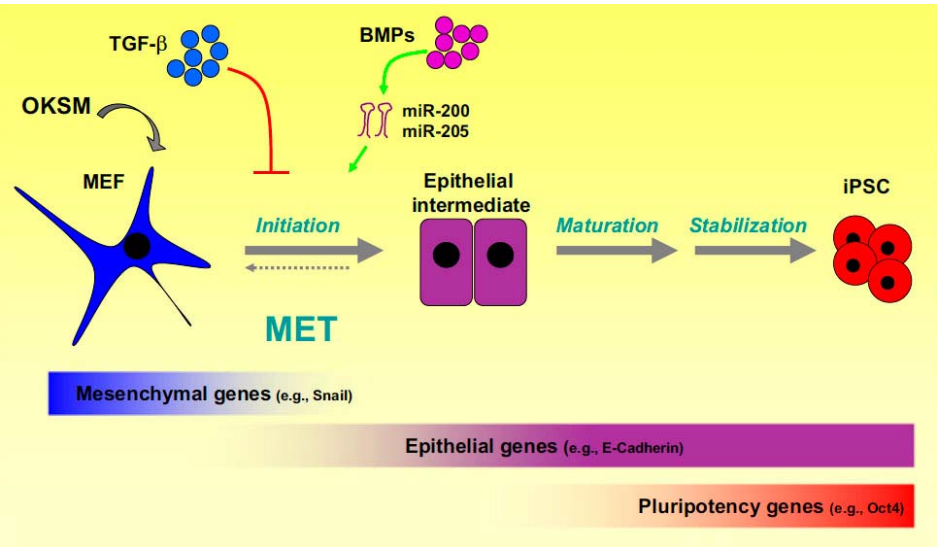


Figure 47. Fibroblast reprogramming into iPSCs entails a mesenchymal-to-epithelial transition. The initiation phase of reprogramming resembles a mesenchymal-to-epithelial transition (MET), generating unstable intermediate cells with epithelial characteristics that revert into fibroblastic cells upon loss of OKSM expression (indicated by dashed reverse arrow). The pluripotency network becomes activated and solidified during the maturation and stabilization phases, respectively. First step: loss of EMT related genes; Second step: gain of E-cadherin expression accompanied with other typical epithelial genes; and Third step: the stabilization of the pluripotent phenotype with the expression of OKSM genes. Polo and Hochedlinger, 2010.

In their work, Li et al. [229] demonstrated that the generation of iPSCs from mouse fibroblasts requires a mesenchymal-to-epithelial transition (MET) orchestrated by suppressing pro-EMT signals from the culture medium and activating an epithelial program inside the cells. At the transcriptional level, Sox2/Oct4 suppressed the EMT mediator Snail, c-Myc downregulated TGF-β1 and TGF-β receptor 2, and thus preventing them to be reactivated by TGF-β, and Klf4 induces epithelial genes including directly induction of E-cadherin expression [229] (Figure 48). Blocking MET impaired the reprogramming of fibroblasts, whereas preventing EMT in epithelial cells cultured with serum could produce iPSCs without Klf4 and c-Myc. This study not only established MET as a key cellular mechanism toward induced pluripotency, but also demonstrated that the generation of iPSCs is the result of a cooperative process between the defined factors and the extracellular milieu.

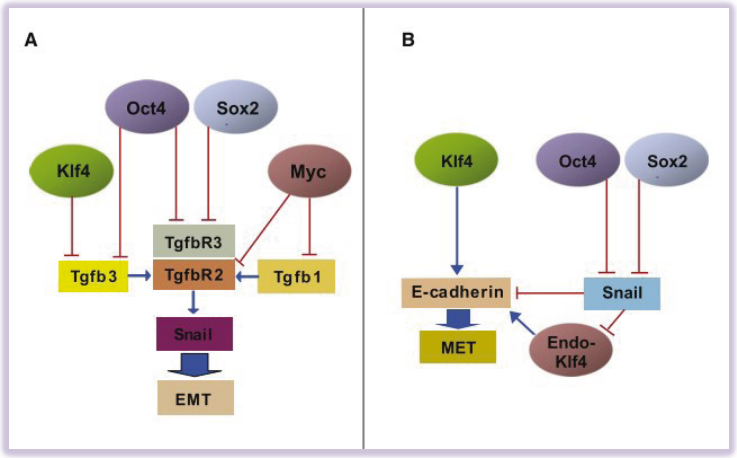


Figure 48. Molecular network for the suppression of EMT and activation of MET. A) Schematic representation of how exogenous factors suppress TGF-β signaling to prevent the self-perpetuation of EMT in the reprogramming fibroblasts. B) Representation of the suppression of Snail by Oct4 and Sox2 and the activation of an epithelial program by Klf4 during the reprogramming of fibroblasts. Li et al., 2010.

Samavarchi-Tehrani et al. [328], used systematic RNAi screening to further uncover a key role for BMP signaling and the induction of mesenchymal-to-epithelial transition (MET) during the initiation phase of fibroblast reprogramming. Importantly, they found that reprogramming is linked to the BMP-dependent induction of miR-205 and the miR-200 family of microRNAs, shown by others to be key regulators of MET (Figure 36). These studies thus define a multistep mechanism that incorporates a BMP-miRNA-MET axis during somatic cell reprogramming. This is in contrast with a report describing how ZEB1 confers stem cell-like properties through inhibition of miR-200 [227].

All these very recent studies associate MET and E-cadherin with a pluripotent state and remark the escape of EMT phenotype as q requisite to acquire the

stabilization of iPSCs. This stands in contrast to the ideas previously described by Mani et al. in transformed mammary cells, because such studies provide preliminary mechanistic insights into how pluripotent-Yamanaka factors can supress SNAI1 expression, either at the transcriptional level or stability of the protein [229], in the case of KLF4, and the downregulation TGF-β receptors. Wether these ciruits, that regulate the reprogramming and pluripotency of adult normal cells, also function in cancer cells is yet to be explored.

One study showed that NANOG blocks BMP-induced mesoderm differentiation of ES cells by physically interacting with Smad1 and interfering with the recruitment of coactivators to the active Smad transcriptional complexes. NANOG also maintains pluripotency by inhibiting typical EMT pathways such as NF-κB in cooperation with STAT3 [59] [316]. An additional study showed that the treatment with the GSK3 inhibitor BIO prevented EMT, upregulated E-cadherin and downregulated EMT genes such SNAI1, SNAI2, MMPs or Vimentin, thus mantaining an epithelial undifferentiated phenotype for hESCs expressing pluripotent transcription factors [61].

On the other hand, studies with nontumorigenic normal mammary epithelial gland cells, such as the MCF10A and MCF12A cell lines, suggest that these cells have an intrinsic capacity for EMT, and the induction of a mesenchymal-like phenotype does not correlate with the acquisition of global stem cell/progenitor features in these cells. Based on their findings, it is proposed that EMT in normal basal cells and claudin-low breast cancers reflects aberrant/incomplete myoepithelial differentiation, concluding that the majority of stem cell/progenitor properties are associated with an epithelial state not with a mesenchymal-like phenotype [317]. In summary, several studies, in particular using embryonic stem cells, suggest the ocurrence of dissociation between self-renewal and EMT, but none thus far in which this dissociation is unequivocally demostrated in cancer cells.

9.2. CSC and MET in tumors: distinct, overlapping or same populations

In a Darwinian view of cancer cell evolution, many phenotypes reflecting the expression of various programs, reversible to irreversible, exclusive, overlapping or linked coexist and interact or compete with each other. Regarding the two major gene programs active in tumor cells, self-renewal and invasive, a prevailing hypothesis considers these two programs as tightly linked or overlapping, for which there is evidence in some in vitro models. As noted above, studies by Mani et al. and Morel et al. and other studies support that cancer cells progress in their malignant evolution through the shared expression of EMT and CSC properties. As also summarized above, using protocols to enrich for breast CSCs, these cells showed evidence of EMT [228]. Likewise, isolation of cells expressing putative CSC markers such as ALDH, showed both migratory behavior and metastastic potential. Finally, a significant number of studies indicate that both programs (CSC and EMT) can share different activating signals and pathways, including TGF-β, RAS or EGF signaling pathways, or hypoxia [27]. Additional studies imply a common regulation of EMT and CSC states by ncRNAs. Specifically, it has been reported that the miR-200 family

microRNAs abolish the EMT program and, at the same time, stemness properties, through a positive feedback loop involving ZEB1 [227]. However, other studies suggest precisely the opposite consequences of excessive miR-200 activation, namely, a gain in epithelial features together with an enhanced metastatic potential [330-332] [350].

In spite of the above evidences, the association of EMT properties with cell proliferation is not necessarily obvious. Indeed, EMT is long known to generally inhibit cell proliferation. On the other hand, a significant fraction of the recent literature suggests that CSC-TPCs are highly proliferative and competitive [27], which stands in contrast to the inhibitory effects of EMT on cell proliferation. This issue can be resolved by considering that the two properties, proliferation/self-renewal and invasion-prone EMT, as two states occurring at different times in the same cell or cell population. If, as generally assumed, the EMT state represents a transient phase in the lifetime of the cells, it is conceivable that the most competitive CSCs, when separated from other cells, transiently adopt an EMT state that allows them to invade at their primary site, metastasize and later on reconvert to their previous highly proliferating state at their secondary site. In this case, some biomarkers of CSC-TPC or EMT would only be expressed at some stages of this process [27].

Further arguments in favor of self-renewal and EMT as alternative cell states come from studies in cell reprogramming, as discussed above. In addition to the evidence that reprogramming of somatic cells into iPSCs requires a mesenchymal-to-epithelial transition, a recent study has demonstrated that the EMT factor SNAI1 induces EMT in ESCs and promotes their differentiation towards early mesoderm commitment, while suppressing their stemness. SNAI1 does this in part through the repression of miR-200, which inhibits ZEB factors and favor E-cadherin expression. Furthermore, miR-200 cooperates with activin to maintain cells in a transcriptional state akin to that of epyblast-like stem cells (EpiSC). Exit from the EpiSC-like state can be regulated through either the induction of Snail or the inhibition of activin, which induce an EMT concomitant with germ layer fate commitment toward mesoderm or neuroectoderm, respectively [318].

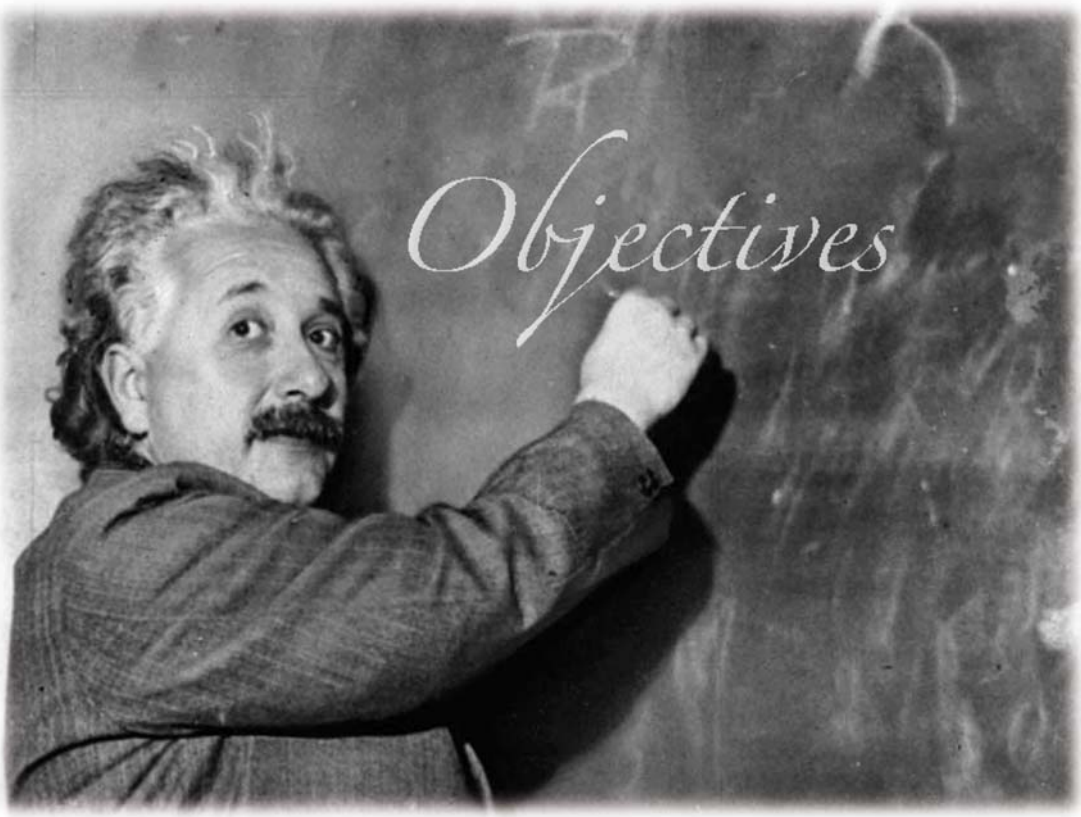
Additional evidences linking stemness to epithelial gene programs come from gene profiling analyses that have identified gene networks specific of stem cell phenotypes or properties, which importantly share transcriptional and post-transcriptional programs with epithelial phenotypes. These ESC-like [229] and ES gene modules [90] are represented in epithelial tumors and include some epithelial genes and remarkably, do not include any of the canonical EMT marker genes. A ChIP-seq and functional analysis of SOX2 gene in colorectal cancer revealed interesting targets related to EMT, including LOXL2, β-catenin, CDH2, FGFR1, SMAD3, ZEB2 which should prompt further studies to determine if SOX2 acts by activating or repressing them [319].

A proposal that attempts to reconcile several of these notions is that cancer cells may exhibit different but partially overlapping programs (for example, TPC, EMT, drug tolerant persisters (DTP) and so on) at different times during their

evolution, and that these programs manifest different rates of reversibility [27]. In this sense, EMT, as the intermediate stage between the solid tumor and invasion and metastasis [146], is a reversible and transient process. This is clearly different from the CSC-TPC state, which is, at best, a relatively stable and slowly reversible state.

From an experimental perspective, the characterization of CSC or EMT programs in tumor cells, and the study of their relative representation in tumor cell subpopulations, requires the isolation of relatively pure populations of presumptive CSC and their corresponding negatively selected counterparts. Putative CSC biomarkers are often used to characterize and isolate such subpopulations. However, the fact that these defining biomarkers show, at best, only partial overlaps between different cancer types suggests, at least, that these markers associate poorly with the core properties attributed to CSCs. As discussed above, some putative CSC markers, such as CD24, define cell populations that display opposite phenotypes in different cancer cells. The use of one criterion to define the CSC will, therefore, also favor the inclusion of different cell populations in this category. A survey of recent literature suggests that, as defined, CSCs could include or be contaminated by other cell types such as the cancer cells that are in EMT state. CSC and EMT are properties of cancer cells that need not be expressed in all cancer cells, nor at all times in a given cell [27]. Whether these populations overlap, *in vivo*, is still an open question. Anyway, the two cell programs CSC and EMT should overlap at least at some time at transitory states. Tumor cells that succeed at generating metastases would therefore be those CSC that have undergone EMT have escaped from the primary tumor site. Floor *et al.* [27] propose that CSC-TPC and EMT programs overlap in some cells at some point in time. The two programs may even favor each other but, at any given time, only a minor fraction of CSC-TPC would be in EMT.

As a speculation, one could envision tumor cells, each regulated by distinct programs, as displaying a sort of "collective intelligence" through competition or collaboration among that optimizes their chances of tumor expansion at the expense of the host. This behavioral independence would be acquired progressively with the progressive shift in gene programs during the development of the tumor, with or without the need to acquire genetic mutations. These concepts further underline the interest of multipronged therapeutic schemes simultaneously or sequentially targeting different gene programs [320].


Objectives

Albert Einstein (Ulm, 1879 - Princeton, 1955)

1) Identification of gene expression programs differentially activated in highly metastatic versus low metastatic tumor cell population

2) Study of mechanisms that modulate the aggressiveness of tumor cell subpopulations

3) Identification of genes of aggressiveness in human tumors

Global summary

Santiago Ramón y Cajal (Petilla de Aragón, 1852 - Madrid, 1934)

This Doctoral Thesis project is part of a broader research program whose global aim is to better understand key molecular and cellular events in tumor metastasis, in particular those associated with prostate cancer, in order to identify regulatory molecules, gene networks and biochemical pathways, to propose relevant biomarkers and, finally, to provide proof-of-concept validations for potential therapeutic targets. As a starting point, this particular project has relied on comparative analyses between isogenic tumor cells with high *vs*. low metastatic potentials. The results of this Thesis project are reflected in two different but closely related studies. Article 1, the main publication of this project, focuses on how phenotypic switches and interactions between distinct tumor cell subpopulations engage or suppress metastasis, and which are the genetic programs that may govern such transformations. Article 2 is more focused on specific molecular post-transcriptional regulatory mechanisms that influence the latter steps of metastatic colonization.

First study. *Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells* (Celià-Terrassa *et al*., J. Clin. Invest., in press).

The starting point of this study consisted of fairly comprehensive comparative analyses of dual-cell models in which tumor cell subpopulations with highly contrasting phenotypes were analyzed. The two models characterized in this study are PC-3Mc *vs*. PC-3/S, two clonal subpopulations derived from the PC-3 prostate cancer cell line, and TSU-Pr1-B2 *vs*. TSU-Pr1, derived from the T24 bladder cancer cell line. PC-3/Mc cells (strongly metastatic) are clonal derivatives of a highly metastatic subpopulation isolated by *in vivo* selection from PC-3 cells, and PC-3/S (poorly metastatic) were directly cloned *in vitro* as single-cell progeny from PC-3 cells [336]. TSU-Pr1 is a poorly metastatic bladder cancer cell line, which, after serial *in vivo* injections and bone metastasis isolations, yielded the TSU-Pr1-B2 population (strongly metastatic) [337] [338]. Remarkably, our study demonstrates that both dual-cell models share significant functional similarities among them and suggest a clear dissociation between (1) the capacity of these cells to autonomously invade in *in vitro* assays and (2) their capacity for self-renewal (anchorage-independent growth) *in vitro* and cause metastasis *in vivo*.

Different *in vivo* approaches, including orthotopic prostatic implantation, intramuscular injection, tail vein injection and intracardiac injections, were employed to confirm that PC-3/Mc cells were clearly much more tumorigenic and metastatic than PC-3/S cells. In agreement with their differential *in vivo* tumorigenic and metastatic potentials, PC-3/Mc cells showed a robust capacity for anchorage-independent growth and self-renewal, as shown in spheroid formation assays *in vitro*, as opposed to the very poor spheroid forming potential shown by PC-3/S cells. Remarkably, and in strong contrast, PC-3/Mc cells were significantly less invasive than PC-3/S cells in *in vitro* assays, in which cells were tested for their autonomous capacity to invade through extracellular matrix. The striking dissociation observed between metastatic-tumorigenic-spheroid growth *vs*. autonomous invasive behavior was contrary to our initial expectations. Results from a global transcriptomic analysis

comparing PC-3/Mc cells *vs*. PC-3/S cells partly provided a hypothesis that might explain these contrasting phenotypes. Interestingly, PC-3/Mc cells differentially expressed genes related to embryonic stem cell (ESCs) signatures, self-renewal capacity, cell cycle and DNA damage checkpoint regulators, and importantly, many epithelial markers, including E-cadherin, EpCAM or desmoplakin. On the other hand, PC-3/S cells expressed many mesenchymal genes and genes related to epithelial-mesenchymal transition (EMT), invasion, activated secretory pathways, cytokines and inflammatory processes, and displayed conspicuously low levels or undetectable expression of proliferation and epithelial genes.

Thus, these functional *in vitro* and *in vivo* approaches, together with transcriptomic and gene expression approaches, revealed a clear and striking functional and gene program dissociation between an "EMT phenotype", expressed by PC-3/S cells with a high *in vitro* autonomous invasive capacity (but poorly metastatic), and a "CSC or TIC phenotype", expressed by PC-3/Mc cells with high self-renewal, tumorigenic and metastatic potentials, but poor autonomous invasive potential *in vitro*. Very similar functional and transcriptomic results were garnered for the TSU bladder cancer dual-cell model, with TSU-Pr1-B2 cells displaying a metastatic epithelial-CSC phenotype and gene program, in contrast to the mesenchymal-invasive phenotype and gene program of TSU-Pr1 cells.

As discussed in the introduction section, a major focus of study of the metastasis problem is to understand the mechanisms by which tumor cells escape the local environment and colonize distant organs [5] [146] and what barriers pose a rate-limiting step to finally achieve metastasis. In this regard, it has been proposed, based mainly on breast cancer models, that the engagement of an EMT program simultaneously leads, through mechanisms not yet fully elucidated, to the acquisition of a self-renewal program endowing tumor cells not only with the capacity to invade through tissues but also to survive and colonize distant sites [306]. However, our initial observations, summarized above, appear to stand in clear contrast to such paradigms, suggesting the occurrence of alternative interactions between self-renewal and EMT gene programs and phenotypes, perhaps depending on specific regulatory circuits that act very differently in different cell types or lineages.

Since tumors are composed of heterogeneous tumor cell populations and this heterogeneity can be under dynamic transitions [339], the determination of such populations in cell lines and from actual human tumors would help us to acquire a better knowledge of their prevalent functional and gene program characteristics, and the relationships among different tumor cell populations. Unfortunately, the isolation of cancer stem-like cells or other tumor cell subpopulations with the assistance of biomarkers it is still fraught with uncertainty and controversy because of the high variability of many purported CSC markers between different cancer types and patients [28]. By using some of the markers, previously proposed as associated with CSCs, we characterized PC-3/Mc cells as $CD44^{+}$, $CD24^{+}$, $CD40^{-}$, $CD71^{+}$ and $CDH1^{+}$ and PC-3/S cells as $CD44^{+}$, $CD24^{-}$, $CD40^{+}$, $CD71^{-}$ and $CDH1^{-}$. In well-characterized breast cancer cell models, $CD44^{high}/CD24^{low}$

cells showed convincing CSC properties [197] [18], as well as in other cancers, such as prostate [340]. However, other studies suggest CD24 as a CSC biomarker in ovarian cancer [323], colorectal cancer [333], prostate cancer [341], pancreatic cancer [342] and also in some cases of breast cancer [90]. We exploited the variable expression of cell-surface E-cadherin (CDH1) in parental, non-clonal PC-3 prostate cancer cells, to sort out high- from low-CDH1 expressors. We found that the $CDH1^{high}$ fraction of parental PC-3 cells displayed features closer to CSC cells (spheroid formation and expression of self-renewal genes) than the $CDH1^{-}$ fraction, and that the latter were more autonomously invasive and expressed more mesenchymal genes than the former. Thus, once more, this time by studying nonclonal tumor cells fractionated only on the basis of their E-cadherin expression, we observed the same *in vitro* functional dissociation between invasion and self-renewal and the dichotomy among CSC-TIC and EMT gene programs that we had found for the PC-3/Mc *vs*. PC-3/S or the TSU-Pr1-B2 *vs*. TSU-Pr1 cell pairs. In other words, we were able to enrich for cells with *in vitro* features consistent with TICs from a heterogeneous tumor cell line such as PC-3, simply by isolating the more epithelial tumor cells expressing high levels of E-cadherin.

In order to further explore the hypothesis that a strong self-renewal and metastatic phenotype requires the maintenance of an epithelial program in our cell models, we have (1) induced constitutive EMT in epithelial tumor subpopulations through the transduction and overexpression of EMT-directing transcription factors, (2) knocked down the same factors in the mesenchymal-like tumor PC-3/S tumor cell subpopulation, (3) knocked down E-cadherin in the epithelial tumor subpopulations, (4) transduced and overexpressed E-cadherin in PC-3/S cells, (5) knocked down self-renewal/pluripotency factors in the strongly epithelial PC-3/Mc subpopulation and (6) transduced and overexpressed the self-renewal factor SOX2 and the epithelial gene E-cadherin in the mesenchymal-like PC-3/S tumor cells. The results from all these complementary approaches have led to the same conclusions, namely that the suppression of an epithelial program (through constitutive expression of EMT transcription factors or knockdown of E-cadherin) inhibits the self-renewal/pluripotency gene network of tumor cells, their capacity to grow under attachment-independent conditions, and their tumorigenic and metastatic potentials. The association between properties attributed to tumor-initiating cells and an epithelial phenotype is further supported by the facts that knockdown of three of the four canonical Yamanaka pluripotency transcription factors (SOX2, KLF4 and MYC) in PC-3/Mc cells reduce their epithelial phenotype and TIC attributes while, very interestingly, increasing their invasiveness. On the other hand, the overexpression of SOX2 in PC-3/S cells was sufficient to enhance the expression of epithelial markers and properties of TICs in these cells, including enhanced tumorigenicity, while inhibiting invasiveness and the expression of mesenchymal markers.

Therefore, these results suggest that the CSC-TIC and EMT programs are mutually exclusive, or under reciprocal negative regulation, at least in our cell models of prostate and bladder cancers. One of the noteworthy aspects of our study is that it shows that the self-renewal properties of these tumor cells depend on the

same core factors that endow normal cells with self-renewal and pluripotency. Moreover, and importantly, the epithelial phenotype is strongly coupled to the self-renewal phenotype in our tumor cell models. This situation is reminiscent of the requirement for normal adult fibroblasts to undergo a mesenchymal-epithelial transition (MET) in order to be susceptible to reprogramming into becoming self-renewing pluripotent cells [229]. This is consistent with the observations by others that the KLF4 core reprogramming factor upregulates E-cadherin through binding to its promoter [229], and that SOX2, OCT4 and MYC suppress the EMT-promoting activities of SNAI1 and TGF-β1 and TGF-βR2, even though the precise mechanisms have not yet been elucidated [329]. In addition, our finding of a strong link in tumor cells between an epithelial gene program and self-renewal is consistent with the observation that, in normal adult fibroblasts, the expression of E-cadherin by itself can facilitate their reprogramming and the acquisition of pluripotency [68] [329], to the point of being capable of functionally replace the requirement for OCT4 to induce pluripotent stem cells from fibroblasts. Thus, expression of E-cadherin would appear to play a pivotal role in the switch between invasive and self-renewal phenotypes, favoring the transition to a more non-invasive pluripotent/self-renewal state. Possible mechanisms include the role played by E-cadherin in promoting cell-cell contacts between pluripotent cells and a direct tie of E-cadherin expression to the regulation of pluripotency transcription factors, in particular OCT4 [229].

Our own results suggest that modulation of E-cadherin alone may not be sufficient for tumor cells to acquire a self-renewal genetic program, but that it may be important to initiate critical functional and physical changes that lead to a self-renewal state. To fully acquire a CSC-TIC state, tumor cells would need the additional activation of one or more of the Yamanaka reprogramming factors, such as SOX2, MYC or KLF4, or yet other factors that, together, confer full self-renewal properties to those cells. The contraposition between a gene program that drives self-renewal and an invasive program also has precedents in other biological settings. During normal vertebrate development, neural crest progenitor cell migration and specification require the activation, among other factors, of Snail/Slug and concomitant suppression of Sox2, events that are induced both by diffusible factors and cell-cell interactions [344]. Moreover, induction of EMT by SNAI1 or others can be anti-proliferative [345] [346]. The late makes biological sense in that cells that need to migrate and invade undergo extensive rearrangements in their cytoskeleton and morphological changes involving high energy expenses and the co-option of mechanisms which may not be compatible with the cellular processes employed to divide and proliferate [27]. Indeed, it has been shown during embryonic development and at the tips of angiogenesis-migratory processes that cells stop dividing when they are migrating [138]. Our observations lead us to propose that tumor cells that depend on a self-renewal/pluripotency gene network for aggressiveness may be susceptible to an inhibition of those properties when they undergo EMT, perhaps through direct or indirect downregulation of the self-renewal/pluripotency gene network by factors such as SNAI1. In spite of the evidences to the contrary offered in our study of prostate and bladder cancer tumor cell models, we speculate that tumor cell types that do not depend on the self-renewal/pluripotency gene network for their tumorigenic and metastatic potentials

may use EMT factors to induce a self-renewal state, as described for several model [347], through a variety of mechanisms. Tissue-specific regulatory networks could also account for different phenotypic consequences of EMT in different cell types. For example, it has been reported that EMT of primary prostate epithelial cells is not accompanied with enhanced anchorage-independent growth or malignant transformation [348], and, as we have found in this study, EMT can suppress the self-renewal states of prostate and bladder cancer cells, while on the other hand the induction of EMT in non-cancerous immortalized MCF10 breast epithelial cells enhances their potential to form mammospheres [193]. Incidentally, our own experiments have confirmed the latter with MCF10CA (data not shown). Further studies address the issue of how the combinatorial stoichiometry of transcription factors may change their function over key gene promoters, leading to the activation of EMT or self-renewal programs. As examples, when OCT4 is expressed above or below 50% of its normal levels, ESCs can be induced to differentiate [50], or due to the absence of other cofactors such as β-catenin [343]. Another example is TCF3, which normally acts by repressing genes, such as CDH1, OCT4 or NANOG but becomes a transcriptional activator when is coupled with β-catennin [48]. Technical approches should be applied in order to better determine quantitative combinations of these factors over specific promoters under different times and conditions.

Our observations are also supported by models in which local invasiveness is inversely correlated with metastatic or organ colonization potential [349]. Additional studies, including some in which EMT has been proposed to enhance tumorigenic and metastatic potentials, have shown that tumor cell subpopulations with clear epithelial phenotypes are endowed with the strongest metastatic potential [350]. Indeed, immunohistochemical analyses of metastatic samples of prostate and other cancers show that they express the epithelial markers E-cadherin [136] [137]. We envision several possible scenarios: (1) if tumor cells that metastatize are those that suffered an EMT, somehow they may revert to an epithelial phenotype (MET) at the metastasis site; (2) they could actually need to revert before reaching the metastatic site in order to colonize it; (3) tumor cells that undergo EMT may not directly contribute to metastatic growth, but they could assist to the metastatic spread by cooperating with other cancer cell populations to escape from the primary tumor site; (4) tumor cells may not require to undergo EMT in order to escape from primary tumor sites, and they would use alternative modalities to invade, as discussed in the Introduction section [232]. The acquisition of EMT and posterior reversion by MET, even the latter being still somewhat controversial, is the preferred and more widely accepted explanation as key events in metastasis [133]. Our observations indicate that it is important to consider EMT and self-renewal as phenotypes that can occur independently in different cell subpopulations at the same time and also in the same cell at different times or at different locations (*in vivo*), setting the stage for highly dynamic interplays, in time and space, between these two gene programs, both of which are critical for the successful development of metastasis by tumors.

We have sought clinical correlation for our cell models, first through gene set enrichment analysis (GSEA) that yielded a subset of an ESC-like gene module [88] enriched in our PC-3/Mc cell subpopulation, revealing interesting associations. After

interrogation of a transcriptomic dataset for 150 prostate human samples, we found that our 70-gene PC-3/Mc geneset predicted those primary tumor samples which were associated with metastasis. Moreover, immunohistochemical analysis revealed a correlation of expression levels of SOX2 in primary tumor samples with more aggressive stages of prostate cancer and, importantly, several metastatic samples displayed a strong expression of SOX2 in association with a strong expression of E-cadherin in 100% of the tumor cells. Therefore, we believe that these studies in human tumors support our model of association of metastasis with active self-renewal gene programs coupled to an epithelial program, the combined phenotype being relevant for the aggressive progression of human prostate cancer.

In our dual-cell tumor models, both cell types are isolates derived from the same parental cell lines, and thus they must coexist in unknown proportions as part of a heterogeneous tumor population. In order to approach the possibility, presented above, that epithelial tumor cell subpopulations displaying distinct epithelial or mesenchymal-like phenotypes could cooperate among them as part of the metastatic behavior of a tumor composed of heterogeneous cell populations, we proceeded to co-culture PC-3/Mc cells with PC-3/S cells. We found that this caused an increased invasiveness of PC-3/Mc cells, which could be interpreted as the result of a passive escape of PC-3/Mc cells, which are poorly invasive under standard growth conditions, thanks to paths opened by the co-cultured PC-3/C cells in the layer of extracellular matrix of our invasion chambers. However, when co-culture between these two cell subpopulations was followed by FACS to obtain separate populations of each cell type, we found that PC3/Mc cells that had been cultured together with PC3/S cells had acquired an autonomous capacity to invade through extracellular matrix, that is, even in the absence of concurrent PC3/S cells. We further found that exposure of PC3/Mc cells to the conditioned medium from PC3/S cells was sufficient to stimulate the invasion of PC3/Mc cells, without a requirement for a direct contact between both cell types. Therefore, either through direct contact or through diffusible factors, or both, the mesenchymal-like PC-3/S cells enhanced the autonomous invasive potential of the epithelila PC-3/Mc cells, being higher the effects observed after direct contact, likely result of the additive help of PC-3/S to break matrigel. Most interestingly, gene expression analyses of PC-3/Mc cells that had been co-cultured with PC-3/Mc and then separated by FACS showed that they had undergone a transient EMT, with a downregulation of epithelial markers, downregulation of self-renewal genes and upregulation of some mesenchymal and EMT genes. The observed downregulation of self-renewal genes was accompanied with an inhibition of spheroid formation by the PC-3/Mc cells that had been co-cultured with PC-3/S cells, once again consistent with the mutual exclusion in these cells of self-renewal and EMT. Therefore, in addition to a possible passive escape mechanism achieved thanks to the path-opening activity of the mesenchymal-like PC-3/S cells, the epithelial-TIC PC-3/Mc cells also engage in an active escape through EMT induced by PC-3/S cells, at least in part mediated by diffusible factors.

The cooperative escape mechanisms shown *in vitro* had *in vivo* correlates, demonstrated by co-injection of both cell types into mice by different *in vivo* approaches. Thus, orthotopic injection and intramuscular injection showed that,

while tumorigenicity of PC-3/Mc cells was decreased, as expected from the downregulation of self-renewal genes and inhibition of spheroid forming potential described above upon co-culture with PC-3/S cells, their capacity to metastatize to lymph nodes, lungs or adrenal glands was significantly accelerated after orthotopic implantation or intravenous or intracardiac injections. We further showed that the epithelial PC-3/Mc cells, but not the mesenchymal-like PC-3/S cells, are the sole responsible for distant organ colonization and growth. These results indicate that when both cell populations coexist together *in vivo* they cooperate among them, such that the mesenchymal-like tumor cells facilitate the escape of from local sites, while the epithelial tumor cells have the capacity to colonize and grow at distant sites. This combinatorial and dynamic interaction of distinct tumor cell subpopulations, each endowed with distinct functions driven by distinct gene programs, also subject to dynamic modulation in time and space, give as a result an optimized compound behavior of tumors, resulting in more efficient metastatic phenotypes. Incidentally, we have found that the interaction between tumor subpopulations not only enhances or accelerates the metastatic spread of epithelial PC-3/Mc cells, but also extends their organotropism to enable them to colonize adrenal glands after co-injection with PC-3/S cells, which they never did when inoculated alone. We suggest that the transient EMT that PC-3/Mc cells undergo as a results of interaction with PC-3/S cells can not only aid them in their escape from primary tumor sites, including intravasation steps, but also affect additional steps in the metastatic cascade, such as extravasation at target metastatic organs, as we have seen for adrenal glands, which they would not normally colonize in the absence of this cooperative influence. The transient EMT induction of PC-3/Mc by PC-3/S cells can also helps the overall process in order to create an important number of invasive cells to overcome invasive limiting steps, these cells start degrading ECM, then intravasate and released into bloodstream where most of them will die, so the possibilities of spreading of invasive cells increase as many of them were in the primary site. Arrived at different organs such as lungs or adrenal gland, they can facilitate the extravasation, and finally, arrived at the final destination they could revert to epithelial-TIC cells as additive input to the following remaining epithelial-TICs coming from the primary site, which are highly endowed with anchorage-independent growth, self-renewal and tumor initiating capacity, in order to colonize the distant tissue. It is worthy to note that in bone colonization there is no increased metastasis after co-injection as was expected since the bone marrow has an open fenestrated vasculature. In all assays performed, we could determine that colonizing and metastatic forming cells were PC-3/Mc cells and never were PC-3/S cells.

Taking into consideration, all these and many other reasons like the high heterogeneity in tumors, is likely that such combinations of different tumor cell populations, finally increase the metastatic efficiency through cooperation among them, developing different roles in order to favor the final goal, metastasis. At least, this is a pattern highly observed in nature and ecological systems, understanding ecological systems as the interaction of different populations to work out adaptability, evolution and final survive.

Collectively, the results of this study suggest that, in some cancers, the acquisition of mesenchymal traits by tumor cells that leads to their loss of epithelial properties occurs at the expense of their self-renewal potential. When the induction of EMT is constitutive, as by forced overexpression of Snai1, the losses in self-renewal, tumorigenic and metastatic potentials are also sustained. On the other hand, transient EMT, such as that induced by the cooperation between epithelial and mesenchymal tumor subpopulations described in this study, could enhance the local invasiveness of the epithelial subpopulation, thus contributing to the overall metastatic potential of a tumor in which heterogeneous epithelial and mesenchymal subpopulations coexist. We propose a model in which the more epithelial/self-renewal tumor populations that leave their primary site either passively, aided by stromal or mesenchymal-like tumor cells, or actively, through their own transient EMT, can form metastases because they have maintained their epithelial phenotypes or, if they have undergone EMT at the primary site, revert to an epithelial/self-renewal program at distant sites. We further propose that, once at their metastatic sites, epithelial TICs may follow a cycle similar to that occurring at the primary sites, with induction of EMT at varying degrees, depending on the environment of the metastatic sites. Formal proof of this model, in particular the demonstration of transient local EMT of epithelial TICs followed by MET at distant sites, would require additional experimental confirmation. Our model is compatible with a more direct participation of stromal or other non-tumoral cells in promoting the local invasiveness of tumor cells.

Second study. *Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization* (Korpal, M. *et al*., Nat. Med. 17:1101-8, 2011).

This study focuses on the role of miR-200s in metastasis, particularly in the late steps of metastasis colonization. Initial results strongly correlated miR-200s family expression levels with metastasis in human primary breast tumors samples, significantly associating miR-200s expression with poor distant relapse-free survival (DRFS). Moreover, human lung-pleural metastasis samples analysis also revealed higher levels of expression of miR-200s in metastasis. Thus, this study provided strong data relating miR-200s with metastasis and poor prognosis in breast cancer. These data contrast with other recent studies in animal metastasis models that have reported conflicting roles of miR-200s in metastatic progression, favoring metastasis in breast cancer [320] or showing an inverse correlation with metastasis in lung adenocarcinoma and pancreatic neuroendocrine cells [354] [355]. These contradictory findings may in part be due to the different models studied. However, it is more likely that the net effect of miR-200s in metastasis progression of various cancers depends on several variables across different models, such as differences in the rate-limiting steps of the metastatic cascade and the importance of MET in colonization in different models. Indeed, to our knowledge, this study is the first large-scale breast cancer clinical study showing a clear clinical association of miR-200 family expression with poor DRFS (Distant relapse-free survival), particularly in ER-positive breast cancers. Furthermore, miR-200 expression levels were higher in lung metastases than in primary tumors. In support of these clinical associations, a survey

across several different isogenic series of cancer cell lines revealed a strong correlation between miR-200 expression and metastatic ability, which was further supported by functional analysis *in vivo*.

In order to further investigate the importance of miR-200s in metastasis, we used the mouse mammary cell lines 4T1 series (67NR, 168FARN, 4TO7 and 4T1), and the MCF10A human epithelial breast cells and the TSU-Pr1 bladder cancer series (also used in the first study).The main cellular model used in this second study is based on the dual comparison between 4TO7 (weakly metastatic) and 4T1 (highly metastatic) cell lines. Remarkably, these models once again support the functional dichotomy observed in the prostate model characterized in the first study, 4TO7 being more invasive, expressing mesenchymal genes and are less metastatic potential than 4T1 cells, which are less invasive and express epithelial genes, including E-cadherin or desmoplakin and furthermore express high levels of miR-200s, in contrast to 4TO7 cells. Importantly, after orthotopic injection, 4T1 cells are capable of spontaneously metastasizing and colonizing distant organs. This metastatic capacity of 4T1 was already reported in other significant studies like that by Yang *et al*. [195]. In that study, the authors contended that 4T1 cells were more metastatic than 4TO7 because they express higher levels of TWIST1 (we have confirmed), which was proposed to engage the EMT program with concurrent E-cadherin repression. This hypothesis was not demonstrated in Yang's study. However, in our hands and as also shown by others [350], 4T1 cells express much higher levels of E-cadherin and other epithelial markers than 4TO7 cells, along with high levels of expression of the mir-200 family of upstream positive regulators of E-cadherin. One possible interpretation of the results of the study by Yang *et al*. is that somehow TWIST1 may facilitate metastasis by orchestrating genes that are not related with the epithelial phenotype of 4T1 cells, given that the expression of this transcription factor does not induce a real EMT (high levels of E-cadherin continue to be expressed). In fact, it could be argued that this study by Yang *et al*. unintentionally reinforces one of the main notions of this thesis, namely that the expression of an epithelial phenotype, possibly coupled with other genetic programs like self-renewal network, may favor the metastatic colonization capacity, although it is clear that the acquisition of EMT is crucial in specific steps of the metastatic cascade, in particular to facilitate the escape of epithelial tumor cells from primary sites.

Although 4TO7 cells were less metastatic than 4T1 cells, with very low lung colonization capacity, they do disseminate faster into the bloodstream than 4T1 after orthotopic injection. Ectopic expression of miR-200s in 4TO7 cells caused their decreased dissemination from the primary site but increased their overall metastatic lung colonization capacity. This was accompanied with an upregulation of E-cadherin and other epithelial genes, along with a downregulation of EMT genes. The upregulation of E-cadherin after overexpression of miR-200s has previously been reported by many studies and is mainly due to the downregulation of the miR-200s targets ZEB1 and ZEB2, potent E-cadherin transcriptional repressors [352] [353]. Knocking-down E-cadherin in 4TO7 cells with ectopic expression of miR-200s modestly decreased lung colonization, while the ectopic expression of E-cadherin in

4TO7 cells couldn't fully recapitulate the metastatic potential of miR-200s overexpressing lines after orthotopic injection. Interestingly, cells expressing ectopic E-cadherin showed a stronger decline in their dissemination from primary site than the miR-200s expressing cells, which also express high levels of E-cadherin. These results suggested that miR-200s enhanced the metastatic potential of 4TO7 cells through factors other than the miR-200s-Zebs-E-cadherin axis. Although transcriptomic profiling showed that the forced expression of E-cadherin did not produce the same global gene expression changes caused by expression of miR-200s, however, a common core of 13 genes were represented within clusters 1 (C1), 2 (C2) and C1+C2 of genes induced by miR-200s overexpression, being E-caderin the top one and other epithelial genes, like EpCAM and MSLN also involved in self-renewal and colony formation, respectively. Another epithelial gene in common was Perp involved in cell-cell desmosome junctions. Additionally, gene ontology analysis of the transcriptomic profiles also revealed many GO categories representing epithelial phenotypes. This suggested that the epithelial phenotype taken as a whole could play an important role in the functional behavior observed *in vivo*.

Tumor cells are inherently highly migratory and invasive *in vitro* and *in vivo*, and although ectopic miR-200 expression reduced dissemination, it had a net effect of enhancing the rate-limiting colonization step. In contrast, this outcome is less likely in model systems where earlier steps of metastasis, such as early dissemination through the acquisition of EMT-like properties, is the rate-limiting process, as may be the case for the model systems showing metastasis-suppressor functions of miR-200s [354] [355].

Transcriptomic and proteomic analysis (mass spectrometry) revealed that miR-200s also repressed the tumor cell secretome by direct targeting of the SEC23A-mediated transport pathway, which affects cell-extrinsic tumor-stromal interactions. This targeting mechanism was also validated in other human cells lines, including MDA-MB-231 and TSU-Pr1. SEC23A was expressed at lower levels in the epithelial highly metastatic cell lines 4T1 and MCF10CA1a than in the mesenchymal-like low metastatic cell lines 4TO7 and MCF10CA1. Moreover, SEC23A levels were significantly lower in clinical metastatic samples relative to primary tumors, consistent with a potential role of SEC23A as a suppressor of metastatic colonization. Knock-down of SEC23A inhibited migration and invasion but increased lung colonization, although its forced overexpression did not suppress metastasis.

Interestingly, SEC23A was indispensable for the secretion of extracellular matrix components. Knock-down of Sec23a in 4TO7 cells produced a global reduction of secreted proteins. A further screening of these reduced factors based on its clinical relevance determined by their low levels of expression in association with low relapse-free survival, yielded six genes. Knock-down of two of them, Tinagl1 and Igfbp4, functionally *in vitro* and *in vivo* phenocopied miR-200s overexpressing and Sec23a knock-down cells. We further identified Tinagl1 and Igfbp4 as key Sec23a-mediated secretory proteins that substantially reduced metastatic colonization, with clinical analysis correlations. This revealed them as factors that suppress lung metastasis that should be amenable to therapeutic targeting since

they are secreted. Our findings support dynamic roles for EMT and MET during different stages of metastasis: whereas EMT and low levels of miR-200s promote invasion and intravasation, MET and high miR-200 expression is required for efficient colonization of secondary organs. In addition to the well-established function of miR-200s in regulating intrinsic cellular properties through ZEB1, ZEB2 and E-cadherin, our study shows that miR-200s may also have the potential to influence the microenvironment via directed targeting of SEC23A-dependent secretome. Through regulation of this pathway, miR-200s may extend their reach to manipulate a neighboring population of tumor and stromal cells, influencing the collective behavior of these cells during metastasis.

In summary, we have combined clinical and experimental studies to establish a biphasic role for miR-200s in metastasis. We showed that miR-200s promote metastatic colonization by enhancing cell-intrinsic epithelial traits via the ZEB–E-cadherin axis and by inhibiting a SEC23A-dependent regulation of tumor secretome. The fact that targeting of SEC23A by miR-200 seems to have dichotomous roles in metastasis, hindering early steps of migration and invasion while promoting the late step of metastatic colonization, may explain the contradictory roles of miR-200s in various models of metastasis. Although miR-200s and the SEC23A pathway may represent new therapeutic opportunities for metastatic cancer, their dichotomous functions warrant careful assessment of potential therapeutic benefits and adverse side effects when treatments are applied at different stages of the disease.

Both studies (article 1 and 2), when considered together, share important and novel views of metastasis, mainly in relationship with the last steps of the metastatic cascade, strongly suggesting that prostate, bladder and breast tumor cells with an epithelial non-secretory phenotype are more prone to metastatize. We propose that mesenchymal-like tumor cells with a relatively stable EMT derive from epithelial-CSC cells at the primary tumor site, largely in response to microenvironmental influences. As we have shown in the first study, both tumor cell populations, epithelial-CSC or TIC and mesenchymal-like, interact with each other, one consequence being a transient EMT of the epithelial-CSC population that had retained their epithelial state. Thus, in a given tumor microenvironment, there is a strong pressure for epithelial-CSC to drift from epithelial to mesenchymal-like phenotypes. Together, these heterogeneous populations of tumor cells are effective at overcoming the initial barriers to metastasis where invasion is required. After completing these initial steps, mesenchymal-like tumor cell populations are less critically required, while epithelial-CSC are more important and critical to finally grow metastastatic secondary tumors. There are many reasons to support that epithelial-CSC-unsecretory cells are more capable of metastatic colonization, including the facts that they can form clusters in the bloodstream for better survival, are resistant to anoikis, they show anchorage-independent growth and strong self-renewal capacity that permit them to adapt to different environments and initiate and grow secondary tumors. Very interestingly, such cells have a reduced secretion of metastatic suppressor factors which may interact in autocrine and paracrine fashion with tumoral and stromal cells at metastatic sites.

Conclusions

- Using prostate, bladder and breast cancer cell models, we have shown a functional dichotomy between two fundamental processes that determine the ability of tumor cells to metastasize, invasion and self-renewal, each associated with a distinct tumor cell subpopulation.

- In our cancer models, epithelial phenotypes are coupled to cancer stem cell-tumor initiating cell properties, being much more efficient than mesenchymal-like tumor cells in growing tumors and metastases in immunodeficient mice.

- Epithelial-mesenchymal transition and self-renewal may be mutually exclusive gene programs in some cancer cells. Thus, induction of epithelial-mesenchymal transition in prostate and bladder epithelial tumor-initianting cells can suppress their self-renewal gene program and their tumorigenic and metastatic potentials.

- In our prostate cancer model, we have shown that interactions among tumor cell subpopulations have significant consequences for the efficiency of metastasis. Thus, mesenchymal-like tumor cells induce a transient epithelial-mesenchymal transition on epithelilal tumor-initiating cells, transiently enhancing their invasiveness and reducing their self-renewal properties in vitro, and enhancing their escape from local tumor sites with a consequent acceleration of metastatic colonization. At least part of this interaction occurs through diffusible factors produced by the mesenchymal-like subpopulation.

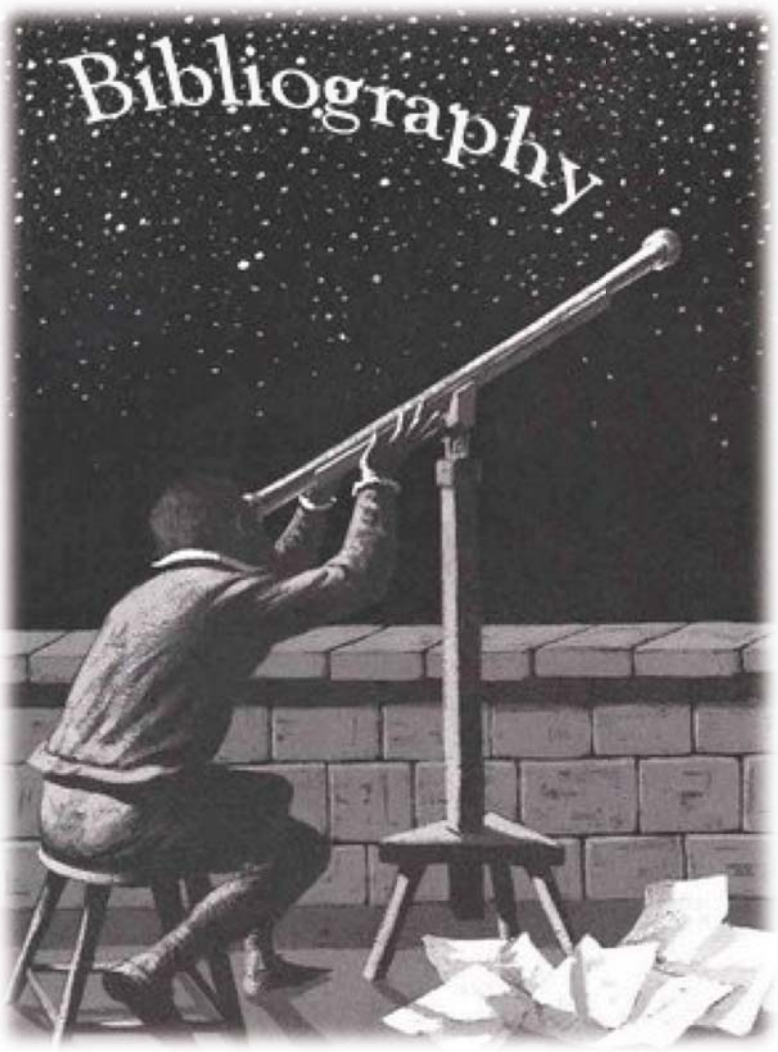
- The expression of SOX2 is significantly associated with more advanced stages of prostate cancer.

- The expression of the miR-200 family of microRNAs is significantly associated with metastasis in breast cancer.

- MiR-200s induce a highly metastatic epithelial-nonsecretory phenotype in breast cancer cells, through the direct targeting of Zeb factors and the secretory transport regulator Sec23a.

- Igfbp4 and Tinagl1 are identified as metastasis suppressors regulated by Sec23a and thus by the upstream control of miR-200s.

- Both studies share a major conclusion: tumor cells to colonize at distant organs and metastatize, they require to engage an active epithelial gene program and phenotype.


Bibliography

Galileo Galilei (Pisa, 1564 - Florencia, 1642)

[1] D. Hanahan, R. A. Weinberg, and S. Francisco, "The Hallmarks of Cancer Review University of California at San Francisco," *Hormone Research*, vol. 100, pp. 57-70, 2000.

[2] D. Hanahan and R. a Weinberg, "Hallmarks of cancer: the next generation.," *Cell*, vol. 144, no. 5, pp. 646-74, Mar. 2011.

[3] G. P. Gupta and J. Massagué, "Cancer metastasis: building a framework.," *Cell*, vol. 127, no. 4, pp. 679-95, Nov. 2006.

[4] B. Weigelt, J. L. Peterse, and L. J. Van T Veer, "Breast cancer metastasis: markers and models.," *Nature Reviews Cancer*, vol. 5, no. 8, pp. 591-602, 2005.

[5] D. X. Nguyen, P. D. Bos, and J. Massagué, "Metastasis: from dissemination to organ-specific colonization.," *Nature reviews. Cancer*, vol. 9, no. 4, pp. 274-84, Apr. 2009.

[6] H. H. Q. Heng et al., "Stochastic cancer progression driven by non-clonal chromosome aberrations.," *Journal of Cellular Physiology*, vol. 208, no. 2, pp. 461-472, 2006.

[7] L. M. F. Merlo and C. C. Maley, "The role of genetic diversity in cancer," *Medicine*, vol. 120, no. 2, pp. 401-403, 2010.

[8] A. P. Feinberg, R. Ohlsson, and S. Henikoff, "The epigenetic progenitor origin of human cancer," *Nature Reviews Genetics*, vol. 7, no. 1, pp. 21-33, 2006.

[9] R. Axelrod, D. E. Axelrod, and K. J. Pienta, "Evolution of cooperation among tumor cells.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 36, pp. 13474-13479, 2006.

[10] F. Balkwill, "Cancer and the chemokine network.," *Nature Reviews Cancer*, vol. 4, no. 7, pp. 540-550, 2004.

[11] M.-H. Yang et al., "Direct regulation of TWIST by HIF-1alpha promotes metastasis.," *Nature Cell Biology*, vol. 10, no. 3, pp. 295-305, 2008.

[12] L. M. F. Merlo, J. W. Pepper, B. J. Reid, and C. C. Maley, "Cancer as an evolutionary and ecological process.," *Nature Reviews Cancer*, vol. 6, no. 12, pp. 924-35, 2006.

[13] E. Moreno, "Is cell competition relevant to cancer?," *Nature Reviews Cancer*, vol. 8, no. 2, pp. 141-147, 2008.

[14] J. C. Pastor-pareja, T. Xu, and M. Wu, "Interaction between RasV12 and scribbled clones induces tumour growth and invasion," *Nature*, vol. 463, no. 7280, pp. 1-5, 2010.

[15] A. E. Karnoub et al., "Mesenchymal stem cells within tumour stroma promote breast cancer metastasis.," *Nature*, vol. 449, no. 7162, pp. 557-563, 2007.

[16] H. H. Q. Heng, S. W. Bremer, J. B. Stevens, K. J. Ye, G. Liu, and C. J. Ye, "Genetic and epigenetic heterogeneity in cancer: a genome-centric perspective.," *Journal of cellular physiology*, vol. 220, no. 3, pp. 538-47, Sep. 2009.

[17] P. C. Nowell, "The Clonal Evolution of Tumor Cell Populations," *Science*, vol. 23-28, pp. 23-28, 1976.

[18] K. Polyak and L. Campbell-Marrotta, "Breast tumor heterogeneity: cancer stem cells or clonal evolution," *Breast cancer research : BCR*, vol. 11 Suppl 1, no. 19, p. S18, Jun. 2007.

[19] G. H. Heppner, "Tumor heterogeneity," *Cancer Research*, vol. 44, no. 6, pp. 2259-65, 1984.

[20] S. R. Wolman, "Cytogenetic heterogeneity: its role in tumor evolution.," *Cancer Genetics and Cytogenetics*, vol. 19, no. 1-2, pp. 129-140, 1986.

[21] C. R. Boland, J. Sato, H. D. Appelman, R. S. Bresalier, and A. P. Feinberg, "Microallelotyping defines the sequence and tempo of allelic losses at tumour suppressor gene loci during colorectal cancer progression.," *Nature Medicine*, vol. 1, no. 9, pp. 902-909, 1995.

[22] T. Reya, S. J. Morrison, M. F. Clarke, and I. L. Weissman, "Stem cells, cancer, and cancer stem cells.," *Nature*, vol. 414, no. 6859, pp. 105-111, 2001.

[23] J. E. Visvader and G. J. Lindeman, "Cancer stem cells in solid tumours: accumulating evidence and unresolved questions.," *Nature reviews. Cancer*, vol. 8, no. 10, pp. 755-68, Oct. 2008.

[24] D. Bonnet and J. E. Dick, "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell.," *Nature Medicine*, vol. 3, no. 7, pp. 730-737, 1997.

[25] F. Li, B. Tiede, J. Massagué, and Y. Kang, "Beyond tumorigenesis: cancer stem cells in metastasis.," *Cell research*, vol. 17, no. 1, pp. 3-14, Jan. 2007.

[26] R. Pardal, M. F. Clarke, and S. J. Morrison, "Applying the principles of stem-cell biology to cancer.," *Nature reviews. Cancer*, vol. 3, no. 12, pp. 895-902, Dec. 2003.

[27] S. Floor, W. C. G. van Staveren, D. Larsimont, J. E. Dumont, and C. Maenhaut, "Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating-cancer stem cells: distinct, overlapping or same populations.," *Oncogene*, pp. 1-13, Jul. 2011.

[28] C. Maenhaut, J. E. Dumont, P. P. Roger, and W. C. G. van Staveren, "Cancer stem cells: a reality, a myth, a fuzzy concept or a misnomer? An analysis.," *Carcinogenesis*, vol. 31, no. 2, pp. 149-58, Mar. 2010.

[29] M. Osawa, K. Hanada, H. Hamada, and H. Nakauchi, "Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell.," *Science*, vol. 273, no. 5272, pp. 242-245, 1996.

[30] J. S. Lam and R. E. Reiter, "Stem cells in prostate and prostate cancer development.," *Urologic Oncology*, vol. 24, no. 2, pp. 131-140, 2006.

[31] S. J. Miller, R. M. Lavker, and T.-T. Sun, "Interpreting epithelial cancer biology in the context of stem cells: tumor properties and therapeutic implications.," *Biochimica et Biophysica Acta*, vol. 1756, no. 1, pp. 25-52, 2005.

[32] M. S. Wicha, S. Liu, and G. Dontu, "Cancer stem cells: an old idea--a paradigm shift," *Cancer Research*, vol. 66, no. 4, pp. 1883-1886, 2006.

[33] M. Hu et al., "Distinct epigenetic changes in the stromal cells of breast cancers.," *Nature Genetics*, vol. 37, no. 8, pp. 899-905, 2005.

[34] S. A. Bapat, "Evolution of cancer stem cells.," *Seminars in Cancer Biology*, vol. 17, no. 3, pp. 204-213, 2007.

[35] R. P. Hill, "Identifying cancer stem cells in solid tumors: case not proven.," *Cancer Research*, vol. 66, no. 4, pp. 1891-1895; discussion 1890, 2006.

[36] D. P. Cahill, K. W. Kinzler, B. Vogelstein, and C. Lengauer, "Genetic instability and darwinian selection in tumours.," *Trends in Cell Biology*, vol. 9, no. 12, p. M57-M60, 1999.

[37] L. Wolpert, "Cancer: the evolutionary legacy," *Medical History*, vol. 47, no. 1, pp. 122-123, 2003.

[38] U. R. Rapp, F. Ceteci, and R. Schreck, "Oncogene-induced plasticity and cancer stem cells.," *Cell cycle Georgetown Tex*, vol. 7, no. 1, pp. 45-51, 2008.

[39] M. Shipitsin and K. Polyak, "The cancer stem cell hypothesis: in search of definitions, markers, and relevance.," *Laboratory investigation a journal of technical methods and pathology*, vol. 88, no. 5, pp. 459-463, 2008.

[40] J. Zhou and Y. Zhang, "Cancer stem cells: Models, mechanisms and implications for improved treatment.," *Cell cycle Georgetown Tex*, vol. 7, no. 10, pp. 1360-1370, 2008.

[41] A. Brock, H. Chang, and S. Huang, "Non-genetic heterogeneity--a mutation-independent driving force for the somatic evolution of tumours.," *Nature Reviews Genetics*, vol. 10, no. 5, pp. 336-342, 2009.

[42] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang, "Transcriptome-wide noise controls lineage choice in mammalian progenitor cells," *Nature*, vol. 453, no. 7194, pp. 544-547, 2008.

[43] J. M. Rosen and C. T. Jordan, "The increasing complexity of the cancer stem cell paradigm.," *Science (New York, N.Y.)*, vol. 324, no. 5935, pp. 1670-3, Jun. 2009.

[44] E. Quintana, M. Shackleton, M. S. Sabel, D. R. Fullen, T. M. Johnson, and S. J. Morrison, "Efficient tumour formation by single human melanoma cells.," *Nature*, vol. 456, no. 7222, pp. 593-8, Dec. 2008.

[45] P. B. Gupta et al., "Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells," *Cell*, vol. 146, no. 4, pp. 633-644, Aug. 2011.

[46] D. Hong et al., "Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia.," *Science*, vol. 319, no. 5861, pp. 336-339, 2008.

[47] P. N. Kelly, A. Dakic, J. M. Adams, S. L. Nutt, and A. Strasser, "Tumor growth need not be driven by rare cancer stem cells.," *Science*, vol. 317, no. 5836, p. 337, 2007.

[48] J. Wray and C. Hartmann, "WNTing embryonic stem cells.," *Trends in cell biology*, vol. 22, no. 3, pp. 159-168, Dec. 2011.

[49] J. Wang et al., "A protein interaction network for pluripotency of embryonic stem cells.," *Nature*, vol. 444, no. 7117, pp. 364-368, 2006.

[50] H. Niwa, J. Miyazaki, and a G. Smith, "Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.," *Nature genetics*, vol. 24, no. 4, pp. 372-6, Apr. 2000.

[51] S. Masui et al., "Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.," *Nature cell biology*, vol. 9, no. 6, pp. 625-35, Jun. 2007.

[52] I. Chambers et al., "Nanog safeguards pluripotency and mediates germline development.," *Nature*, vol. 450, no. 7173, pp. 1230-4, Dec. 2007.

[53] J. Jiang et al., "A core Klf circuitry regulates self-renewal of embryonic stem cells.," *Nature cell biology*, vol. 10, no. 3, pp. 353-60, Mar. 2008.

[54] L. a Boyer et al., "Core transcriptional regulatory circuitry in human embryonic stem cells.," *Cell*, vol. 122, no. 6, pp. 947-56, Sep. 2005.

[55] J. Kim, J. Chu, X. Shen, J. Wang, and S. H. Orkin, "An extended transcriptional network for pluripotency of embryonic stem cells.," *Cell*, vol. 132, no. 6, pp. 1049-61, Mar. 2008.

[56] X. Chen et al., "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells.," *Cell*, vol. 133, no. 6, pp. 1106-17, Jun. 2008.

[57] H. Niwa, K. Ogawa, D. Shimosato, and K. Adachi, "A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells.," *Nature*, vol. 460, no. 7251, pp. 118-22, Jul. 2009.

[58] T. Miki, S.-ya Yasuda, and M. Kahn, "Wnt/β-catenin signaling in embryonic stem cell self-renewal and somatic cell reprogramming.," *Stem cell reviews*, vol. 7, no. 4, pp. 836-46, Nov. 2011.

[59] A. Suzuki et al., "Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 27, pp. 10294-9, Jul. 2006.

[60] N. Sato, L. Meijer, L. Skaltsounis, P. Greengard, and A. H. Brivanlou, "Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor.," *Nature medicine*, vol. 10, no. 1, pp. 55-63, Jan. 2004.

[61] U. Ullmann, C. Gilles, M. De Rycke, H. Van de Velde, K. Sermon, and I. Liebaers, "GSK-3-specific inhibitor-supplemented hESC medium prevents the epithelial-mesenchymal transition process and the up-regulation of matrix metalloproteinases in hESCs cultured in feeder-free conditions.," *Molecular human reproduction*, vol. 14, no. 3, pp. 169-79, Mar. 2008.

[62] L. Pereira, F. Yi, and B. J. Merrill, "Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal.," *Molecular and cellular biology*, vol. 26, no. 20, pp. 7479-91, Oct. 2006.

[63] W.-L. Tam et al., "T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways.," *Stem cells (Dayton, Ohio)*, vol. 26, no. 8, pp. 2019-31, Aug. 2008.

[64] M. F. Cole, S. E. Johnstone, J. J. Newman, M. H. Kagey, and R. a Young, "Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells.," *Genes & development*, vol. 22, no. 6, pp. 746-55, Mar. 2008.

[65] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.," *Cell*, vol. 126, no. 4, pp. 663-676, 2006.

[66] J. Yu et al., "Induced pluripotent stem cell lines derived from human somatic cells.," *Science*, vol. 318, no. 5858, pp. 1917-20, 2007.

[67] M. Nakagawa et al., "Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts.," *Nature Biotechnology*, vol. 26, no. 1, pp. 101-106, 2008.

[68] T. Redmer, S. Diecke, T. Grigoryan, A. Quiroga-Negreira, W. Birchmeier, and D. Besser, "E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming.," *EMBO reports*, vol. 12, no. 7, pp. 720-6, Jan. 2011.

[69] S. Yamanaka, "Elite and stochastic models for induced pluripotent stem cell generation.," *Nature*, vol. 460, no. 7251, pp. 49-52, 2009.

[70] L. Hu, C. McArthur, and R. B. Jaffe, "Ovarian cancer stem-like side-population cells are tumourigenic and chemoresistant.," *British journal of cancer*, vol. 102, no. 8, pp. 1276-83, May 2010.

[71] V. Clément et al., "Marker-independent identification of glioma-initiating cells.," *Nature methods*, vol. 7, no. 3, pp. 224-8, Mar. 2010.

[72] A. K. Croker et al., "High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability.," *Journal of Cellular and Molecular Medicine*, vol. 13, no. 8, pp. 2236-2252, 2009.

[73] A. D. Lander, "The 'stem cell' concept: is it holding us back?," *Journal of biology*, vol. 8, no. 8, p. 70, Jan. 2009.

[74] M. F. Pera and P. P. L. Tam, "Extrinsic regulation of pluripotent stem cells.," *Nature*, vol. 465, no. 7299, pp. 713-20, Jun. 2010.

[75] K. Gosselin et al., "Senescence-associated oxidative DNA damage promotes the generation of neoplastic cells.," *Cancer research*, vol. 69, no. 20, pp. 7917-25, Oct. 2009.

[76] M.-T. Mueller, P. C. Hermann, and C. Heeschen, "Cancer stem cells as new therapeutic target to prevent tumour progression and metastasis.," *Frontiers in bioscience elite edition*, vol. 2, pp. 602-613, 2010.

[77] J. N. Rich, "Cancer stem cells in radiation resistance.," *Cancer biology therapy*, vol. 67, no. 19, pp. 8980-8984, 2006.

[78] B. M. Boman and M. S. Wicha, "Cancer stem cells: a step toward the cure.," *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 26, no. 17, pp. 2795-9, Jun. 2008.

[79] D. T. Scadden, "The stem-cell niche as an entity of action.," *Nature*, vol. 441, no. 7097, pp. 1075-1079, 2006.

[80] T. Borovski, F. De Sousa E Melo, L. Vermeulen, and J. P. Medema, "Cancer stem cell niche: the place to be.," *Cancer research*, vol. 71, no. 3, pp. 634-9, Feb. 2011.

[81] C. Calabrese et al., "A perivascular niche for brain tumor stem cells.," *Cancer Cell*, vol. 11, no. 1, pp. 69-82, 2007.

[82] K. E. Hovinga et al., "Inhibition of notch signaling in glioblastoma targets cancer stem cells via an endothelial cell intermediate.," *Stem Cells*, vol. 28, no. 6, pp. 1019-1029, 2010.

[83] C. Folkins, S. Man, P. Xu, Y. Shaked, D. J. Hicklin, and R. S. Kerbel, "Anticancer therapies combining antiangiogenic and tumor cell cytotoxic effects reduce the tumor stem-like cell fraction in glioma xenograft tumors.," *Cancer Research*, vol. 67, no. 8, pp. 3560-3564, 2007.

[84] R. N. Kaplan et al., "VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche.," *Nature*, vol. 438, no. 7069, pp. 820-827, 2005.

[85] S. Hiratsuka, A. Watanabe, H. Aburatani, and Y. Maru, "Tumour-mediated upregulation of chemoattractants and recruitment of myeloid cells predetermines lung metastasis.," *Nature Cell Biology*, vol. 8, no. 12, pp. 1369-1375, 2006.

[86] E. Olaso, A. Santisteban, J. Bidaurrazaga, A. M. Gressner, J. Rosenbaum, and F. Vidal-Vanaclocha, "Tumor-dependent activation of rodent hepatic stellate cells during experimental melanoma metastasis.," *Hepatology*, vol. 26, no. 3, pp. 634-642, 1997.

[87] E. H. Van Der Horst et al., "Discovery of Fully Human Anti-MET Monoclonal Antibodies with Antitumor Activity against Colon Cancer Tumor Models In Vivo1," *Neoplasia New York Ny*, vol. 11, no. 4, pp. 355-364, 2009.

[88] D. J. Wong, H. Liu, T. W. Ridky, D. Cassarino, E. Segal, and H. Y. Chang, "Module map of stem cell genes guides creation of epithelial cancer stem cells.," *Cell stem cell*, vol. 2, no. 4, pp. 333-44, Apr. 2008.

[89] S. Sinha, A. S. Adler, Y. Field, H. Y. Chang, and E. Segal, "Systematic functional characterization of cis-regulatory motifs in human core promoters," *Genome Research*, vol. 18, no. 3, pp. 477-488, 2008.

[90] I. Ben-Porath et al., "An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors.," *Nature genetics*, vol. 40, no. 5, pp. 499-507, May 2008.

[91] M. Widschwendter et al., "Epigenetic stem cell signature in cancer," *Nature Genetics*, vol. 39, no. 2, pp. 157-158, 2007.

[92] J. Kim et al., "A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs.," *Cell*, vol. 143, no. 2, pp. 313-24, Oct. 2010.

[93] R. Bernards and R. A. Weinberg, "A progression puzzle.," *Nature*, vol. 418, no. 6900, p. 823, 2002.

[94] I. J. Fidler and J. E. Talmadge, "Evidence that intravenously derived murine pulmonary melanoma metastases can originate from the expansion of a single tumor cell.," *Cancer Research*, vol. 46, no. 10, pp. 5167-5171, 1986.

[95] U. Cavallaro and G. Christofori, "Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.," *Nature Reviews Cancer*, vol. 4, no. 2, pp. 118-132, 2004.

[96] W. Guo and F. G. Giancotti, "Integrin signalling during tumour progression," *Nature Reviews Molecular Cell Biology*, vol. 5, no. 10, pp. 816-826, 2004.

[97] P. Friedl and K. Wolf, "Tumour-cell invasion and migration: diversity and escape mechanisms.," *Nature Reviews Cancer*, vol. 3, no. 5, pp. 362-74, 2003.

[98] L. A. Liotta and E. C. Kohn, "The microenvironment of the tumour-host interface.," *Nature*, vol. 411, no. 6835, pp. 375-379, 2001.

[99] M. Egeblad and Z. Werb, "New functions for the matrix metalloproteinases in cancer progression.," *Nature Reviews Cancer*, vol. 2, no. 3, pp. 161-174, 2002.

[100] C. M. Overall and O. Kleifeld, "Tumour microenvironment - opinion: validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy.," *Nature Reviews Cancer*, vol. 6, no. 3, pp. 227-239, 2006.

[101] L. M. Coussens, B. Fingleton, and L. M. Matrisian, "Cancer therapy - Matrix metalloproteinase inhibitors and cancer: Trials and tribulations," *Science*, vol. 295, no. 5564, pp. 2387-2392, 2002.

[102] D. Hanahan and J. Folkman, "Patterns and Emerging Mechanisms of the Angiogenic Switch during Tumorigenesis," *Cell*, vol. 86, no. 3, pp. 353-364, 1996.

[103] K. Alitalo, T. Tammela, and T. V. Petrova, "Lymphangiogenesis in development and human disease.," *Nature*, vol. 438, no. 7070, pp. 946-953, 2005.

[104] J. Condeelis and J. E. Segall, "Intravital imaging of cell movement in tumours.," *Nature reviews. Cancer*, vol. 3, no. 12, pp. 921-30, Dec. 2003.

[105] G. F. Nash, L. F. Turner, M. F. Scully, and A. K. Kakkar, "Platelets and cancer.," *The lancet oncology*, vol. 3, no. 7, pp. 425-430, 2002.

[106] D. J. Ruiter, J. H. Van Krieken, G. N. Van Muijen, and R. M. De Waal, "Tumour metastasis: is tissue an issue?," *The lancet oncology*, vol. 2, no. 2, pp. 109-112, 2001.

[107] S. Douma, T. V. Laar, J. Zevenhoven, and R. Meuwissen, "Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB," *Nature*, vol. 430, no. August, pp. 1034-9, 2004.

[108] H. Wang et al., "Tumor cell alpha3beta1 integrin and vascular laminin-5 mediate pulmonary arrest and metastasis.," *The Journal of Cell Biology*, vol. 164, no. 6, pp. 935-941, 2004.

[109] D. M. Brown and E. Ruoslahti, "Metadherin, a cell surface protein in breast tumors that mediates lung metastasis.," *Cancer Cell*, vol. 5, no. 4, pp. 365-374, 2004.

[110] A. Müller et al., "Involvement of chemokine receptors in breast cancer metastasis.," *Nature*, vol. 410, no. 6824, pp. 50-56, 2001.

[111] A. B. Al-Mehdi, K. Tozawa, A. B. Fisher, L. Shientag, A. Lee, and R. J. Muschel, "Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis.," *Nature Medicine*, vol. 6, no. 1, pp. 100-102, 2000.

[112] C. Khanna et al., "The membrane-cytoskeleton linker ezrin is necessary for osteosarcoma metastasis.," *Nature Medicine*, vol. 10, no. 2, pp. 182-186, 2004.

[113] S. M. Weis and D. A. Cheresh, "Pathophysiological consequences of VEGF-induced vascular permeability.," *Nature*, vol. 437, no. 7058, pp. 497-504, 2005.

[114] M. L. Criscuoli, M. Nguyen, and B. P. Eliceiri, "Tumor metastasis but not tumor growth is dependent on Src-mediated vascular permeability.," *Blood*, vol. 105, no. 4, pp. 1508-1514, 2005.

[115] S. Paget, "The distribution of secondary growths in cancer of the breast," *The Lancet*, vol. 133, no. 3421, pp. 571-573, 1889.

[116] I. J. Fidler, "The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited," *Nature Reviews Cancer*, vol. 3, no. 6, pp. 453-458, 2003.

[117] D. Tarin, J. E. Price, M. G. Kettlewell, R. G. Souter, A. C. Vass, and B. Crossley, "Mechanisms of human tumor metastasis studied in patients with peritoneovenous shunts.," *Cancer Research*, vol. 44, no. 8, pp. 3584-3592, 1984.

[118] S. Braun et al., "A pooled analysis of bone marrow micrometastasis in breast cancer.," *The New England Journal of Medicine*, vol. 353, no. 8, pp. 793-802, 2005.

[119] E. D. Hay, "An overview of epithelio-mesenchymal transformation," *Acta Anatomica*, vol. 154, no. 1, pp. 8-20, 1995.

[120] G. Greenburg and E. D. Hay, "Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells.," *The Journal of Cell Biology*, vol. 95, no. 1, pp. 333-9, 1982.

[121] B. Boyer and J. P. Thiery, "Epithelium-mesenchyme interconversion as example of epithelial plasticity.," *APMIS acta pathologica microbiologica et immunologica Scandinavica*, vol. 101, no. 4, pp. 257-268, 1993.

[122] H. Peinado, F. Portillo, and A. Cano, "Transcriptional regulation of cadherins during development and carcinogenesis.," *The International journal of developmental biology*, vol. 48, no. 5-6, pp. 365-75, Jan. 2004.

[123] J. Yang and R. a Weinberg, "Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis.," *Developmental cell*, vol. 14, no. 6, pp. 818-29, Jun. 2008.

[124] B. Christ and C. P. Ordahl, "Early stages of chick somite development.," *Anatomy and Embryology*, vol. 191, no. 5, pp. 381-396, 1995.

[125] M. Zeisberg and E. G. Neilson, "Review series personal perspective Biomarkers for epithelial-mesenchymal transitions," *Cell*, vol. 119, no. 6, pp. 1429-1437, 2009.

[126] J. P. Thiery and J. P. Sleeman, "Complex networks orchestrate epithelial-mesenchymal transitions.," *Nature reviews. Molecular cell biology*, vol. 7, no. 2, pp. 131-42, Feb. 2006.

[127] J. Wang et al., "Atrioventricular cushion transformation is mediated by ALK2 in the developing mouse heart.," *Developmental Biology*, vol. 286, no. 1, pp. 299-310, 2005.

[128] J. Zavadil and E. P. Böttinger, "TGF-beta and epithelial-to-mesenchymal transitions.," *Oncogene*, vol. 24, no. 37, pp. 5764-5774, 2005.

[129] a Cano et al., "The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.," *Nature cell biology*, vol. 2, no. 2, pp. 76-83, Feb. 2000.

[130] M. A. Nieto, M. G. Sargent, D. G. Wilkinson, and J. Cooke, "Control of cell behavior during vertebrate development by Slug, a zinc finger gene.," *Science*, vol. 264, no. 5160, pp. 835-839, 1994.

[131] E. Batlle et al., "The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.," *Nature cell biology*, vol. 2, no. 2, pp. 84-9, Mar. 2000.

[132] H. Peinado, D. Olmeda, and A. Cano, "Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?," *Nature reviews. Cancer*, vol. 7, no. 6, pp. 415-28, Jun. 2007.

[133] J. P. Thiery, "Epithelial-mesenchymal transitions in tumour progression.," *Nature reviews. Cancer*, vol. 2, no. 6, pp. 442-54, Jun. 2002.

[134] D. Tarin, E. W. Thompson, and D. F. Newgreen, "The fallacy of epithelial mesenchymal transition in neoplasia.," *Cancer research*, vol. 65, no. 14, pp. 5996-6000; discussion 6000-1, Jul. 2005.

[135] E. W. Thompson and Æ. E. D. Williams, "EMT and MET in carcinoma — clinical observations , regulatory pathways and new models," *Clinical Experimental Metastasis*, pp. 591-592, 2008.

[136] U. Jeschke et al., "Expression of E-cadherin in human ductal breast cancer carcinoma in situ, invasive carcinomas, their lymph node metastases, their distant metastases, carcinomas with recurrence and in recurrence.," *Anticancer research*, vol. 27, no. 4, pp. 1969-74, 2007.

[137] D. Park, R. Kåresen, U. Axcrona, T. Noren, and T. Sauer, "Expression pattern of adhesion molecules on survival in primary breast carcinoma , and their corresponding axillary lymph node metastasis," *October*, pp. 52-65, 2007.

[138] P. Carmeliet, F. De Smet, S. Loges, and M. Mazzone, "Branching morphogenesis and antiangiogenesis candidates: tip cells lead the way.," *Nature reviews Clinical oncology*, vol. 6, no. 6, pp. 315-326, 2009.

[139] P. Friedl and D. Gilmour, "Collective cell migration in morphogenesis, regeneration and cancer.," *Nature reviews. Molecular cell biology*, vol. 10, no. 7, pp. 445-57, Jul. 2009.

[140] A. Barrallo-Gimeno and M. A. Nieto, "The Snail genes as inducers of cell movement and survival: implications in development and cancer.," *Development (Cambridge, England)*, vol. 132, no. 14, pp. 3151-61, Jul. 2005.

[141] S. Vega, A. V. Morales, O. H. Ocaña, F. Valdés, I. Fabregat, and M. A. Nieto, "Snail blocks the cell cycle and confers resistance to cell death.," *Genes & development*, vol. 18, no. 10, pp. 1131-43, May 2004.

[142] S. Lehn et al., "Down-regulation of the oncogene cyclin D1 increases migratory capacity in breast cancer and is linked to unfavorable prognostic features.," *The American journal of pathology*, vol. 177, no. 6, pp. 2886-97, Dec. 2010.

[143] X. Wang et al., "Krüppel-like factor 8 induces epithelial to mesenchymal transition and epithelial cell invasion.," *Cancer Research*, vol. 67, no. 15, pp. 7184-7193, 2007.

[144] V. R. Sobrado et al., "The class I bHLH factors E2-2A and E2-2B regulate EMT.," *Journal of Cell Science*, vol. 122, no. 7, pp. 1014-1024, 2009.

[145] D. Olmeda, a Montes, G. Moreno-Bueno, J. M. Flores, F. Portillo, and a Cano, "Snai1 and Snai2 collaborate on tumor growth and metastasis properties of mouse skin carcinoma cell lines.," *Oncogene*, vol. 27, no. 34, pp. 4690-701, Aug. 2008.

[146] J. P. Thiery, H. Acloque, R. Y. J. Huang, and M. A. Nieto, "Epithelial-mesenchymal transitions in development and disease.," *Cell*, vol. 139, no. 5, pp. 871-90, Nov. 2009.

[147] E. Casas, J. Kim, A. Bendesky, L. Ohno-Machado, C. J. Wolfe, and J. Yang, "Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis.," *Cancer research*, vol. 71, no. 1, pp. 245-54, Jan. 2011.

[148] R. Lander, K. Nordin, and C. Labonne, "The F-box protein Ppa is a common regulator of core EMT factors Twist, Snail, Slug, and Sip1.," *The Journal of cell biology*, Jul. 2011.

[149] V. Bolós, H. Peinado, M. A. Pérez-Moreno, M. F. Fraga, M. Esteller, and A. Cano, "The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors.," *Journal of Cell Science*, vol. 116, no. 3, pp. 499-511, 2003.

[150] A. Barrallo-Gimeno and M. A. Nieto, "Evolutionary history of the Snail/Scratch superfamily.," *Trends in genetics : TIG*, vol. 25, no. 6, pp. 248-52, Jun. 2009.

[151] A. K. Perl, P. Wilgenbus, U. Dahl, H. Semb, and G. Christofori, "A causal role for E-cadherin in the transition from adenoma to carcinoma.," *Nature*, vol. 392, no. 6672, pp. 190-193, 1998.

[152] M. A. Nieto, "The snail superfamily of zinc-finger transcription factors," *Nature Reviews Molecular Cell Biology*, vol. 3, no. 3, pp. 155-166, 2002.

[153] M. A. Nieto, "The snail superfamily of zinc-finger transcription factors.," *Nature reviews. Molecular cell biology*, vol. 3, no. 3, pp. 155-66, Mar. 2002.

[154] S. Guaita et al., "Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression.," *The Journal of biological chemistry*, vol. 277, no. 42, pp. 39209-16, Oct. 2002.

[155] T. Ohkubo and M. Ozawa, "The transcription factor Snail downregulates the tight junction components independently of E-cadherin downregulation.," *Journal of Cell Science*, vol. 117, no. 9, pp. 1675-1685, 2004.

[156] H. Peinado, E. Ballestar, M. Esteller, and A. Cano, "Snail Mediates E-Cadherin Repression by the Recruitment of the Sin3A / Histone Deacetylase 1 (HDAC1)/ HDAC2 Complex," *Society*, vol. 24, no. 1, pp. 306-319, 2004.

[157] N. Herranz et al., "Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor.," *Molecular and cellular biology*, vol. 28, no. 15, pp. 4772-81, Aug. 2008.

[158] Q. Cao et al., "Repression of E-cadherin by the polycomb group protein EZH2 in cancer.," *Oncogene*, vol. 27, no. 58, pp. 7274-84, Dec. 2008.

[159] Y. Lin et al., "The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1.," *The EMBO journal*, vol. 29, no. 11, pp. 1803-16, Jun. 2010.

[160] G. Christofori, "Snail1 links transcriptional control with epigenetic regulation.," *The EMBO journal*, vol. 29, no. 11, pp. 1787-9, Jun. 2010.

[161] Y. Wu, B. M. Evers, and B. P. Zhou, "Small C-terminal Domain Phosphatase Enhances Snail Activity through Dephosphorylation," *The Journal of Biological Chemistry*, vol. 284, no. 1, pp. 640-648, 2009.

[162] B. P. Zhou et al., "Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.," *Nature cell biology*, vol. 6, no. 10, pp. 931-40, Oct. 2004.

[163] S. J. Grille et al., "The protein kinase Akt induces epithelial mesenchymal transition and promotes enhanced motility and invasiveness of squamous cell carcinoma lines.," *Cancer Research*, vol. 63, no. 9, pp. 2172-2178, 2003.

[164] P. Polakis, "Wnt signaling and cancer Wnt signaling and cancer," *Genes Development*, vol. 14, no. 15, pp. 1837-1851, 2000.

[165] Z. Yang, S. Rayala, D. Nguyen, R. K. Vadlamudi, S. Chen, and R. Kumar, "Pak1 phosphorylation of snail, a master regulator of epithelial-to-mesenchyme transition, modulates snail's subcellular localization and functions.," *Cancer Research*, vol. 65, no. 8, pp. 3179-3184, 2005.

[166] H. Peinado et al., "A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression.," *The EMBO journal*, vol. 24, no. 19, pp. 3446-58, Oct. 2005.

[167] J. T. Erler et al., "Lysyl oxidase is essential for hypoxia-induced metastasis.," *Nature*, vol. 440, no. 7088, pp. 1222-6, Apr. 2006.

[168] Y. Wu, J. Deng, P. G. Rychahou, S. Qiu, B. M. Evers, and B. P. Zhou, "Stabilization of snail by NF-kappaB is required for inflammation-induced cell migration and invasion.," *Cancer cell*, vol. 15, no. 5, pp. 416-28, May 2009.

[169] V. Evdokimova et al., "Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition.," *Cancer Cell*, vol. 15, no. 5, pp. 402-415, 2009.

[170] S. Yamashita, C. Miyagi, T. Fukada, N. Kagara, Y.-S. Che, and T. Hirano, "Zinc transporter LIVI controls epithelial-mesenchymal transition in zebrafish gastrula organizer.," *Nature*, vol. 429, no. 6989, pp. 298-302, 2004.

[171] J. Unno et al., "LIV-1 enhances the aggressive phenotype through the induction of epithelial to mesenchymal transition in human pancreatic carcinoma cells.," *International Journal of Oncology*, vol. 35, no. 4, pp. 813-821, 2009.

[172] N. S. Spoelstra et al., "The transcription factor ZEB1 is aberrantly expressed in aggressive uterine cancers.," *Cancer Research*, vol. 66, no. 7, pp. 3893-3902, 2006.

[173] S. Spaderna et al., "The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer.," *Cancer Research*, vol. 68, no. 2, pp. 537-544, 2008.

[174] J. E. Remacle et al., "New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites.," *the The European Molecular Biology Organization Journal*, vol. 18, no. 18, pp. 5073-5084, 1999.

[175] A. Eger et al., "DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells.," *Oncogene*, vol. 24, no. 7, pp. 2375-2385, 2005.

[176] J. Comijn et al., "The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion.," *Molecular Cell*, vol. 7, no. 6, pp. 1267-1278, 2001.

[177] C. Vandewalle et al., "SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell–cell junctions," *Nucleic Acids Research*, vol. 33, no. 20, pp. 6566-6578, 2005.

[178] M. Taki, K. Verschueren, K. Yokoyama, M. Nagayama, and N. Kamata, "Involvement of Ets-1 transcription factor in inducing matrix metalloproteinase-2 expression by epithelial-mesenchymal transition in human squamous carcinoma cells," *International Journal of Oncology*, vol. 27, no. 2, pp. 439-448, 2006.

[179] M. Beltran et al., "A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial–mesenchymal transition," *Genes Development*, vol. 22, no. 6, pp. 756-769, 2008.

[180] J. Lu et al., "MicroRNA expression profiles classify human cancers.," *Nature*, vol. 435, no. 7043, pp. 834-838, 2005.

[181] M. Korpal and Y. Kang, "The emerging role of miR-200 family of microRNAs in epithelila-mesenchymal transition and cancer metastasis," *Rna Biology*, no. September, pp. 115-119, 2008.

[182] N. R. Christoffersen, A. Silahtaroglu, U. A. Orom, S. Kauppinen, and A. H. Lund, "miR-200b mediates post-transcriptional repression of ZFHX1B," *Rna New York Ny*, vol. 13, no. 8, pp. 1172-1178, 2007.

[183] G. J. Hurteau, J. A. Carlson, S. D. Spivack, and G. J. Brock, "Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin.," *Cancer Research*, vol. 67, no. 17, pp. 7972-7976, 2007.

[184] S.-M. Park, A. B. Gaur, E. Lengyel, and M. E. Peter, "The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2.," *Genes & development*, vol. 22, no. 7, pp. 894-907, Apr. 2008.

[185] P. a Gregory et al., "The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.," *Nature cell biology*, vol. 10, no. 5, pp. 593-601, May 2008.

[186] M. Korpal, E. S. Lee, G. Hu, and Y. Kang, "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*, vol. 283, no. 22, pp. 14910-4, May 2008.

[187] X. Wang et al., "KLF8 promotes human breast cancer cell invasion and metastasis by transcriptional activation of MMP9.," *Oncogene*, vol. 30, no. 16, pp. 1901-11, Apr. 2011.

[188] Y. Hamamori, H.-yi Wu, and V. Sartorelli, "The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein , Twist . The Basic Domain of Myogenic Basic Helix-Loop-Helix (bHLH) Proteins Is the Novel Target for Direct Inhibiti," *Microbiology*, 1997.

[189] M. a Perez-Moreno et al., "A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions.," *The Journal of biological chemistry*, vol. 276, no. 29, pp. 27424-31, Jul. 2001.

[190] G. Moreno-Bueno et al., "Genetic profiling of epithelial cells expressing E-cadherin repressors reveals a distinct role for Snail, Slug, and E47 factors in epithelial-mesenchymal transition.," *Cancer research*, vol. 66, no. 19, pp. 9543-56, Oct. 2006.

[191] Q. Qin, Y. Xu, T. He, C. Qin, and J. Xu, "Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms.," *Cell Research*, vol. 22, no. 1, pp. 90-106, 2011.

[192] M. E. Massari and C. Murre, "Helix-Loop-Helix Proteins : Regulators of Transcription in Eucaryotic Organisms," *Society*, vol. 20, no. 2, pp. 429-440, 2000.

[193] S. Ansieau et al., "Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence.," *Cancer cell*, vol. 14, no. 1, pp. 79-89, Jul. 2008.

[194] M. Leptin, "twist and snail as positive and negative regulators during Drosophila mesoderm development.," *Genes & Development*, vol. 5, no. 9, pp. 1568-1576, 1991.

[195] J. Yang et al., "Twist , a Master Regulator of Morphogenesis , Plays an Essential Role in Tumor Metastasis Ben Gurion University of the Negev," *Plays*, vol. 117, pp. 927-939, 2004.

[196] M.-H. Yang et al., "Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition.," *Nature cell biology*, vol. 12, no. 10, pp. 982-92, Oct. 2010.

[197] S. a Mani et al., "The epithelial-mesenchymal transition generates cells with properties of stem cells.," *Cell*, vol. 133, no. 4, pp. 704-15, May 2008.

[198] J. Fu et al., "The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis.," *Cell research*, vol. 21, no. 2, pp. 275-89, Feb. 2011.

[199] F. Vesuna, P. Van Diest, J. H. Chen, and V. Raman, "Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer.," *Biochemical and Biophysical Research Communications*, vol. 367, no. 2, pp. 235-241, 2008.

[200] K. a Hartwell, B. Muir, F. Reinhardt, A. E. Carpenter, D. C. Sgroi, and R. a Weinberg, "The Spemann organizer gene, Goosecoid, promotes tumor metastasis.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 50, pp. 18969-74, Dec. 2006.

[201] D. M. Dykxhoorn et al., "miR-200 Enhances Mouse Breast Cancer Cell Colonization to Form Distant Metastases," *PLoS ONE*, vol. 4, no. 9, p. 14, 2009.

[202] P. J. Coffer and B. M. T. Burgering, "Forkhead-box transcription factors and their role in the immune system.," *Nature Reviews Immunology*, vol. 4, no. 11, pp. 889-899, 2004.

[203] S. S. Myatt and E. W.-F. Lam, "The emerging roles of forkhead box (Fox) proteins in cancer.," *Nature Reviews Cancer*, vol. 7, no. 11, pp. 847-859, 2007.

[204] L. Xu and J. Massagué, "Nucleocytoplasmic shuttling of signal transducers.," *Nature Reviews Molecular Cell Biology*, vol. 5, no. 3, pp. 209-219, 2004.

[205] E. L. Greer and A. Brunet, "FOXO transcription factors at the interface between longevity and tumor suppression.," *Oncogene*, vol. 24, no. 50, pp. 7410-7425, 2005.

[206] S. a Mani et al., "Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 24, pp. 10069-74, Jun. 2007.

[207] T. Kume, "The Role of FoxC2 Transcription Factor in Tumor Angiogenesis.," *Journal of oncology*, vol. 2012, p. 204593, Jan. 2012.

[208] L. Wyatt, C. Wadham, L. A. Crocker, M. Lardelli, and Y. Khew-Goodall, "The protein tyrosine phosphatase Pez regulates TGFβ, epithelial–mesenchymal transition, and organ development," *The Journal of Cell Biology*, vol. 178, no. 7, pp. 1223-1235, 2007.

[209] H. Wang, S. Y. Quah, J. M. Dong, E. Manser, J. P. Tang, and Q. Zeng, "PRL-3 down-regulates PTEN expression and signals through PI3K to promote epithelial-mesenchymal transition.," *Cancer Research*, vol. 67, no. 7, pp. 2922-2926, 2007.

[210] F. Liang, J. Liang, W.-Q. Wang, J.-P. Sun, E. Udho, and Z.-Y. Zhang, "PRL3 promotes cell invasion and proliferation by down-regulation of Csk leading to Src activation.," *The Journal of Biological Chemistry*, vol. 282, no. 8, pp. 5413-5419, 2007.

[211] E. Martín-Villar, D. Megías, S. Castel, M. M. Yurrita, S. Vilaró, and M. Quintanilla, "Podoplanin binds ERM proteins to activate RhoA and promote epithelial-mesenchymal transition.," *Journal of Cell Science*, vol. 119, no. 21, pp. 4541-4553, 2006.

[212] A. Wicki, F. Lehembre, N. Wick, B. Hantusch, D. Kerjaschki, and G. Christofori, "Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton.," *Cancer cell*, vol. 9, no. 4, pp. 261-72, Apr. 2006.

[213] M. Shtutman, E. Levina, P. Ohouo, M. Baig, and I. B. Roninson, "Cell adhesion molecule L1 disrupts E-cadherin-containing adherens junctions and increases scattering and motility of MCF7 breast carcinoma cells.," *Cancer research*, vol. 66, no. 23, pp. 11370-80, Dec. 2006.

[214] T. Waerner et al., "ILEI: a cytokine essential for EMT, tumor formation, and late events in metastasis in epithelial cells.," *Cancer cell*, vol. 10, no. 3, pp. 227-39, Sep. 2006.

[215] N. J. Sullivan et al., "Interleukin-6 induces an epithelial-mesenchymal transition phenotype in human breast cancer cells.," *Oncogene*, vol. 28, no. 33, pp. 2940-7, Aug. 2009.

[216] D. C. Radisky et al., "Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability.," *Nature*, vol. 436, no. 7047, pp. 123-7, Jul. 2005.

[217] S. A. Illman, K. Lehti, J. Keski-Oja, and J. Lohi, "Epilysin (MMP-28) induces TGF-beta mediated epithelial to mesenchymal transition in lung carcinoma cells.," *Journal of Cell Science*, vol. 119, no. 18, pp. 3856-3865, 2006.

[218] C. Billottet et al., "Modulation of several waves of gene expression during FGF-1 induced epithelial-mesenchymal transition of carcinoma cells.," *Journal of Cellular Biochemistry*, vol. 104, no. 3, pp. 826-839, 2008.

[219] G. Moreno-Bueno, F. Portillo, and A. Cano, "Transcriptional regulation of cell polarity in EMT and cancer.," *Oncogene*, vol. 27, no. 55, pp. 6958-6969, 2008.

[220] A. Gal, T. Sjöblom, L. Fedorova, S. Imreh, H. Beug, and A. Moustakas, "Sustained TGFb exposure suppresses Smad and non-Smad signalling in mammary epithelial cells , leading to EMT and inhibition of growth arrest and apoptosis," *Oncogene*, vol. 27, no. 2007, pp. 1218-30, 2008.

[221] Y. Liu, S. El-Naggar, D. S. Darling, Y. Higashi, and D. C. Dean, "Zeb1 links epithelial-mesenchymal transition and cellular senescence.," *Development Cambridge England*, vol. 135, no. 3, pp. 579-588, 2008.

[222] J.-P. Coppé et al., "Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor.," *PLoS biology*, vol. 6, no. 12, pp. 2853-68, Dec. 2008.

[223] K. L. Knutson et al., "Immunoediting of cancers may lead to epithelial to mesenchymal transition.," *The Journal of Immunology*, vol. 177, no. 3, pp. 1526-1533, 2006.

[224] S. E. Moody et al., "The transcriptional repressor Snail promotes mammary tumor recurrence.," *Cancer Cell*, vol. 8, no. 3, pp. 197-209, 2005.

[225] C. Kudo-Saito, H. Shirako, T. Takeuchi, and Y. Kawakami, "Cancer metastasis is accelerated through immunosuppression during Snail-induced EMT of cancer cells.," *Cancer Cell*, vol. 15, no. 3, pp. 195-206, 2009.

[226] W. Guo et al., "Slug and Sox9 Cooperatively Determine the Mammary Stem Cell State," *Cell*, vol. 148, no. 5, pp. 1015-1028, Mar. 2012.

[227] U. Wellner et al., "The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs.," *Nature cell biology*, vol. 11, no. 12, pp. 1487-95, Dec. 2009.

[228] A.-P. Morel, M. Lièvre, C. Thomas, G. Hinkal, S. Ansieau, and A. Puisieux, "Generation of breast cancer stem cells through epithelial-mesenchymal transition.," *PloS one*, vol. 3, no. 8, p. e2888, Jan. 2008.

[229] R. Li et al., "A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts.," *Cell stem cell*, vol. 7, no. 1, pp. 51-63, Jul. 2010.

[230] T. Tsuji et al., "Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth.," *Cancer research*, vol. 68, no. 24, pp. 10377-86, Dec. 2008.

[231] P. Friedl and K. Wolf, "Plasticity of cell migration: a multiscale tuning model.," *The Journal of cell biology*, vol. 188, no. 1, pp. 11-9, Jan. 2010.

[232] P. Friedl and S. Alexander, "Cancer Invasion and the Microenvironment: Plasticity and Reciprocity," *Cell*, vol. 147, no. 5, pp. 992-1009, Nov. 2011.

[233] A. J. Ridley et al., "Cell migration: integrating signals from front to back.," *Science*, vol. 302, no. 5651, pp. 1704-1709, 2003.

[234] P. Friedl, "Prespecification and plasticity: shifting mechanisms of cell migration.," *Current Opinion in Cell Biology*, vol. 16, no. 1, pp. 14-23, 2004.

[235] E. A. Cox, S. K. Sastry, and A. Huttenlocher, "Integrin-mediated adhesion regulates cell polarity and membrane protrusion through the Rho family of GTPases.," *Molecular Biology of the Cell*, vol. 12, no. 2, pp. 265-77, 2001.

[236] P. Friedl and K. Wolf, "Proteolytic interstitial cell migration: a five-step process.," *Cancer metastasis reviews*, vol. 28, no. 1-2, pp. 129-35, Jun. 2009.

[237] A. Estecha et al., "Moesin orchestrates cortical polarity of melanoma tumour cells to initiate 3D invasion.," *Journal of Cell Science*, vol. 122, no. 19, pp. 3492-3501, 2009.

[238] P. Friedl and K. Wolf, "Tube travel: the role of proteases in individual and collective cancer cell invasion.," *Cancer research*, vol. 68, no. 18, pp. 7247-9, Sep. 2008.

[239] C. Hidalgo-Carcedo et al., "Collective cell migration requires suppression of actomyosin at cell-cell contacts mediated by DDR1 and the cell polarity regulators Par3 and Par6.," *Nature Cell Biology*, vol. 13, no. 1, pp. 49-58, 2011.

[240] K. Wolf, R. Müller, S. Borgmann, E.-B. Bröcker, and P. Friedl, "Amoeboid shape change and contact guidance: T-lymphocyte crawling through fibrillar collagen is independent of matrix remodeling by MMPs and other proteases.," *Blood*, vol. 102, no. 9, pp. 3262-3269, 2003.

[241] D. Kedrin et al., "Intravital imaging of metastatic behavior through a mammary imaging window.," *Nature Methods*, vol. 5, no. 12, pp. 1019-1021, 2008.

[242] A. a Khalil and P. Friedl, "Determinants of leader cells in collective cell migration.," *Integrative biology : quantitative biosciences from nano to macro*, vol. 2, no. 11-12, pp. 568-74, Nov. 2010.

[243] C. Gaggioli et al., "Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells.," *Nature cell biology*, vol. 9, no. 12, pp. 1392-400, Dec. 2007.

[244] O. Ilina, G.-J. Bakker, A. Vasaturo, R. M. Hofmann, and P. Friedl, "Two-photon laser-generated microtracks in 3D collagen lattices: principles of MMP-dependent and -independent collective cancer cell invasion.," *Physical Biology*, vol. 8, no. 1, p. 015010, 2011.

[245] G. W. McLean, N. O. Carragher, E. Avizienyte, J. Evans, V. G. Brunton, and M. C. Frame, "The role of focal-adhesion kinase in cancer - a new therapeutic opportunity.," *Nature reviews. Cancer*, vol. 5, no. 7, pp. 505-15, Jul. 2005.

[246] K. T. Chan, C. L. Cortesio, and A. Huttenlocher, "FAK alters invadopodia and focal adhesion composition and dynamics to regulate breast cancer invasion.," *The Journal of cell biology*, vol. 185, no. 2, pp. 357-70, Apr. 2009.

[247] J. S. Desgrosellier and D. a Cheresh, "Integrins in cancer: biological implications and therapeutic opportunities.," *Nature reviews. Cancer*, vol. 10, no. 1, pp. 9-22, Jan. 2010.

[248] M. Zöller, "CD44: can a cancer-initiating cell profit from an abundantly expressed molecule?," *Nature reviews. Cancer*, vol. 11, no. 4, pp. 254-67, Apr. 2011.

[249] R. L. Brown, L. M. Reinke, M. S. Damerow, and D. Perez, "CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression," *Pharmacology*, vol. 121, no. 3, 2011.

[250] P. Grytsenko, O. Ilina, and P. Friedl, "Interstitial guidance of cancer invasion.," *The Journal of pathology*, no. 1, pp. 185-199, Oct. 2011.

[251] T. J. C. Harris and U. Tepass, "Adherens junctions: from molecules to morphogenesis.," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 7, pp. 502-514, 2010.

[252] G. Christofori, "New signals from the invasive front.," *Nature*, vol. 441, no. 7092, pp. 444-50, May 2006.

[253] T. T. Onder, P. B. Gupta, S. a Mani, J. Yang, E. S. Lander, and R. a Weinberg, "Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.," *Cancer research*, vol. 68, no. 10, pp. 3645-54, May 2008.

[254] N. Gavert, A. Ben-Shmuel, V. Lemmon, T. Brabletz, and A. Ben-Ze'ev, "Nuclear factor-kappaB signaling and ezrin are essential for L1-mediated metastasis of colon cancer cells.," *Journal of Cell Science*, vol. 123, no. 12, pp. 2135-2143, 2010.

[255] M. Munz, P. a Baeuerle, and O. Gires, "The emerging role of EpCAM in cancer and stem cell signaling.," *Cancer research*, vol. 69, no. 14, pp. 5627-9, Jul. 2009.

[256] S. D. Mason and J. a Joyce, "Proteolytic networks in cancer.," *Trends in cell biology*, vol. 21, no. 4, pp. 228-37, Apr. 2011.

[257] V. Gocheva and J. A. Joyce, "Cysteine Cathepsins and the Cutting Edge of Cancer Invasion ND ES SC Key words RIB," *New York*, vol. 6, no. 1, pp. 60-64, 2007.

[258] S. Alexander, G. E. Koehl, M. Hirschberg, E. K. Geissler, and P. Friedl, "Dynamic imaging of cancer growth and invasion: a modified skin-fold chamber model.," *Histochemistry and cell biology*, vol. 130, no. 6, pp. 1147-54, Dec. 2008.

[259] N. Gavert et al., "Expression of L1-CAM and ADAM10 in human colon cancer cells induces metastasis.," *Cancer Research*, vol. 67, no. 16, pp. 7703-7712, 2007.

[260] K. Wolf et al., "Compensation mechanism in tumor cell migration," *The Journal of Cell Biology*, vol. 160, no. 2, pp. 267-277, 2003.

[261] V. Sanz-Moreno and C. J. Marshall, "The plasticity of cytoskeletal dynamics underlying neoplastic cell migration.," *Current opinion in cell biology*, vol. 22, no. 5, pp. 690-6, Oct. 2010.

[262] V. Sanz-Moreno et al., "ROCK and JAK1 signaling cooperate to control actomyosin contractility in tumor cells and stroma.," *Cancer cell*, vol. 20, no. 2, pp. 229-45, Aug. 2011.

[263] Y. Soda et al., "Transdifferentiation of glioblastoma cells into vascular endothelial cells.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 11, pp. 4274-4280, 2011.

[264] A. Raza, M. J. Franklin, and A. Z. Dudek, "Pericytes and vessel maturation during tumor angiogenesis and metastasis.," *American Journal of Hematology*, vol. 85, no. 8, pp. 593-598, 2010.

[265] E. C. Finger and A. J. Giaccia, "Hypoxia, inflammation, and the tumor microenvironment in metastatic disease.," *Cancer metastasis reviews*, vol. 29, no. 2, pp. 285-93, Jun. 2010.

[266] C. Murdoch, M. Muthana, S. B. Coffelt, and C. E. Lewis, "The role of myeloid cells in the promotion of tumour angiogenesis.," *Nature Reviews Cancer*, vol. 8, no. 8, pp. 618-631, 2008.

[267] M. De Palma, C. Murdoch, M. A. Venneri, L. Naldini, and C. E. Lewis, "Tie2-expressing monocytes: regulation of tumor angiogenesis and therapeutic implications.," *Trends in Immunology*, vol. 28, no. 12, pp. 519-524, 2007.

[268] J. A. Joyce and J. W. Pollard, "Microenvironmental regulation of metastasis.," *Nature Reviews Cancer*, vol. 9, no. 4, pp. 239-52, 2009.

[269] B.-Z. Qian and J. W. Pollard, "Macrophage diversity enhances tumor progression and metastasis.," *Cell*, vol. 141, no. 1, pp. 39-51, 2010.

[270] R. Kalluri and M. Zeisberg, "Fibroblasts in cancer.," *Nature reviews. Cancer*, vol. 6, no. 5, pp. 392-401, May 2006.

[271] M. Shimoda, K. T. Mellody, and A. Orimo, "Carcinoma-associated fibroblasts are a rate-limiting determinant for tumour progression," *Seminars in cell developmental biology*, vol. 21, no. 1, pp. 19-25, 2010.

[272] D. C. Radisky, "news and views Leading the charge," *Group*, vol. 9, no. 12, pp. 1341-1342, 2007.

[273] F. T. Martin et al., "Potential role of mesenchymal stem cells (MSCs) in the breast tumour microenvironment: stimulation of epithelial to mesenchymal transition (EMT).," *Breast cancer research and treatment*, vol. 124, no. 2, pp. 317-26, Nov. 2010.

[274] K. Kessenbrock, V. Plaks, and Z. Werb, "Matrix metalloproteinases: regulators of the tumor microenvironment.," *Cell*, vol. 141, no. 1, pp. 52-67, 2010.

[275] D. a Chan and A. J. Giaccia, "Hypoxia, gene expression, and metastasis.," *Cancer metastasis reviews*, vol. 26, no. 2, pp. 333-9, Jun. 2007.

[276] C. Sahlgren, M. V. Gustafsson, S. Jin, L. Poellinger, and U. Lendahl, "Notch signaling mediates hypoxia-induced tumor cell migration and invasion.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 17, pp. 6392-7, Apr. 2008.

[277] A. J. Evans et al., "VHL promotes E2 box-dependent E-cadherin transcription by HIF-mediated regulation of SIP1 and snail.," *Molecular and cellular biology*, vol. 27, no. 1, pp. 157-69, Jan. 2007.

[278] B. Krishnamachary et al., "Hypoxia-inducible factor-1-dependent repression of E-cadherin in von Hippel-Lindau tumor suppressor-null renal cell carcinoma mediated by TCF3, ZFHX1A, and ZFHX1B.," *Cancer research*, vol. 66, no. 5, pp. 2725-31, Mar. 2006.

[279] H. Peinado and A. Cano, "A hypoxic twist in metastasis.," *Nature cell biology*, vol. 10, no. 3, pp. 253-4, Mar. 2008.

[280] G. Kroemer and J. Pouyssegur, "Tumor cell metabolism: cancer's Achilles' heel.," *Cancer cell*, vol. 13, no. 6, pp. 472-82, Jun. 2008.

[281] G. L. Semenza, "Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy.," *Trends in pharmacological sciences*, pp. 1-8, Mar. 2012.

[282] X. Lu and Y. Kang, "Hypoxia and hypoxia-inducible factors: master regulators of metastasis.," *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 16, no. 24, pp. 5928-35, Dec. 2010.

[283] S. J. Lunt, N. Chaudary, and R. P. Hill, "The tumor microenvironment and metastatic disease.," *Clinical & experimental metastasis*, vol. 26, no. 1, pp. 19-34, Jan. 2009.

[284] C. Jose, N. Bellance, and R. Rossignol, "Choosing between glycolysis and oxidative phosphorylation: A tumor's dilemma?," *Biochimica et biophysica acta*, vol. 1807, no. 6, pp. 552-61, Jun. 2011.

[285] M. I. Koukourakis, A. Giatromanolaki, A. L. Harris, and E. Sivridis, "Comparison of metabolic pathways between cancer cells and stromal cells in colorectal carcinomas: a metabolic survival role for tumor-associated stroma.," *Cancer research*, vol. 66, no. 2, pp. 632-7, Jan. 2006.

[286] S. Dhup, R. K. Dadhich, P. E. Porporato, and P. Sonveaux, "Multiple biological activities of lactic Acid in cancer: influences on tumor growth, angiogenesis and metastasis.," *Current pharmaceutical design*, vol. 18, no. 10, pp. 1319-30, Apr. 2012.

[287] B. W. Booth, D. L. Mack, A. Androutsellis-Theotokis, R. D. G. McKay, C. a Boulanger, and G. H. Smith, "The mammary microenvironment alters the differentiation repertoire of neural stem cells.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 39, pp. 14891-6, Sep. 2008.

[288] M. J. C. Hendrix, E. a Seftor, R. E. B. Seftor, J. Kasemeier-Kulesa, P. M. Kulesa, and L.-M. Postovit, "Reprogramming metastatic tumour cells with embryonic microenvironments.," *Nature reviews. Cancer*, vol. 7, no. 4, pp. 246-55, Apr. 2007.

[289] A. H. Ting, K. M. Mcgarvey, and S. B. Baylin, "The cancer epigenome — components and functional correlates The cancer epigenome — components and functional correlates," *Genes & Development*, pp. 3215-3231, 2006.

[290] B. Li, M. Carey, and J. L. Workman, "Review The Role of Chromatin during Transcription," *Cell*, pp. 707-719, 2007.

[291] R. Margueron and D. Reinberg, "Chromatin structure and the inheritance of epigenetic information.," *Nature reviews. Genetics*, vol. 11, no. 4, pp. 285-96, Apr. 2010.

[292] I. B. Weinstein, "Addiction to Oncogenes — the Achilles Heal of Cancer," *Science*, vol. 297, no. July, pp. 64-65, 2002.

[293] S. B. Baylin and J. E. Ohm, "Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction?," *Nature reviews. Cancer*, vol. 6, no. 2, pp. 107-16, Feb. 2006.

[294] O. G. Mcdonald, H. Wu, W. Timp, A. Doi, and A. P. Feinberg, "Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition," *Online*, no. May, 2011.

[295] N. Dumont, M. B. Wilson, Y. G. Crawford, P. a Reynolds, M. Sigaroudinia, and T. D. Tlsty, "Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 39, pp. 14867-72, Sep. 2008.

[296] H. P. Mohammad and S. B. Baylin, "c o m m e n tA r y Linking cell signaling and the epigenetic machinery," *Nature Publishing Group*, vol. 28, no. 10, pp. 1033-1038, 2010.

[297] B. E. Bernstein et al., "A bivalent chromatin structure marks key developmental genes in embryonic stem cells.," *Cell*, vol. 125, no. 2, pp. 315-326, 2006.

[298] D. R. Hurst, M. D. Edmonds, and D. R. Welch, "Metastamir: the field of metastasis-regulatory microRNA is spreading.," *Cancer research*, vol. 69, no. 19, pp. 7495-8, Oct. 2009.

[299] A. L. Kasinski and F. J. Slack, "MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy," *Nature Reviews Cancer*, vol. 11, no. 12, pp. 849-864, Nov. 2011.

[300] S. F. Tavazoie et al., "Endogenous human microRNAs that suppress breast cancer metastasis.," *Nature*, vol. 451, no. 7175, pp. 147-52, Jan. 2008.

[301] D. Iliopoulos, M. Lindahl-Allen, C. Polytarchou, H. a Hirsch, P. N. Tsichlis, and K. Struhl, "Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells.," *Molecular cell*, vol. 39, no. 5, pp. 761-72, Sep. 2010.

[302] P. S. Mongroo and A. K. Rustgi, "The role of the miR-200 family in epithelial-mesenchymal transition," *Therapy*, pp. 219-222, 2010.

[303] Peiro, S., Escriva, M., Puig, I., Barbera, M.J., Dave, N., Herranz, N., Larriba, M.J., Takkunen, M., Franci, C., Munoz, A., et al. 2006. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res* 34:2077-2084

[304] R. Kalluri and R. A. Weinberg, "Review series The basics of epithelial-mesenchymal transition," vol. 119, no. 6, 2009.

[305] H. Zhang, Y. Li, and M. Lai, "The microRNA network and tumor metastasis.," *Oncogene*, vol. 29, no. 7, pp. 937-48, Feb. 2010.

[306] A. Barroso-delJesus et al., "Embryonic Stem Cell-Specific miR302-367 Cluster: Human Gene Structure and Functional Characterization of Its Core Promoter," *Molecular and Cellular Biology*, vol. 28, no. 21, pp. 6609-6619, 2008.

[307] R. R. Chivukula and J. T. Mendell, "Abate and switch: miR-145 in stem cell differentiation.," *Cell*, vol. 137, no. 4, pp. 606-8, May 2009.

[308] A. Barroso-del Jesus, G. Lucena-Aguilar, and P. Menendez, "The miR-302-367 cluster as a potential stemness regulator in ESCs.," *Cell cycle Georgetown Tex*, vol. 8, no. 3, pp. 394-398, 2009.

[309] D. a G. Card et al., "Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells.," *Molecular and cellular biology*, vol. 28, no. 20, pp. 6426-38, Oct. 2008.

[310] Y. Tay, J. Zhang, A. M. Thomson, B. Lim, and I. Rigoutsos, "MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation.," *Nature*, vol. 455, no. 7216, pp. 1124-1128, 2008.

[311] I. Büssing, F. J. Slack, and H. Grosshans, "let-7 microRNAs in development, stem cells and cancer.," *Trends in molecular medicine*, vol. 14, no. 9, pp. 400-9, Sep. 2008.

[312] N. Xu, T. Papagiannakopoulos, G. Pan, J. a Thomson, and K. S. Kosik, "MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells.," *Cell*, vol. 137, no. 4, pp. 647-58, May 2009.

[313] I. Ibarra, Y. Erlich, S. K. Muthuswamy, R. Sachidanandam, and G. J. Hannon, "A role for microRNAs in maintenance of mouse mammary epithelial progenitor cells.," *Genes & development*, vol. 21, no. 24, pp. 3238-43, Dec. 2007.

[314] C. Liu et al., "The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44.," *Nature medicine*, vol. 17, no. 2, pp. 211-5, Feb. 2011.

[315] Y. J. Choi et al., "miR-34 miRNAs provide a barrier for somatic cell reprogramming.," *Nature cell biology*, vol. 13, no. 11, pp. 1353-60, Nov. 2011.

[316] J. Torres and F. M. Watt, "Nanog maintains pluripotency of mouse embryonic stem cells by inhibiting NFkappaB and cooperating with Stat3.," *Nature cell biology*, vol. 10, no. 2, pp. 194-201, Feb. 2008.

[317] D. Sarrio, C. K. Franklin, A. Mackay, J. S. Reis-Filho, and C. M. Isacke, "Epithelial and mesenchymal subpopulations within normal basal breast cell lines exhibit distinct stem cell/progenitor properties.," *Stem cells (Dayton, Ohio)*, vol. 30, no. 2, pp. 292-303, Feb. 2012.

[318] J. G. Gill et al., "Snail and the microRNA-200 family act in opposition to regulate epithelial-to-mesenchymal transition and germ layer fate restriction in differentiating ESCs.," *Stem cells (Dayton, Ohio)*, vol. 29, no. 5, pp. 764-76, May 2011.

[319] X. Fang et al., "ChIP-seq and functional analysis of the SOX2 gene in colorectal cancers.," *Omics : a journal of integrative biology*, vol. 14, no. 4, pp. 369-84, Aug. 2010.

[320] a Singh and J. Settleman, "EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer.," *Oncogene*, vol. 29, no. 34, pp. 4741-51, Aug. 2010.

[321] I. Uribesalgo, C. Ballaré, and L. Di Croce, "Polycomb Regulates NF-κB Signaling in Cancer through miRNA," *Cancer Cell*, vol. 21, no. 1, pp. 5-7, Jan. 2012.

[322] V. Kashyap et al., "Regulation of stem cell pluripotency and differentiation involves a mutual regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs.," *Stem cells and development*, vol. 18, no. 7, pp. 1093-108, Sep. 2009.

[323] M.-Q. Gao, Y.-P. Choi, S. Kang, J. H. Youn, and N.-H. Cho, "CD24+ cells from hierarchically organized ovarian cancer are enriched in cancer stem cells.," *Oncogene*, vol. 29, no. 18, pp. 2672-80, May 2010.

[324] P. C. Hermann et al., "Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer.," *Cell stem cell*, vol. 1, no. 3, pp. 313-23, Sep. 2007.

[325] M. F. Pittenger, "Multilineage Potential of Adult Human Mesenchymal Stem Cells," *Science*, vol. 284, no. 5411, pp. 143-147, 1999.

[326] R. Pang et al., "A subpopulation of CD26+ cancer stem cells with metastatic capacity in human colorectal cancer.," *Cell stem cell*, vol. 6, no. 6, pp. 603-15, Jun. 2010.

[327] P. A. Lysy, C. Sibille, M. Najimi, and E. M. Sokal, "Human Skin Fibroblasts: From Mesodermal to Hepatocyte-Like Differentiation."

[328] P. Samavarchi-Tehrani et al., "Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming.," *Cell stem cell*, vol. 7, no. 1, pp. 64-77, Jul. 2010.

[329] J. M. Polo and K. Hochedlinger, "When fibroblasts MET iPSCs.," *Cell stem cell*, vol. 7, no. 1, pp. 5-6, Jul. 2010.

[330] J. G. Gill et al., "Snail and the microRNA-200 family act in opposition to regulate epithelial-to-mesenchymal transition and germ layer fate restriction in differentiating ESCs.," *Stem cells (Dayton, Ohio)*, vol. 29, no. 5, pp. 764-76, May 2011.

[331] Y. Shimono et al., "Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells.," *Cell*, vol. 138, no. 3, pp. 592-603, Aug. 2009.

[332] C.-H. Lin, A. L. Jackson, J. Guo, P. S. Linsley, and R. N. Eisenman, "Myc-regulated microRNAs attenuate embryonic stem cell differentiation.," *The EMBO journal*, vol. 28, no. 20, pp. 3157-70, Oct. 2009.

[333] T. M. Yeung, S. C. Gandhi, J. L. Wilding, R. Muschel, and W. F. Bodmer, "Cancer stem cells from colorectal cancer-derived cell lines.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 8, pp. 3722-7, Feb. 2010.

[334] R. L. Brown et al., "CD44 splice isoformswitching in human and mouse epithelium is essential for epithelila-mesenchymal transition and breast cancer progression," *Journal of Clinical Investigation*, vol. 121, no. 3, pp. 1064-1074, March 2011.

[335] L. Ricci-Vitiani et al., "Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells.," *Nature*, vol. 468, no. 7325, pp. 824-8, Dec. 2010.

[336] J. M. Kozlowski, I. J. Fidler, D. Campbell, Z. L. Xu, M. E. Kaighn, and I. R. Hart, "Metastatic behavior of human tumor cell lines grown in the nude mouse.," *Cancer research*, vol. 44, no. 8, pp. 3522-9, Aug. 1984.

[337] C. L. Chaffer, J. P. Brennan, J. L. Slavin, T. Blick, E. W. Thompson, and E. D. Williams, "Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2.," *Cancer research*, vol. 66, no. 23, pp. 11271-8, Dec. 2006.

[338] C. L. Chaffer et al., "Upregulated MT1-MMP/TIMP-2 axis in the TSU-Pr1-B1/B2 model of metastatic progression in transitional cell carcinoma of the bladder.," *Clinical & experimental metastasis*, vol. 22, no. 2, pp. 115-25, Jan. 2005.

[339] A. Marusyk and K. Polyak, "Tumor heterogeneity: causes and consequences.," *Biochimica et biophysica acta*, vol. 1805, no. 1, pp. 105-17, Jan. 2010.

[340] E. M. Hurt, B. T. Kawasaki, G. J. Klarmann, S. B. Thomas, and W. L. Farrar, "CD44+CD24− prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis," *British Journal of Cancer*, vol. 98, no. 4, pp. 756-765, 2008.

[341] A. P. Rybak, L. He, A. Kapoor, J.-C. Cutz, and D. Tang, "Characterization of sphere-propagating cells with stem-like properties from DU145 prostate cancer cells.," *Biochimica et Biophysica Acta*, vol. 1813, no. 5, pp. 683-94, 2011.

[342] C. Li et al., "Identification of pancreatic cancer stem cells.," *Cancer research*, vol. 67, no. 3, pp. 1030-7, Feb. 2007.

[343] K. F. Kelly, D. Y. Ng, G. Jayakumaran, G. a Wood, H. Koide, and B. W. Doble, "β-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism.," *Cell stem cell*, vol. 8, no. 2, pp. 214-27, Feb. 2011.

[344] T. Sauka-Spengler and M. Bronner-Fraser, "A gene regulatory network orchestrates neural crest formation.," *Nature Reviews Molecular Cell Biology*, vol. 9, no. 7, pp. 557-568, 2008.

[345] S. Vega, A. V. Morales, O. H. Ocaña, F. Valdés, I. Fabregat, and M. A. Nieto, "Snail blocks the cell cycle and confers resistance to cell death," *Genes & Development*, pp. 1131-1143, 2004.

[346] J. Liu et al., "Slug inhibits proliferation of human prostate cancer cells via downregulation of cyclin D1 expression.," *The Prostate*, vol. 70, no. 16, pp. 1768-77, Dec. 2010.

[347] D. Olmeda, G. Moreno-Bueno, J. M. Flores, A. Fabra, F. Portillo, and A. Cano, "SNAI1 is required for tumor growth and lymph node metastasis of human breast carcinoma MDA-MB-231 cells.," *Cancer research*, vol. 67, no. 24, pp. 11721-31, Dec. 2007.

[348] X.-S. Ke et al., "Epithelial to mesenchymal transition of a primary prostate cell line with switches of cell adhesion modules but without malignant transformation.," *PloS one*, vol. 3, no. 10, p. e3368, Jan. 2008.

[349] T. Tsuji, S. Ibaragi, and G.-fu Hu, "Epithelial-mesenchymal transition and cell cooperativity in metastasis.," *Cancer research*, vol. 69, no. 18, pp. 7135-9, Sep. 2009.

[350] D. M. Dykxhoorn et al., "miR-200 enhances mouse breast cancer cell colonization to form distant metastases.," *PloS one*, vol. 4, no. 9, p. e7181, Jan. 2009.

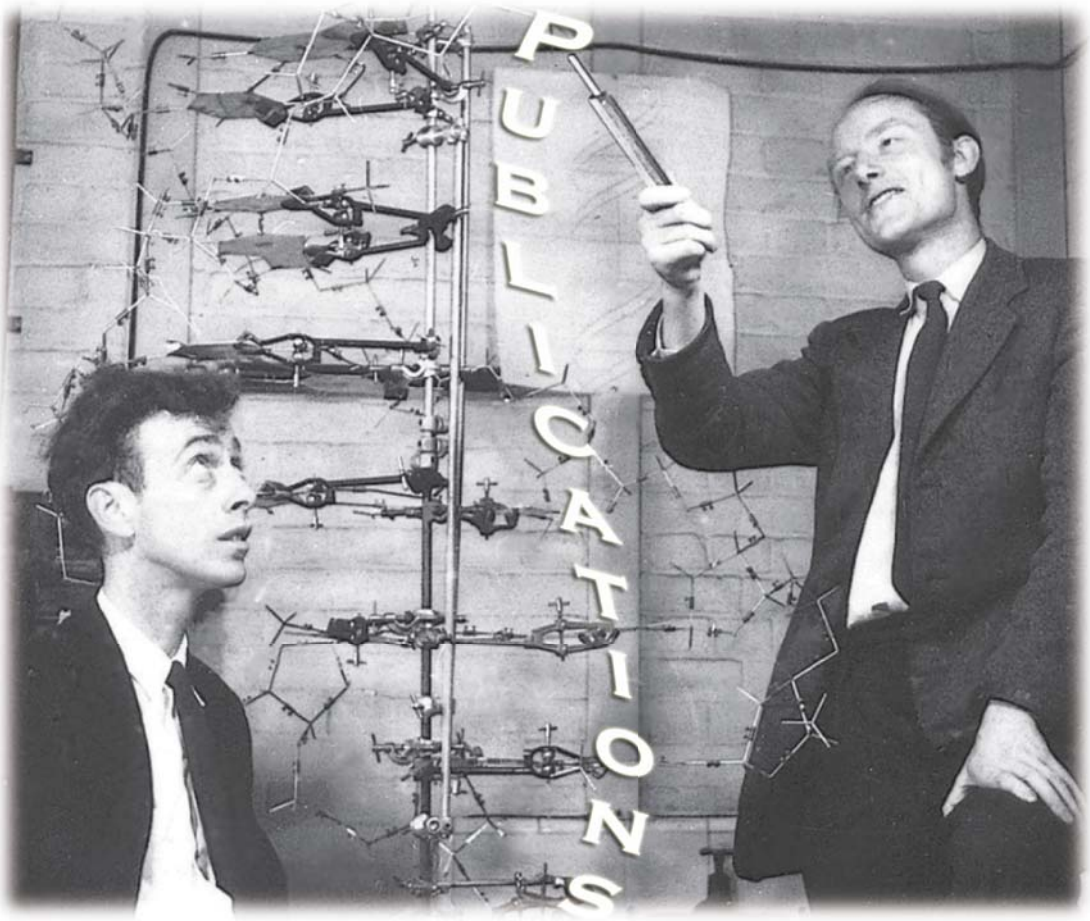
[351] Y. Lou et al., "Epithelial-mesenchymal transition (EMT) is not sufficient for spontaneous murine breast cancer metastasis.," *Developmental dynamics : an official publication of the American Association of Anatomists*, vol. 237, no. 10, pp. 2755-68, Oct. 2008.

[352] U. Burk et al., "A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells.," *EMBO reports*, vol. 9, no. 6, pp. 582-9, Jun. 2008.

[353] C. P. Bracken et al., "A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition.," *Cancer research*, vol. 68, no. 19, pp. 7846-54, Oct. 2008.

[354] D. L. Gibbons et al., "Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression.," *Genes & Development*, vol. 23, no. 18, pp. 2140-2151, 2009.

[355] P. Olson et al., "MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer," *Genes & Development*, vol. 23, no. 18, pp. 2152-2165, 2009.


PUBLICATIONS

James D. Watson (Chicago, 1928 - Lives in New York); left side of the picture.
Francis Crick (Northampton, 1916 - San Diego, 2004); right side of the picture.

INFORMATION AND VALIDATION FROM THESIS SUPERVISOR OF PUBLICATIONS IMPACT FACTOR AND TASK DEVELOPED IN EACH PUBLICATION

The work developed by Toni Celià-Terrassa during his PhD thesis has yielded the publication of two articles of high profile in prestigious international journals. He is the first and main author of a study published in the *Journal of Clinical Investigation* and the sixth author of a study published in *Nature* Medicine. The *Journal of Clinical Investigation* is a top-tier venue for critical advances in biomedical research with a 2010 impact factor of 14.152. The 2010 impact factor for *Nature Medicine* was 25.430. This places *Nature Medicine* as the top primary research journal in Medicine.

Articles and participation of the PhD student Toni-Celià Terrassa:

- Article 1: **Toni Celià-Terrassa,** Óscar Meca-Cortés, Francesca Mateo, Alexia Martínez de Paz, Nuria Rubio, Anna Arnal-Estapé, Brian J. Ell, Raquel Bermudo, Alba Díaz, Marta Guerra-Rebollo, Juan José Lozano, Conchi Estarás, Catalina Ulloa, Daniel Álvarez-Simón, Jordi Milà, Ramón Vilella, Rosanna Paciucci, Marian Martínez-Balbás, Antonio García de Herreros, Roger R. Gomis, Yibin Kang, Jerónimo Blanco, Pedro L. Fernández, and Timothy M. Thomson. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. Published in *Journal of Clinical Investigation;* May 2012 (Impact factor: 14.152).

  Toni Celià-Terrassa was responsible for initiating this study with no prior information on the cellular models that form the core of this work, for which he personally designed, developed and executed experiments to tackle most of the issues addressed in this study. In addition to performing many of the experiments of this paper, he analyzed and interpreted the results, and coordinated, along with the senior author and his thesis advisor, many of the experiments performed by contributing co-authors. He also participated in the writing of the manuscript.

- *Article 2:* Manav Korpal, Brian J Ell, Francesca M Buffa, Toni Ibrahim, Mario A Blanco, **Toni Celià-Terrassa**, Laura Mercatali, Zia Khan, Hani Goodarzi, Yuling Hua, Yong Wei, Guohong Hu, Benjamin A Garcia, Jiannis Ragoussis, Dino Amadori, Adrian L Harris & Yibin Kang. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. Published in *Nature Medicine*; 7 August 2011 (Impact factor: 25.430).

  In this work, Toni Celià-Terrassa participated in the identification of novel metastatatic secreted suppressors and its experimental validation, analysis and interpretation. He also participated in the review process for several of the experimental data.

Timothy M. Thomson Okatsu, MD, PhD
Thesis supervisor

Article 1
Journal of Clinical Investigation
Published online 16/04/12

Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells

Toni Celià-Terrassa, Óscar Meca-Cortés, Francesca Mateo, Alexia Martínez de Paz, Nuria Rubio, Anna Arnal-Estapé, Brian J. Ell, Raquel Bermudo, Alba Díaz, Marta Guerra-Rebollo, Juan José Lozano, Conchi Estarás, Catalina Ulloa, Daniel Álvarez-Simón, Jordi Milà, Ramón Vilella, Rosanna Paciucci, Marian Martínez-Balbás, Antonio García de Herreros, Roger R. Gomis, Yibin Kang, Jerónimo Blanco, Pedro L. Fernández, and Timothy M. Thomson

Research article

Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells

Toni Celià-Terrassa,[1] Óscar Meca-Cortés,[1] Francesca Mateo,[1] Alexia Martínez de Paz,[1] Nuria Rubio,[2] Anna Arnal-Estapé,[3] Brian J. Ell,[4] Raquel Bermudo,[5,6] Alba Díaz,[6] Marta Guerra-Rebollo,[2] Juan José Lozano,[7] Conchi Estarás,[8] Catalina Ulloa,[1] Daniel Álvarez-Simón,[1] Jordi Milà,[9] Ramón Vilella,[9] Rosanna Paciucci,[10] Marian Martínez-Balbás,[8] Antonio García de Herreros,[11] Roger R. Gomis,[3,12] Yibin Kang,[4] Jerónimo Blanco,[2] Pedro L. Fernández,[5,6,13] and Timothy M. Thomson[1]

[1]Department of Cell Biology, Barcelona Institute of Molecular Biology, Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain. [2]Cardiovascular Research Center, CSIC–Institut Català de Ciències Cardiovaculars (CSIC-ICCC), Centro de Investigaciones Biomédicas en Red en Biomateriales, Bioingeniería y Naomedicina (CIBER-BBN), Barcelona, Spain. [3]Oncology Programme, Institute for Research in Biomedicine (IRB-Barcelona), Barcelona, Spain. [4]Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA. [5]Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain. [6]Department of Pathology, Hospital Clínic, Barcelona, Spain. [7]CIBER de Enfermedades Hepáticas y Digestivas, Barcelona, Spain. [8]Department of Molecular Biology, Barcelona Institute of Molecular Biology, CSIC, Barcelona, Spain. [9]Department of Immunology, Hospital Clínic, Barcelona, Spain. [10]Vall d'Hebrón Research Institute, Barcelona, Spain. [11]Programa de Recerca en Càncer, Institut Municipal d'Investigaciones Mèdiques (IMIM) Hospital del Mar, Barcelona, Spain. [12]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. [13]University of Barcelona Medical School, Barcelona, Spain.

Malignant progression in cancer requires populations of tumor-initiating cells (TICs) endowed with unlimited self renewal, survival under stress, and establishment of distant metastases. Additionally, the acquisition of invasive properties driven by epithelial-mesenchymal transition (EMT) is critical for the evolution of neoplastic cells into fully metastatic populations. Here, we characterize 2 human cellular models derived from prostate and bladder cancer cell lines to better understand the relationship between TIC and EMT programs in local invasiveness and distant metastasis. The model tumor subpopulations that expressed a strong epithelial gene program were enriched in highly metastatic TICs, while a second subpopulation with stable mesenchymal traits was impoverished in TICs. Constitutive overexpression of the transcription factor Snai1 in the epithelial/TIC-enriched populations engaged a mesenchymal gene program and suppressed their self renewal and metastatic phenotypes. Conversely, knockdown of EMT factors in the mesenchymal-like prostate cancer cell subpopulation caused a gain in epithelial features and properties of TICs. Both tumor cell subpopulations cooperated so that the nonmetastatic mesenchymal-like prostate cancer subpopulation enhanced the in vitro invasiveness of the metastatic epithelial subpopulation and, in vivo, promoted the escape of the latter from primary implantation sites and accelerated their metastatic colonization. Our models provide new insights into how dynamic interactions among epithelial, self-renewal, and mesenchymal gene programs determine the plasticity of epithelial TICs.

Introduction

There is a wealth of evidence that the acquisition of aggressive traits of cancer, or malignant progression, can be determined both by the occurrence of genetic mutations and by the imposition of heritable epigenetic marks on relevant genes (1). Within a tumor, these newly acquired genetic and epigenetic events can emerge either sequentially within a single lineage or in parallel in multiple, independent lineages (2). In either scenario of cancer cell evolution, the final outcome is the coexistence in a given tumor of different subpopulations of tumor cells, each endowed with particular phenotypes (intratumoral heterogeneity). There is also evidence that transcriptional reprogramming in tumor cells can be induced in response to nontumor environmental cues that include factors such as TGF-β, PDGF, or EGF (3), hormones, or hypoxic stress (4). Therefore, cancer cells endowed with a capacity for indefinite self renewal (cancer stem cells [CSCs]), but still retaining some capacity for differentiation, could evolve into distinct phenotypes in response to environmental cues and to new mutations. It has been proposed that, as in any ecological niche (5), these subpopulations could interact among each other, either by competing for common resources (6) or by cooperating for mutual benefit (2, 7). These tumoral subpopulations can also interact with, and use to their advantage, nontumoral elements, as has been convincingly shown in many models of tumor progression and metastasis (8).

Tumor-initiating cells (TICs) constitute subpopulations of cells capable of initiating and sustaining the growth of tumors in immunodeficient mice (9–11). In turn, TICs and CSCs share with ES and adult stem cells gene networks that are essential for self renewal and pluripotency (12, 13). Independent of their origin, it is still unclear whether CSCs are a population of tumor cells endowed with irreversible self-renewal properties or whether they are subject to dynamic influences that can affect their phenotypes (14, 15). A second process and gene program critical for cancer progression is epithelial-mesenchymal transition (EMT) (16–19). Whether induced by environmental cues or by other

Conflict of interest: The authors have declared that no conflict of interest exists.
Citation for this article: *J Clin Invest* doi:10.1172/JCI59218.

research article

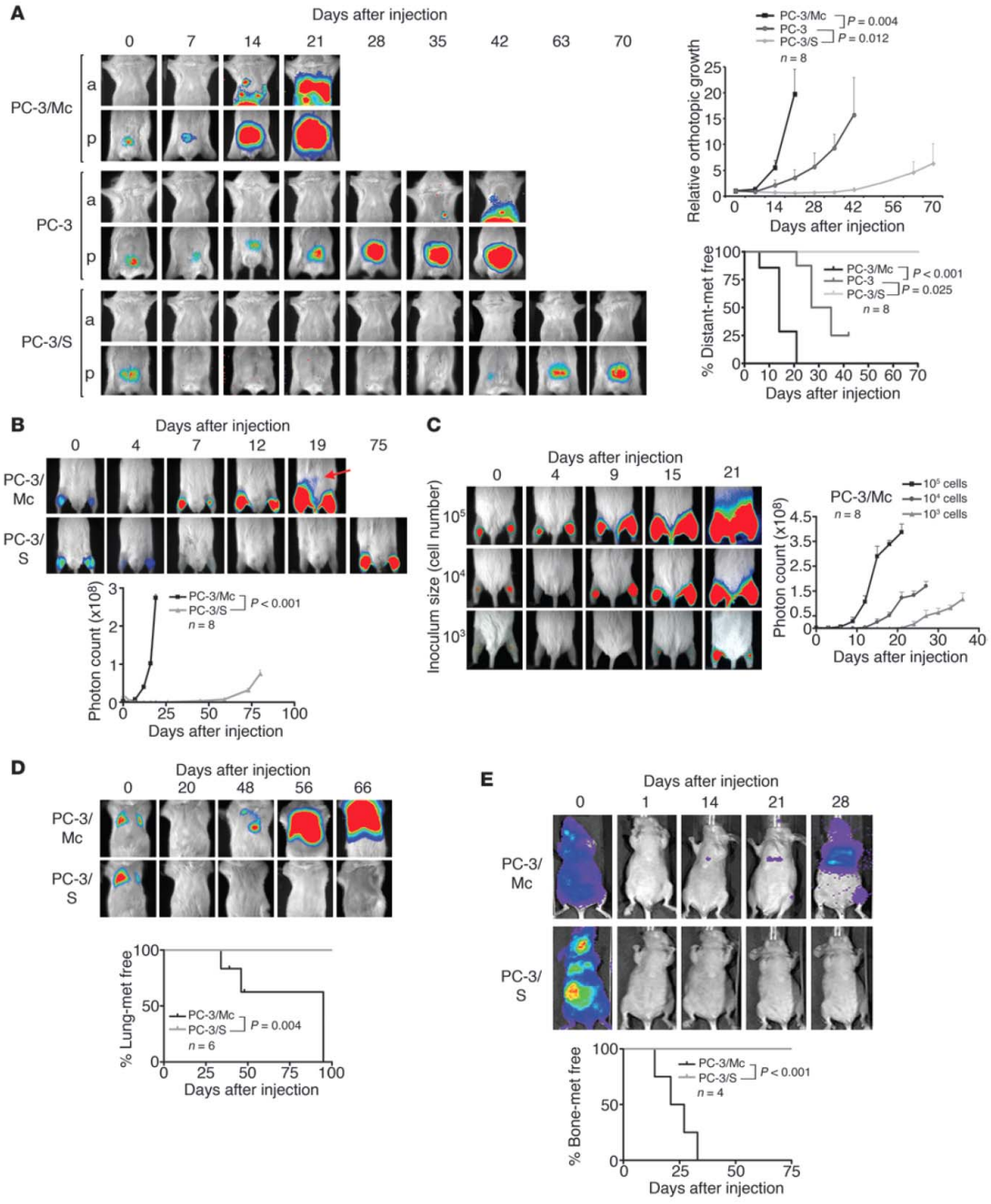


Figure 1
Divergent growth and metastatic potentials of 2 clonal populations derived from PC-3 prostate cancer cells. (**A**) PC-3/Mc, but not PC-3/S, cells rapidly formed tumors upon orthotopic implantation in NOD-SCID mice, developing lymph node and distant metastases as early as 14 days after implantation. Parental PC-3 cells grew and metastasized with efficiencies intermediate between the 2 clonal populations. Cells (1.0×10^5) were implanted in the ventral lobes of 6-week-old male mice. Anterior (a) or posterior (p) halves were imaged independently for enhanced resolution. Upper right panel: growth curves of orthotopic tumors, with photon counts normalized to values on day 0. Lower right panel: Kaplan-Meier plots for metastasis-free (met free) mice. (**B**) PC-3/Mc cells grew rapidly after i.m. grafting (2.0×10^5 cells), with detection in lymph nodes after 19 days (arrow). PC-3/S cells formed tumors after 75 days, without detectable distant spread. Bottom panel: growth curves at the i.m. implantation sites. (**C**) Grafting of limited numbers of PC-3/Mc cells readily produced tumors. 10^5, 10^4, or 10^3 cells were injected i.m. in each hind limb. Right panel: growth curves at the i.m. implantation site. (**D**) PC-3/Mc, but not PC-3/S, cells readily colonized lungs upon i.v. injection (2.5×10^5 cells). Bottom panel: Kaplan-Meier plots for lung colonization–free mice at each time point. (**E**) PC-3/Mc, but not PC-3/S, cells readily colonized bones upon i.c. injection (2.0×10^5 cells). Bottom panel: Kaplan-Meier plots for bone metastasis–free mice. Results are expressed as mean ± SEM.

mechanisms, EMT is driven by transcriptional factors such as *SNAI1/2*, *ZEB1/2*, or *TWIST1/2*, results in enhanced migration and invasive potentials of epithelial cells, and is critical for the metastatic spread of epithelial tumors (16, 20, 21). In several models of cancer, the induction of EMT potentiates self renewal and the acquisition of CSC properties (22–24). Consequently, a common notion is that EMT may be a general feature of cancer stem or progenitor populations, associating local invasiveness with the ability to colonize distant organs as expressions of 2 tightly interdependent gene programs borne by the same tumor cells (15, 21). However, other models of neoplasia have found an inverse correlation between local invasiveness and the ability of tumor cells to colonize distant organs (25), suggesting a dichotomy between these 2 critical features of the metastatic process, possibly expressed by separate tumor cell subpopulations (26, 27) in which tumor cells that display a strong epithelial phenotype are endowed with the strongest capacity to survive in circulation and to establish distant metastases (14, 25, 28).

To better understand the relationship between CSC and EMT programs in local invasiveness and distant metastasis, we have characterized 2 cellular models, derived from prostate and bladder cancer cell lines, displaying a dissociation between these 2 programs. We have found that forced induction of constitutive EMT in subpopulations of tumor cells displaying relatively stable epithelial/TIC features caused the suppression of major properties associated with TICs, including anchorage-independent growth and metastatic potential. Conversely, knockdown in the TIC-enriched epithelial subpopulations of self-renewal genes and of E-cadherin led, in addition to an inhibition of anchorage-independent growth, to a loss of their epithelial features, enhanced invasiveness, and an inhibition of their capacity to colonize distant organs. These observations closely link properties of metastatic TICs to an epithelial phenotype and gene program and suggest that, in our models, EMT, while enabling the invasiveness of tumor cells, opposes the self-renewal gene program that drives their local and metastatic growth.

Results

Cellular models of metastasis in which TIC and EMT properties are dissociated. The PC-3 prostate cancer cell line was used to generate 2 distinct clonal populations. PC-3/S cells were isolated in vitro by single-cell cloning from luciferase-expressing PC-3 cells (29). A second single-cell progeny, hereafter designated PC-3/Mc, was isolated from luciferase-expressing PC-3/M cells, a PC-3 subline that had been selected in vivo for its high metastatic potential (30). Orthotopic implantation in the ventral prostate lobe of NOD-SCID mice of 1.0×10^5 PC-3/Mc cells quickly produced large tumors, spreading to regional lymph nodes shortly after implantation and to distant organs at later times (Figure 1A). In contrast, PC-3/S cells grew slowly and were not detected outside of the implantation site for the duration of monitoring (70 days). Parental PC-3 cells displayed an intermediate behavior in local growth rate and in the dissemination to regional lymph nodes and distant sites (Figure 1A). Intramuscular (i.m.) grafting corroborated the remarkable differences in tumorigenicity between these 2 cell subpopulations (Figure 1B). Grafting of limited numbers of PC-3/Mc cells yielded robust tumor growth (Figure 1C), and they could be serially transplanted in immunodeficient mice, maintaining or gaining their efficiency for local growth upon successive transplantations (Supplemental Figure 1A; supplemental material available online with this article; doi:10.1172/JCI59218DS1). They also readily colonized lungs and bones after i.v. (Figure 1D) or intracardiac (i.c.) (Figure 1E and Supplemental Figure 1B) injection, suggesting enrichment in metastasis-prone TICs. In contrast, PC-3/S cells did not grow detectable colonies in lungs or bones after i.v. or i.c. injection at any time during monitoring (Figure 1, D and E). Therefore, the PC-3/Mc and PC-3/S subpopulations of PC-3 prostate cancer cells display highly contrasting phenotypes, with PC-3/Mc cells enriched in serially transplantable TICs with high metastatic potential.

In vitro, PC-3/Mc cells grew much faster and had 1.5-fold more cells in the S phase of the cell cycle than PC-3/S cells (Figure 2A and Supplemental Figure 2A). Likewise, PC-3/Mc cells readily formed large spheroids under nonadherent growth conditions (Figure 2B and Supplemental Figure 2B) and maintained this capacity upon serial plating (Supplemental Figure 2C), whereas PC-3/S cells showed limited anchorage-independent growth (Figure 2B and Supplemental Figure 2B). Contrary to our expectations, PC-3/Mc cells were barely invasive in Matrigel-Boyden chamber assays, while PC-3/S cells were highly invasive and motile (Figure 2, C and F). Thus, the in vitro invasiveness and motility of PC-3/Mc and PC-3/S cells are inversely correlated to their in vitro spheroid-forming and proliferative potentials. This suggests a dichotomy in these cells between 2 processes that determine the capacity of tumor cells to metastasize, namely the capacities to grow serially transplantable tumors in vivo and spheroids in vitro and to invade through extracellular matrix in vitro.

Gene profiling revealed a striking divergence in transcriptional programs between these 2 subpopulations derived from a common parental cell line (Supplemental Table 1 and Supplemental Figure 3). PC-3/Mc cells expressed an epithelial gene program, including E-cadherin (*CDH1*), EpCAM (*TACSTD1*), and desmoplakin (*DSP*), and also genes associated with pluripotency and self renewal (31, 32), including *KLF4*, *MYC*, *SOX2*, *KLF9*, and *LIN28A* (Figure 2, D, E, G, and Supplemental Figure 3A). Gene set enrichment analysis revealed that PC-3/Mc cells have very active DNA repair, DNA replication, and mitotic transition and checkpoint

research article

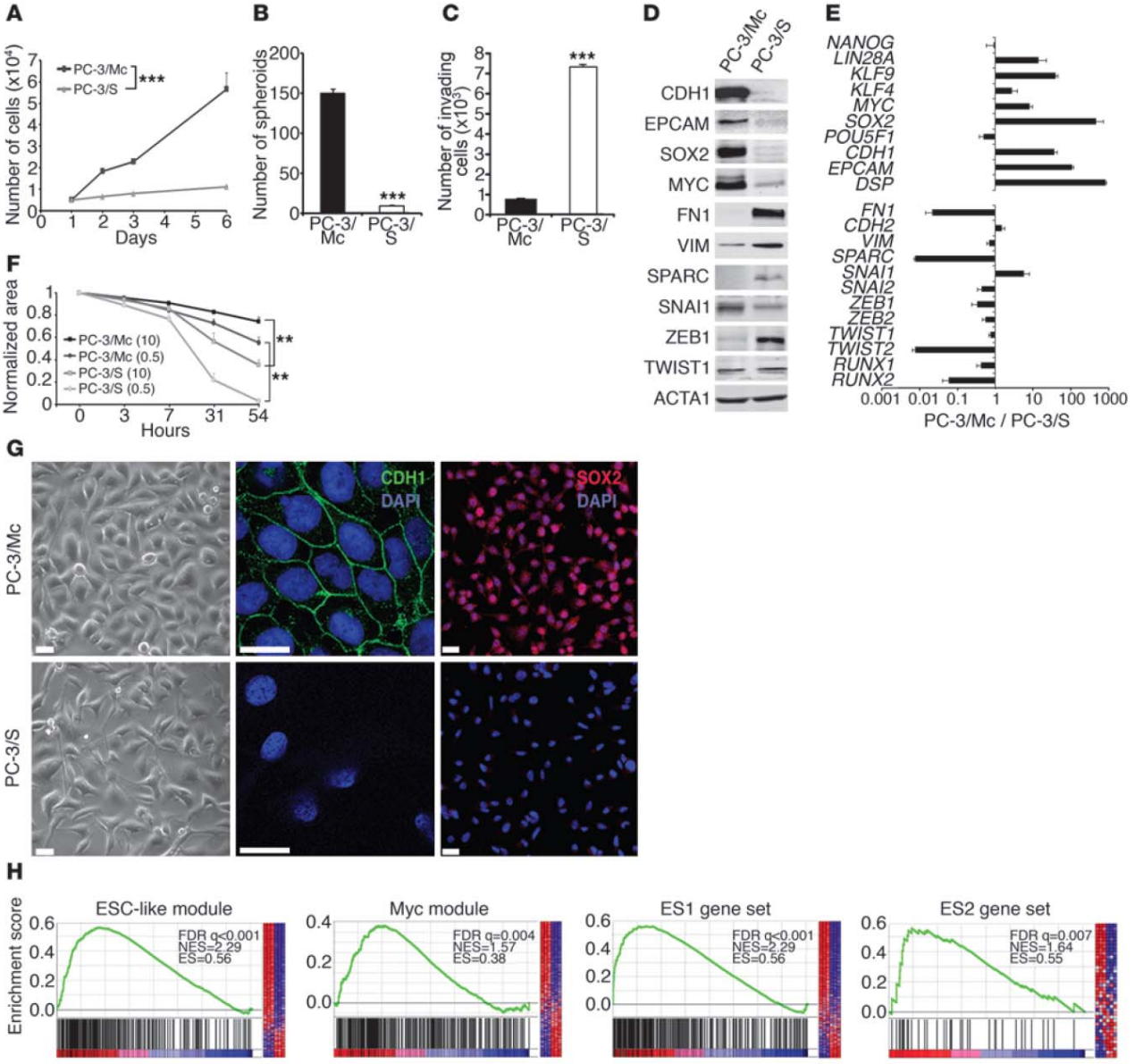


Figure 2

Opposing phenotypes and distinct gene programs expressed by 2 clonal populations derived from PC-3 cells. (**A**) PC-3/Mc cells grew with short doubling times (22–24 hours), while PC-3/S cells grew with long doubling times (60–72 hours). (**B**) PC-3/Mc, but not PC-3/S, cells displayed robust anchorage-independent growth. Cells (10^3) seeded in low-attachment plates in the presence of 0.5% methyl cellulose were scored for spheroids after 14 days (triplicate assays). (**C**) PC-3/Mc cells were barely invasive, while PC-3/S cells were highly invasive. Cells seeded on the upper chamber of Matrigel- and hyaluronic acid–coated Transwell units were scored for invading cells after 24 hours (triplicate assays). (**D**) PC-3/Mc cells expressed higher levels than PC-3/S cells of E-cadherin and EpCAM. PC-3/S cells expressed higher levels than PC-3/Mc cells of fibronectin, vimentin, and SPARC, by Western blotting. (**E**) PC-3/Mc cells expressed higher levels than PC-3/S cells of genes associated with self renewal and pluripotency. PC-3/S cells expressed higher levels than PC-3/Mc cells of genes associated with mesenchymal phenotypes and EMT. Relative transcript levels are represented as the $\log_{10}$ of ratios between the 2 cell lines of their $2^{-\Delta\Delta Cp}$ real-time PCR values. (**F**) PC-3/S cells were more motile than PC-3/Mc cells in wound-healing assays (triplicate assays). Parentheses denote percentages of FBS. (**G**) PC-3/Mc cells were round and expressed membrane-associated E-cadherin and nuclear SOX2. PC-3/S cells were flat and spindled and with undetectable E-cadherin. Scale bar: 20 μm. (**H**) Gene-set enrichment analysis (GSEA) showing significant enrichment in PC-3/Mc cells of the ESC-like, MYC, ES1, and ES2 gene modules. FDR q, false discovery rate *q* value; NES, normalized enrichment score; ES, enrichment score. Results are expressed as mean ± SEM. $^{**}P < 0.01$; $^{***}P < 0.001$.

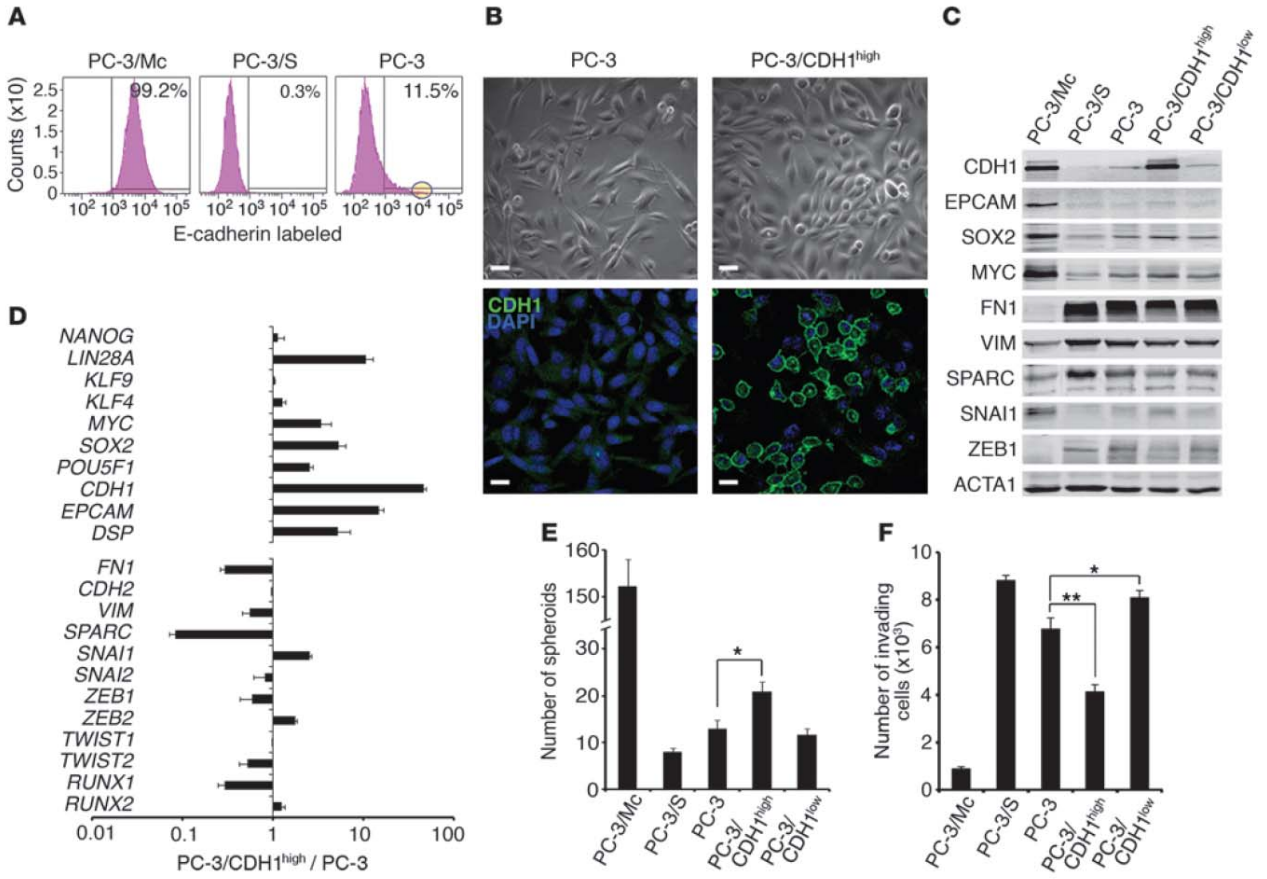


Figure 3

E-cadherin–positive PC-3 cells show an enhanced anchorage-independent growth and a stronger expression of a self-renewal gene program relative to parental or E-cadherin–negative cells. (**A**) Over 99% of PC-3/Mc cells were positive, and 0.3% of PC-3/S cells were positive for surface E-cadherin. A minor fraction (11.5%) of parental PC-3 prostate cancer cells expressed cell-surface E-cadherin. The circle on the right panel indicates the 1% sorted population with the highest CDH1 expression (PC-3/CDH1^{hi}). (**B**) The bulk of parental PC-3 cells displayed a spindled morphology and low levels of membrane-bound E-cadherin. Most PC-3/CDH1^{hi} cells displayed a round morphology and a strong expression of membrane-bound E-cadherin. Scale bars: 20 μm. (**C**) PC-3/CDH1^{hi} cells expressed higher levels of MYC and SOX2 and lower levels of the mesenchymal markers fibronectin or ZEB1 than PC-3/S or PC-3/CDH1^{lo} cells, as determined by Western blotting. (**D**) PC-3/CDH1^{hi} cells expressed self-renewal/ pluripotency genes at levels significantly higher than parental PC-3 cells, as determined by real-time qPCR. Relative transcript levels are represented as the $\log_{10}$ of ratios between the 2 subpopulations of their $2^{-\Delta\Delta Cp}$ real-time PCR values. (**E**) PC-3/CDH1^{hi} cells grew more spheroids than E-cadherin–negative (PC-3/CDH1^{lo}) or parental PC-3 cells. For comparison, the spheroid growth of PC-3/Mc and PC-3/S cells is also illustrated. (**F**) PC-3/CDH1^{hi} cells were less invasive in Transwell-Matrigel assays than PC-3/CDH1^{lo} or parental PC-3 cells. For comparison, the invasiveness of PC-3/Mc and PC-3/S cells is also illustrated. Results are expressed as mean ± SEM. *$P < 0.05$; **$P < 0.01$.

gene networks (Supplemental Table 2 and Supplemental Figure 3, A and B). Importantly, PC-3/Mc cells were strongly enriched in an ES cell–like module (ESC-like module) shown to be highly active in epithelial cancers associated with metastasis and death (13), with 265 of the 335 genes of this module overrepresented in PC-3/Mc cells and also in a *MYC* gene module (33) and ES1 and ES2 gene sets (ref. 12, Figure 2H, and Supplemental Table 4). This supports the conclusion that PC-3/Mc cells, which have a high potential for anchorage-independent and metastatic growth but are poorly invasive in vitro, display both an epithelial phenotype and a very active self-renewal/pluripotency gene program. In contrast, PC-3/S cells expressed high levels of many mesenchymal markers (e.g., *VIM*, *SPARC*, and *FN1*) and genes linked to EMT, such as *TWIST2*, *SNAI2*, *ZEB1*, and *RUNX2* (Figure 2, D and E, and Supplemental Figure 3, C and D). Of interest, PC-3/S cells expressed many genes for chemokines and inflammatory cytokines and their receptors at levels much higher than those of PC-3/Mc cells (Supplemental Tables 3 and 5 and Supplemental Figure 3, C and D), suggesting that this subpopulation has engaged a proinflammatory program similar to that induced in cells under stress or in presenescent states (34, 35). Intriguingly, PC-3/Mc cells expressed higher levels than PC-3/S cells of the EMT factor SNAI1. The endogenous SNAI1 protein showed a correct nuclear localization in PC-3/Mc cells, clearly visible when allowed to accumulate upon treatment with the GSK3 inhibitor LiCl or the proteasome inhibitor MG132 (Supplemental Figure 4A).

research article

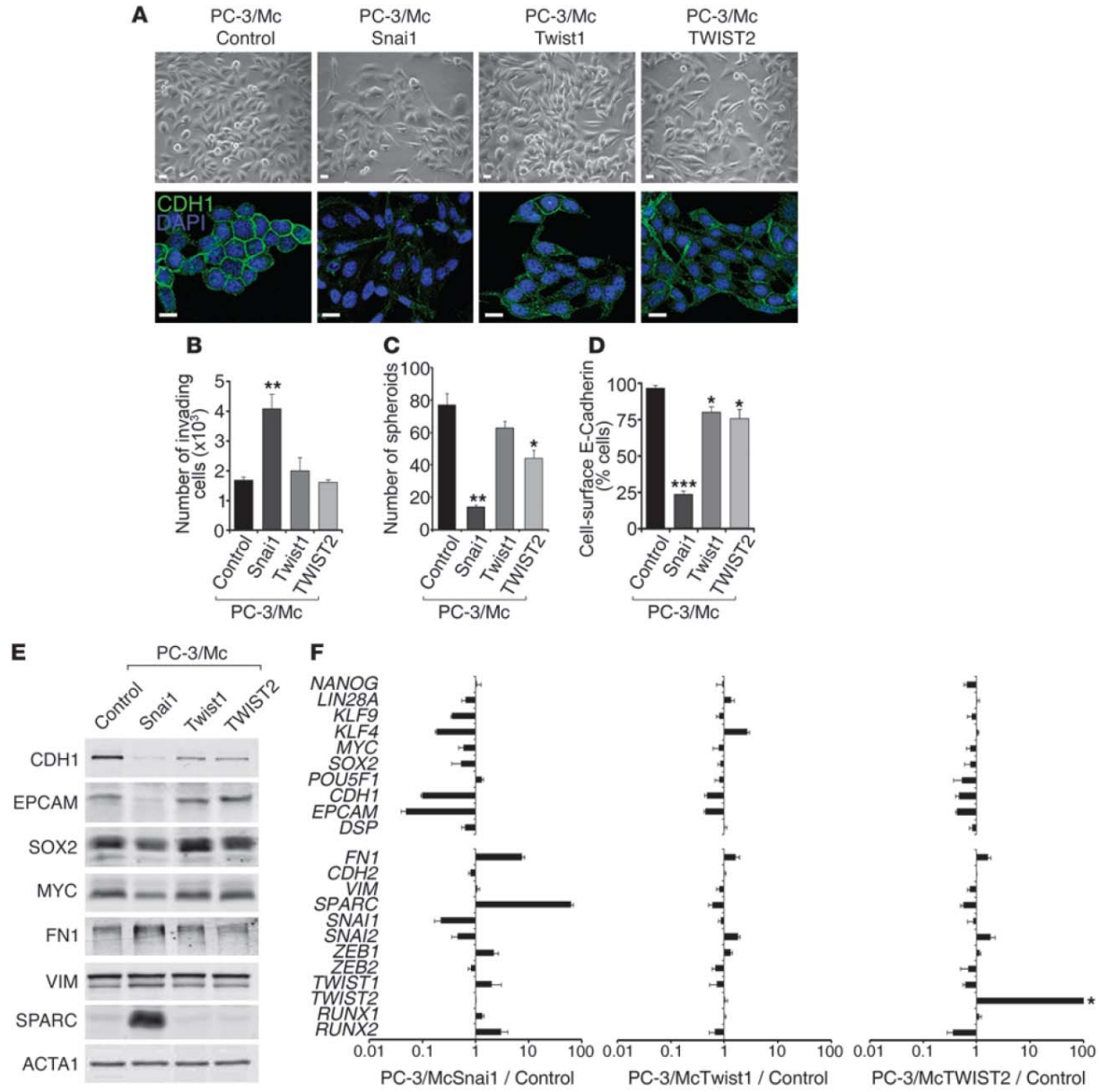


Figure 4

Overexpression of *Snai1* in PC-3/Mc cells induces EMT and suppresses anchorage-independent growth and the expression of a self-renewal gene program. (**A**) Overexpression of *Snai1*, *Twist1*, or *TWIST2* in PC-3/Mc cells induced a fibroblastoid morphology and a downregulation of membrane-associated E-cadherin. Cells were transduced with retroviruses for the expression of mouse *Snai1* or *Twist1* or human *TWIST2*. Controls were PC-3/Mc cells transduced with pBABE and selected for puromycin resistance. Scale bars: 20 μm. (**B**) Overexpression of *Snai1* strongly induced the invasiveness of PC-3/Mc cells, with a moderate effect by *Twist1* or *TWIST2*. (**C**) Overexpression of *Snai1* strongly inhibited spheroid growth by PC-3/Mc cells, with a moderate effect by *TWIST2*. (**D**) Overexpression of *Snai1* in PC-3/Mc cells caused a strong downregulation of cell-surface E-cadherin, with a moderate effect by *Twist1* or *TWIST2*, as determined by flow cytometry. (**E**) Overexpression of *Snai1* in PC-3/Mc cells induced a downregulation of E-cadherin and EpCAM, a modest downregulation of SOX2 and MYC, and an upregulation of fibronectin and SPARC, as determined by Western blotting. Overexpression of *Twist1* or *TWIST2* induced a moderate downregulation of E-cadherin. (**F**) Overexpression of *Snai1* and, more moderately, *Twist1* or *TWIST2*, caused a downregulation of self-renewal and epithelial genes and an upregulation of mesenchymal genes. Relative transcript levels are represented as the $\log_{10}$ of ratios between experimental and control cells of their $2^{-\Delta\Delta Cp}$ real-time PCR values. The levels of *SNAI1* correspond to the endogenous, human transcripts, downregulated by overexpression of the exogenous (mouse) *Snai1*. Asterisk in **F** indicates that values for ectopic *TWIST2* are off scale. Results are expressed as mean ± SEM. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

However, knockdown of endogenous *SNAI1* in PC-3/Mc cells did not significantly alter the levels of expression of E-cadherin or other epithelial markers, suggesting a defect in the function of endogenous *SNAI1* in these cells (Supplemental Figure 4, B and C), possibly explaining why the expression of this factor in PC-3/Mc cells still allows the expression of high levels of E-cadherin and a strong epithelial phenotype.

Analysis of histone marks associated with relevant promoters supported the transcriptional basis for these divergent expression profiles (Supplemental Figure 5). Thus, the *SOX2* and E-cadherin (*CDH1*) promoters were enriched in acetylated histone H4 in PC-3/Mc but not in PC-3/S and were more enriched in the H3K27me$_3$ repressive mark in PC-3/S. Conversely, the promoters of the mesenchymal genes *TWIST2* and *RUNX2* were enriched in acetylated histones H3 and H4 only in PC-3/S cells and were impoverished in H3K27me$_3$ in either cell type.

We next determined whether the epithelial-aggressive versus mesenchymal-nonaggressive dichotomy observed in our PC-3 prostate cancer cell line subpopulations applied to other models for which epithelial tumor cell subpopulations with distinct potentials for growth and metastasis had been characterized. We chose a cellular model derived from the human bladder cancer cell line T24 (TSU-Pr1 and TSU-Pr1-B2) (36, 37). It had been previously shown that the more epithelial TSU-Pr1-B2 cells are more tumorigenic and metastatic than the more mesenchymal TSU-Pr1 subpopulation and can colonize bones after i.c. inoculation in immunodeficient mice (36, 37). We confirmed that TSU-Pr1-B2 cells display features of epithelial cells when compared with the more mesenchymal TSU-Pr1 subpopulation: higher E-cadherin and desmoplakin expression levels and lower levels of fibronectin and EMT factors (*SNAI2*, *ZEB1*, *ZEB2*) (Supplemental Figure 6, A–C). Compared with the mesenchymal-like TSU-Pr1 cells, epithelial TSU-Pr1-B2 cells expressed higher levels of the pluripotency factors *SOX2*, *LIN28A*, *NANOG*, and *KLF9* (Supplemental Figure 6, A–C). Functionally, the more epithelial TSU-Pr1-B2 cells formed significantly more and larger spheroids than the mesenchymal-like TSU-Pr1 cells, but were significantly less invasive in vitro (Supplemental Figure 6, D and E). Upon i.c. injection in immunocompromised mice, the more epithelial TSU-Pr1-B2 cells established metastases to the bones and other organs significantly more efficiently and at earlier times than the more mesenchymal TSU-Pr1 cells (Supplemental Figure 6F). Thus, the TSU-Pr1 and TSU-Pr1-B2 bladder cancer dual cell model share gene expression and functional features with the PC-3 prostate cancer model described above.

The maintenance of critical properties of TICs is associated with an epithelial gene program in prostate and bladder cancer cells. We next determined whether selection from parental PC-3 cells of a subpopulation with epithelial features could enrich for cells with higher anchorage-independent growth potential. About 11% of parental PC-3 cells expressed high levels of E-cadherin (Figure 3A). Sorting of cells at the top 1% level of CDH1 expression selected for a subpopulation (PC-3/CDH1^{hi}) that expressed higher levels than parental or CDH1^{lo} PC-3 cells of epithelial markers such as *CDH1* and *EPCAM*, and also *LIN28A*, *SOX2*, *MYC*, *POU5F1/OCT4*, and *KLF4* (Figure 3, B–D). PC-3/CDH1^{hi} cells formed significantly more spheroids (Figure 3E) and were less invasive (Figure 3F) than parental or CDH1^{lo} PC-3 cells. Additional cell-surface phenotyping showed that most PC-3/Mc cells were CD44^{hi}CD24^{hi}CD71^{hi}CD40^{lo}, while most PC-3/S cells were CD44^{hi}CD24^{lo}CD71^{lo}CD40^{hi} (Supplemental Table 6). Although several studies have associated prostate and breast CSCs with a CD44^{hi}CD24^{lo} profile (10, 38, 39), other reports have found that CSCs and aggressive tumors can express high levels of CD24 (12, 40, 41). Transferrin receptor (CD71) is expressed in actively cycling compartments in different tissues (42). Therefore, PC-3 cells enriched for an epithelial phenotype show a stronger expression of self-renewal/pluripotency gene networks.

To explore whether maintenance of an epithelial gene program is important for the properties of PC-3/Mc or TSU-Pr1-B2 cells, we induced a mesenchymal program by transduction and overexpression of the EMT transcription factors *Snai1*, *Twist1*, and *TWIST2* (Supplemental Figure 7). This caused, in addition to the expected changes to more fibroblastoid morphologies (Figure 4A and Supplemental Figure 8A) and enhanced invasiveness (Figure 4B and Supplemental Figure 8B), a reduced formation of spheroids (Figure 4C and Supplemental Figure 8C), in particular in response to the overexpression of *Snai1*. These phenotypic changes were accompanied with a downregulation of the self-renewal/pluripotency factors *KLF4*, *SOX2*, and *MYC* in addition to a downregulation of the epithelial markers E-cadherin, EpCAM, and desmoplakin and upregulation of the mesenchymal markers fibronectin and SPARC (Figure 4, D–F, and Supplemental Figure 8, D–F). As expected (43), the constitutive overexpression of exogenous *SNAI1* in PC-3/Mc cells strongly suppressed the expression of endogenous *Snai1* transcripts (Figure 4F and Supplemental Figure 8F). The switch in transcriptional programs caused by high levels of exogenous *Snai1* was accompanied by an enrichment of the repressive histone mark H3K27me$_3$ and depletion of the active transcription marks acetylated histones H3 and H4 at the promoters of *SOX2* and *CDH1* (Supplemental Figure 9). In addition, the overexpression of *Snai1* in PC-3/Mc cells caused a decreased growth rate (Supplemental Figure 10A) and decreased the number of cells in the S phase of the cell cycle (Supplemental Figure 10B). In vivo, overexpression of *Snai1* in PC-3/Mc cells led to a significant inhibition of local growth upon orthotopic (Figure 5A) or i.m. implantation (Figure 5B) as well as inhibition of their capacity to spread to regional lymph nodes and distant sites (Figure 5A) and to colonize lungs (Figure 5C) or bones (Figure 5D). Likewise, constitutive overexpression of *Snai1* in TSU-Pr1-B2 bladder cancer cells led to a marked suppression of their potential for distant organ colonization (Supplemental Figure 8G).

In reciprocal experiments, the mesenchymal-like PC-3/S tumor cells were manipulated to reduce the levels of EMT factors. Knockdown in these cells of *SNAI1*, *ZEB1*, or *TWIST2* or a triple knockdown (*SNAI1*, *ZEB1*, and *TWIST2*) (Supplemental Figure 11) caused a loss of their fibroblastoid morphology (Figure 6A), an upregulation of E-cadherin (Figure 6, A, B, and E), decreased invasiveness (Figure 6C), and enhanced spheroid formation (Figure 6D), features that were more evident with the triple knockdown. This phenotypic switch was accompanied with the upregulation of genes characteristic of epithelial and self-renewal programs (Figure 6E).

In support of the importance of an epithelial phenotype in the maintenance of properties of self renewal and metastatic potential, knockdown of E-cadherin in PC-3/Mc cells (Figure 7, A, D, and E) caused, in addition to the expected enhanced invasiveness (Figure 7B), a significant reduction in their capacity to form spheroids (Figure 7C) and to colonize lungs in NOD-SCID mice (Figure 7F). Knockdown of E-cadherin in PC-3/Mc

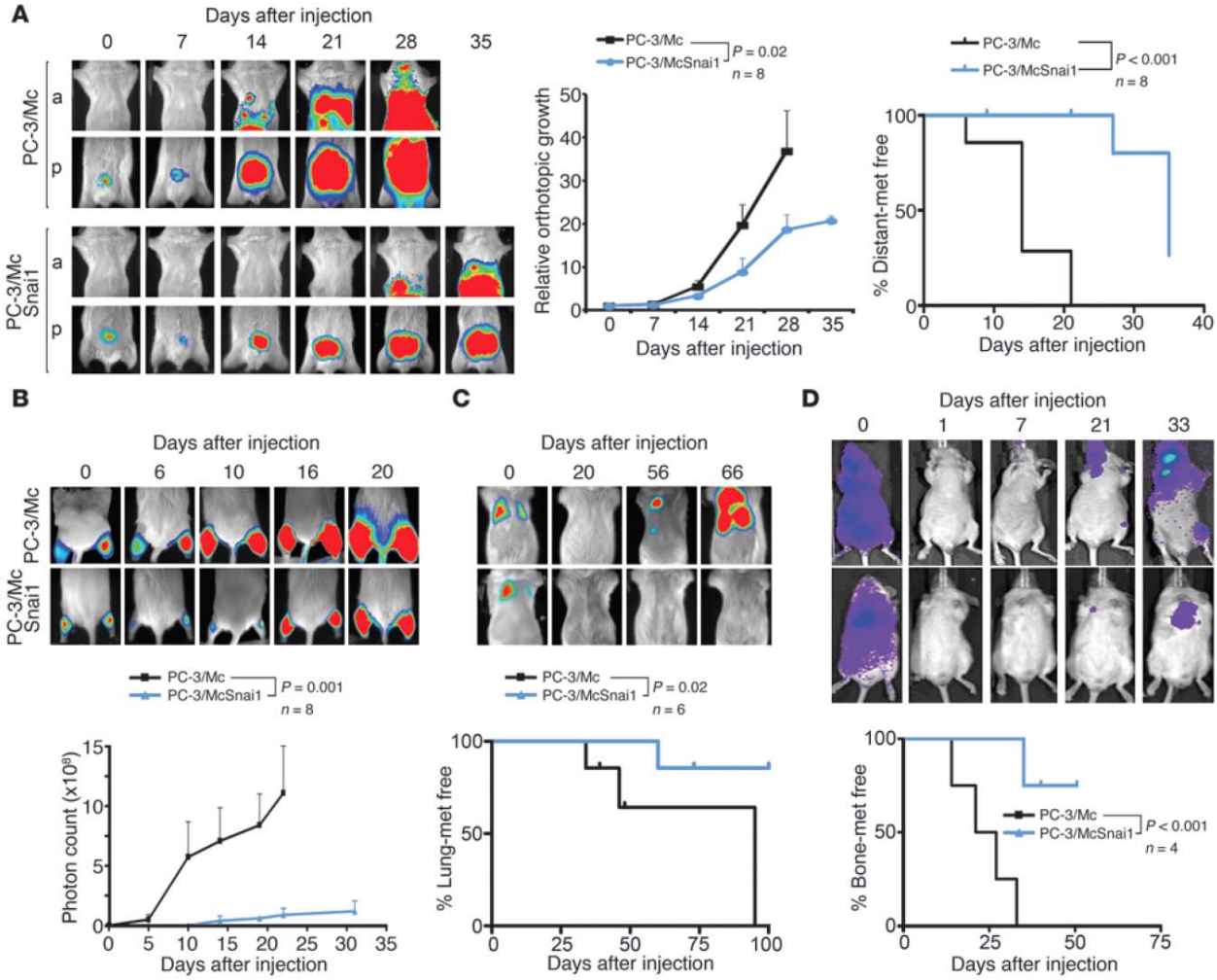


Figure 5

Constitutive overexpression of *Snai1* inhibits local growth, metastatic spread, and distant organ colonization of PC-3/Mc cells. (**A**) Overexpression of *Snai1* strongly inhibited local growth and metastatic spread after orthotopic prostatic implantation of PC-3/Mc-SNAI1 cells (1.0×10^5) in 6-week-old male NOD-SCID mice. Anterior or posterior halves were imaged independently for enhanced resolution. Middle panel: growth curves of orthotopic tumors, with photon counts normalized to values on day 0. Right panel: Kaplan-Meier plots for metastasis-free mice. (**B**) Overexpression of *Snai1* strongly inhibited the growth of PC-3/Mc cells (2.5×10^5) grafted i.m. Mice grafted with control PC-3/Mc cells were euthanized at day 22 after grafting. Bottom panel: growth curve at the i.m. implantation site. (**C**) Overexpression of *Snai1* prevented lung colonization of PC-3/Mc cells (2.5×10^5) inoculated i.v. Bottom panel: Kaplan-Meier plots for lung colonization–free mice. (**D**) Overexpression of *Snai1* suppressed bone colonization of PC-3/Mc cells (2.0×10^5) inoculated i.c. Bottom panel: Kaplan-Meier plots for bone colonization–free mice. Results are expressed as mean ± SEM.

cells was accompanied with a modest but relatively broad downregulation of self-renewal/pluripotency transcription factors, including *SOX2*, *KLF4*, and *MYC*, and the upregulation of several mesenchymal genes, such as *FN1* and *ZEB2* (Figure 7, A and D), suggesting that the expression of E-cadherin may play an active role in the maintenance of a epithelial gene program.

Conversely, overexpression of exogenous E-cadherin in PC-3/S cells, which do not express this epithelial marker under standard growth conditions (Figure 7G), caused a strong inhibition of invasiveness (Figure 7H) and a striking gain in the capacity of cells to form spheroids (Figure 7I). Upon i.m. implantation in immunocompromised mice, and consistent with their in vitro phenotypes, PC-3/S-CDH1 cells grew tumors at significantly faster rates than control cells (Figure 7J). This phenotypic switch was accompanied with a modest upregulation of self-renewal/pluripotency factors and an inhibition of the mesenchymal-like gene profile characteristic of PC-3/S cells (Figure 7K).

The above results suggest a tight association between the expression of an epithelial gene program and the maintenance of a self-renewal gene program and properties of TICs, as well as an inhibition of the latter properties by induction of a mesenchymal gene program. To further explore the relationship between the self-renewal gene network and the growth properties of PC-3/Mc cells, we knocked down *SOX2*, *KLF4*, or *MYC*, or all 3 transcripts in

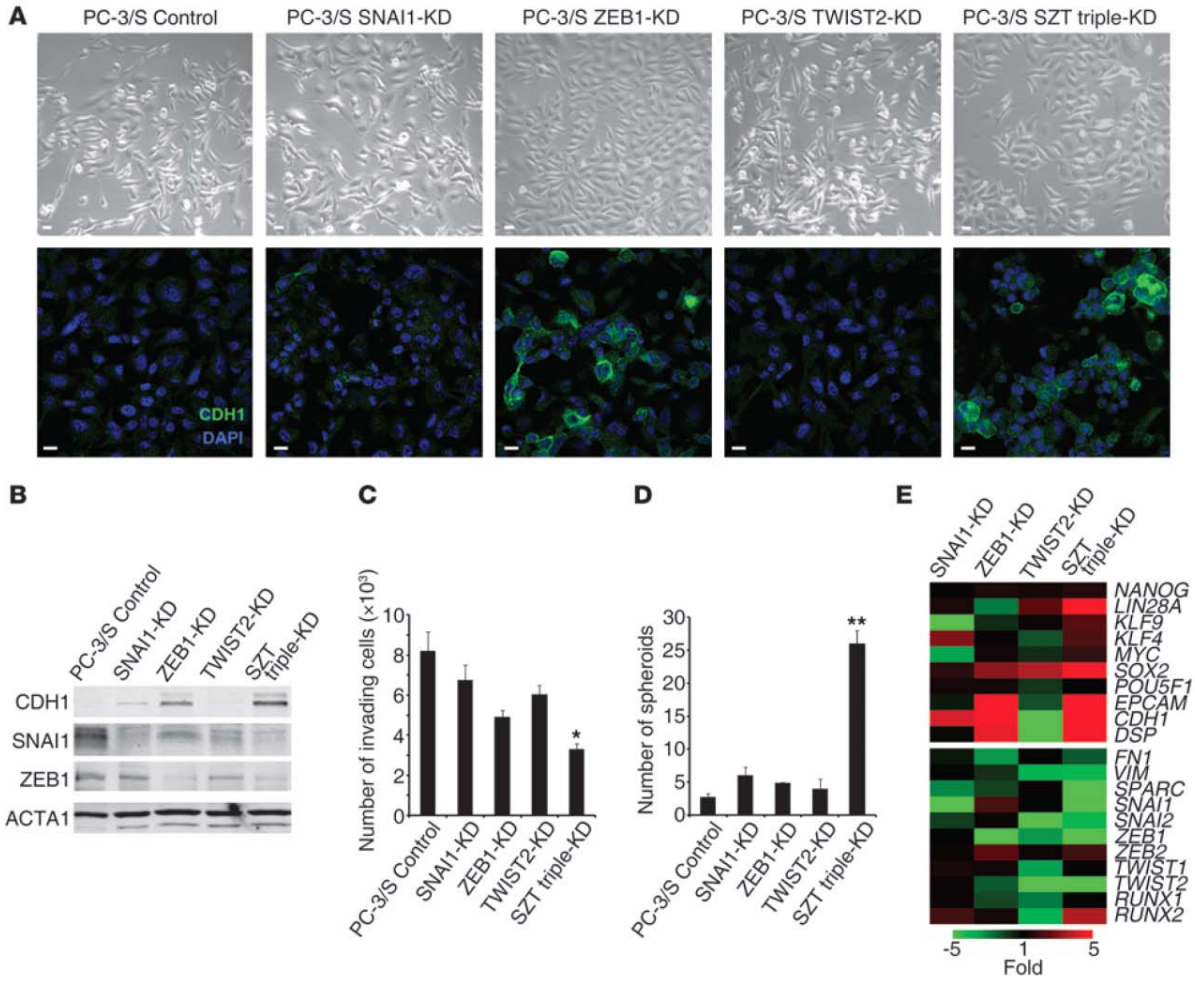


Figure 6

Knockdown of EMT transcription factors in mesenchymal-like PC-3/S cells causes a gain in anchorage-independent growth and the expression of a self-renewal gene network. (**A**). Knockdown of *SNAI1*, *ZEB1*, *TWIST2*, or a triple SZT knockdown in PC-3/S cells was associated with fewer cells with fibroblastoid morphologies and a gain in the expression of E-cadherin, most evident in *ZEB1* knockdowns (single or triple SZT). Scale bars: 20 μm. (**B**) Knockdown in mesenchymal-like PC-3/S cells of *SNAI1*, *ZEB1*, or a triple SZT knockdown caused an upregulation of E-cadherin, as determined by Western blotting, with the strongest effect observed in the triple knockdown. (**C**) Knockdown of *SNAI1*, *ZEB1*, *TWIST2*, or a triple SZT knockdown caused a diminished invasive capacity of PC-3/S cells in Transwell-Matrigel assays, with the triple SZT knockdown showing the strongest effects. (**D**) Knockdown of *SNAI1*, *ZEB1*, *TWIST2*, or a triple SZT knockdown caused a gain in the capacity of PC-3/S cells to grow spheroids, with the triple knockdown showing the strongest effects. (**E**) Knockdown of *SNAI1*, *ZEB1*, *TWIST2*, or a triple SZT knockdown in mesenchymal-like PC-3/S cells caused an upregulation of the epithelial genes *CDH1*, *EPCAM*, and *DSP* and of the self-renewal/pluripotency genes *LIN28*, *SOX2*, *MYC*, and *KLF4*, most evident for the triple SZT knockdown. Real-time RT-PCR values, determined by the ΔΔCp method, are represented as a heat map with pseudocoloring ranging from green (underexpressed relative to values in control PC-3/S cell) to red (overexpressed relative to control PC- 3/S cells). Controls were puromycin-selected PC-3/S cells bearing control pLK0-scrambled lentiviral vector. Results are expressed as mean ± SEM. $^{*}P < 0.05$; $^{**}P < 0.01$.

these cells (Supplemental Figure 12). This caused a downregulation of E-cadherin and other epithelial markers (Figure 8, A–C), a decrease in the formation of spheroids (Figure 8D), and enhanced invasiveness (Figure 8E), changes that were most evident in cells with a triple knockdown for all 3 self-renewal/pluripotency factors. In vivo, knockdown of *SOX2* was sufficient to inhibit the local growth of PC-3/Mc cells (Figure 8F) and their lung colonization (Figure 8G). Knockdown of *SOX2* in TSU-Pr1-B2 bladder cancer cells also resulted in a downregulation of E-cadherin (Supplemental Figure 13, A and B), loss of spheroid-forming potential (Supplemental Figure 13C), gain in invasiveness (Supplemental Figure 13D), and a strong inhibition of distant organ colonization (Supplemental Figure 13E).

In a reciprocal approach, the transduction and overexpression of *SOX2* in PC-3/S cells caused the upregulation of E-cadherin and downregulation of fibronectin (Supplemental Figure 14A) and enhanced the formation of spheroids (Supplemental Figure 14B) concomitant with an inhibition of invasiveness (Supplemental Figure 14C), and a strong enhancement of tumorigenicity upon i.m. implantation (Supplemental Figure 14D).

research article

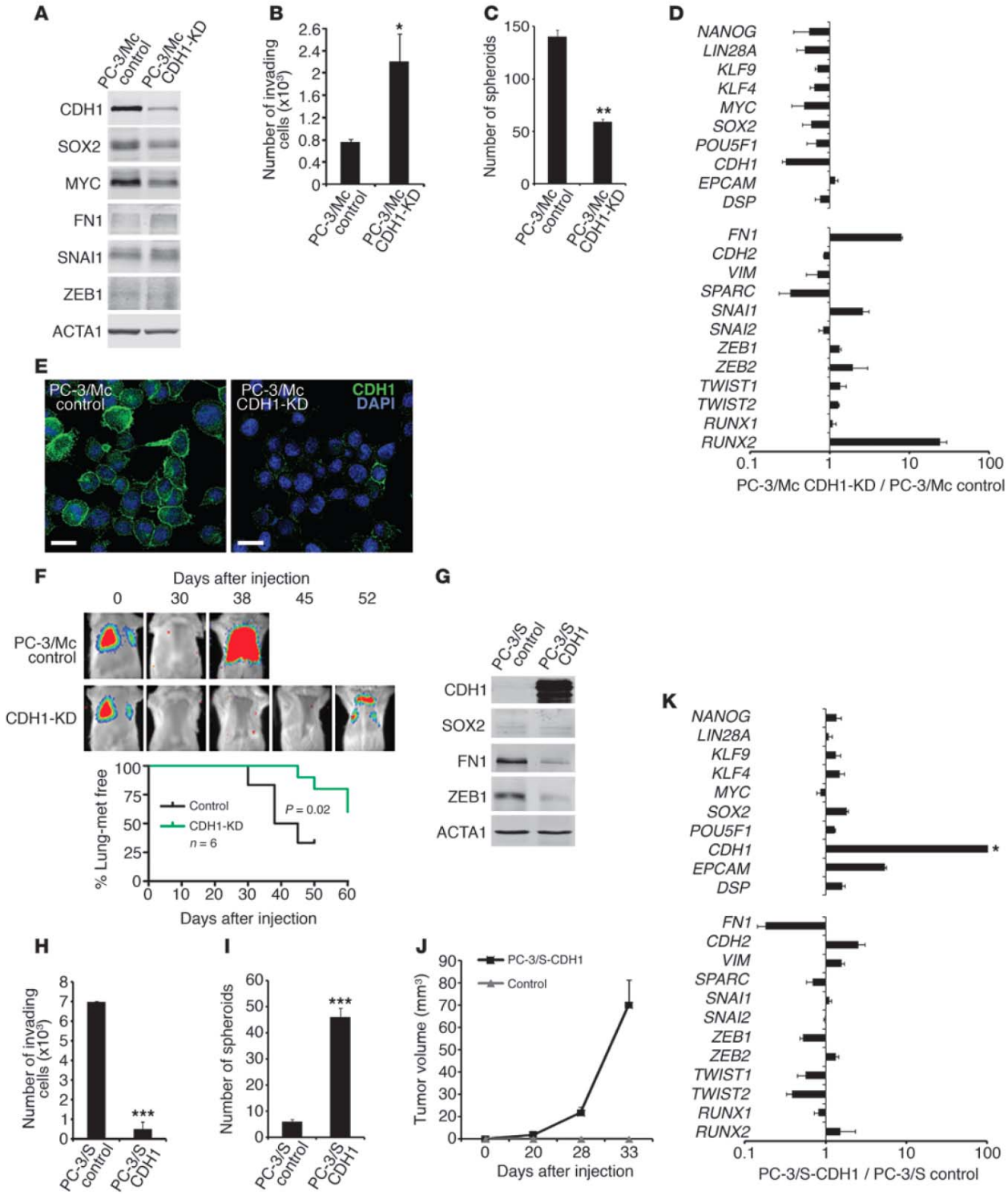


Figure 7
E-cadherin is required for anchorage-independent growth and lung colonization of PC-3/Mc cells. (**A**) Knockdown of E-cadherin in PC-3/Mc cells downregulated *SOX2* and *MYC*. Controls were puromycin-selected PC-3/Mc cells bearing pLK0-scrambled lentiviral vector. (**B**) Knockdown of E-cadherin enhanced the invasiveness of PC-3/Mc cells. (**C**) Knockdown of E-cadherin inhibited the spheroid-forming potential of PC-3/Mc cells. (**D**) Knockdown of E-cadherin in PC-3/Mc cells caused a modest downregulation of self-renewal/pluripotency genes. Relative transcript levels are represented as the $\log_{10}$ of ratios between experimental and control cells of $2^{-\Delta\Delta Cp}$ real-time PCR values. Controls were PC-3/Mc cells bearing pLK0-scrambled vector. (**E**) Knockdown of E-cadherin in PC-3/Mc cells detected by indirect immunofluorescence. Scale bars: 20 μm. (**F**) Knockdown of E-cadherin in PC-3/Mc cells inhibited their lung colonization after i.v. injection into SCID mice. The Kaplan-Meier plot reflects the actuarial numbers of lung colonization–free mice. (**G**) Overexpression of E-cadherin in PC-3/S cells caused a downregulation of *FN1*. (**H**) Overexpression of E-cadherin strongly inhibited the invasiveness of PC-3/S cells. (**I**) Overexpression of E-cadherin strongly enhanced the spheroid-forming potential of PC-3/S cells. (**J**) Overexpression of E-cadherin strongly enhanced the tumorigenicity of PC-3/S cells. PC-3/S-CDH1 and control cells (5×10^5) were implanted i.m. in the hind limbs of male Swiss-nude mice and tumor growth monitored with a caliper. (**K**) Overexpression of E-cadherin induced a moderate upregulation of self-renewal/pluripotency genes and a moderate downregulation of mesenchymal genes. Asterisk in **K** shows E-cadherin levels determined in murine E-cadherin–overexpressing cells reflect the exogenous transcripts, quantified with mouse-specific primers and probes (values are off scale). Results are expressed as mean ± SEM. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

Taken together, these observations reinforce the notion that expression of an epithelial gene program and phenotype is critical for the maintenance of a self-renewal gene program and more aggressive attributes of these tumor cells.

A cell subpopulation enriched in TICs cooperates with a subpopulation with traits of stable EMT for enhanced in vitro invasiveness and in vivo organ colonization. The above results suggest that tumor cells with strong epithelial phenotypes and low autonomous (in vitro) invasive potential display strong metastatic potential. However, in order to develop distant metastases, tumor cells must first breach local barriers that contain them within their primary site. That the highly metastatic PC-3/Mc cells are poorly invasive in vitro may contradict this principle unless they become invasive under certain conditions. Indeed, shortly after i.m. grafting in immunodeficient mice, PC-3/Mc cells downregulated E-cadherin and upregulated fibronectin (Figure 9A), suggesting that murine factors may induce EMT in these cells in vivo. On the other hand, tumors and lung colonies formed by PC-3/Mc cells in NOD-SCID mice coexpressed SOX2 and E-cadherin (Figure 9B), suggesting that PC-3/Mc cells, when implanted alone in vivo, may escape their primary implantation sites aided by EMT induced by murine factors and that, after leaving their primary site, they may revert to an epithelial phenotype in order to grow distant metastases.

In addition, the mirror-image phenotypes of subpopulations with either epithelial or mesenchymal phenotypes in our prostate and bladder cancer models raise the question of whether diverse populations isolated from a common parental tumor cell line might interact with each other in order to compound a collective behavior that has an impact on the tumor's potential for local invasiveness or establishment of distant metastases (2, 6, 7, 26). Upon coculture with PC-3/S cells, PC-3/Mc cells became invasive (Figure 10A), suggesting a cooperation between these 2 subpopulations in order to facilitate the local invasiveness of the more epithelial tumor cell subpopulations, which display a poor autonomous invasive potential in vitro. In addition, we found that PC-3/Mc cells that had been cocultured with PC-3/S cells were still invasive after separation from PC-3/S cells by FACS (Supplemental Figure 15), a phenotypic change that was maintained for at least 7 days after coculture (Figure 10B) and was reversible, following a time-dependent decline after separation from PC-3/S cells (Figure 10, B and F). Coculture of PC-3/Mc cells with NIH3T3 fibroblasts also stimulated their invasiveness (Supplemental Figure 16), suggesting that the invasiveness of epithelial PC-3/Mc can be enhanced by exposure to both tumoral and nontumoral mesenchymal cell types. We further found that conditioned medium (CM) from PC-3/S cells markedly induced the invasiveness of PC-3/Mc cells (Figure 10C), suggesting that diffusible factors play a significant role in the stimulation of invasiveness of PC-3/Mc cells induced by PC-3/S cells. Moreover, PC-3/Mc cells that had been cocultured with PC-3/S cells not only gained in invasive potential, but were inhibited in their anchorage-independent growth potential (Figure 10D), reminiscent of the inhibition of self renewal and anchorage-independent growth by EMT observed in the experiments described above.

The observed phenotypic switch was accompanied with a downregulation of the epithelial genes *CDH1*, *EPCAM*, and *DSP* and the self-renewal/pluripotency genes *SOX2*, *OCT4*, *KLF4*, and *LIN28A* along with upregulation of the mesenchymal genes *FN1*, *SPARC*, *TWIST2*, and *RUNX2*, which also declined with time after coculture (Figure 10, E and F). Changes in histone marks at relevant promoters support an epigenetic basis for this gene program switch (Supplemental Figure 17) and may help explain the persistence of the invasive program of PC-3/Mc cells 7 or more days after their coculture and separation from PC-3/S cells (Figure 10, B and F). These observations suggest that the escape of epithelial/TIC subpopulations from local environments may follow not only passive mechanisms through the action of stromal components and mesenchymal tumor cells (25), but also an active mechanism through transient EMT of epithelial tumor subpopulations induced by mesenchymal tumor subpopulations.

In vivo, PC-3/Mc cells coinjected with PC-3/S cells grew at significantly slower rates than PC-3/Mc cells alone upon orthotopic (Figure 11A) or i.m. (Figure 11B) implantation, but metastasized to regional lymph nodes after orthotopic implantation at earlier times than control PC-3/Mc cells alone (Figure 11A). In these experiments, *Renilla* luciferase–expressing PC-3/S cells were transiently detected outside of the site of orthotopic implantation at early times after orthotopic implantation (Figure 11A), but ceased to be detected after more prolonged monitoring both at the site of implantation or at distant sites (Figure 11A), consistent with the highly invasive but poorly metastatic properties of these cells. These results are also consistent with the in vitro observations, described above, of reduced growth of PC-3/Mc cells when cocultured with PC-3/S cells, along with an increased capacity to escape from implantation sites, followed by a subsequent restoration of their growth potential at distant sites where PC-3/S cells are no longer present. Upon i.v. injection, PC-3/Mc cells coinjected with PC-3/S cells colonized lungs at earlier times than PC-3/Mc cells alone (Figure 11C), suggesting that interaction with PC-3/S cells also rendered them more effi-

research article

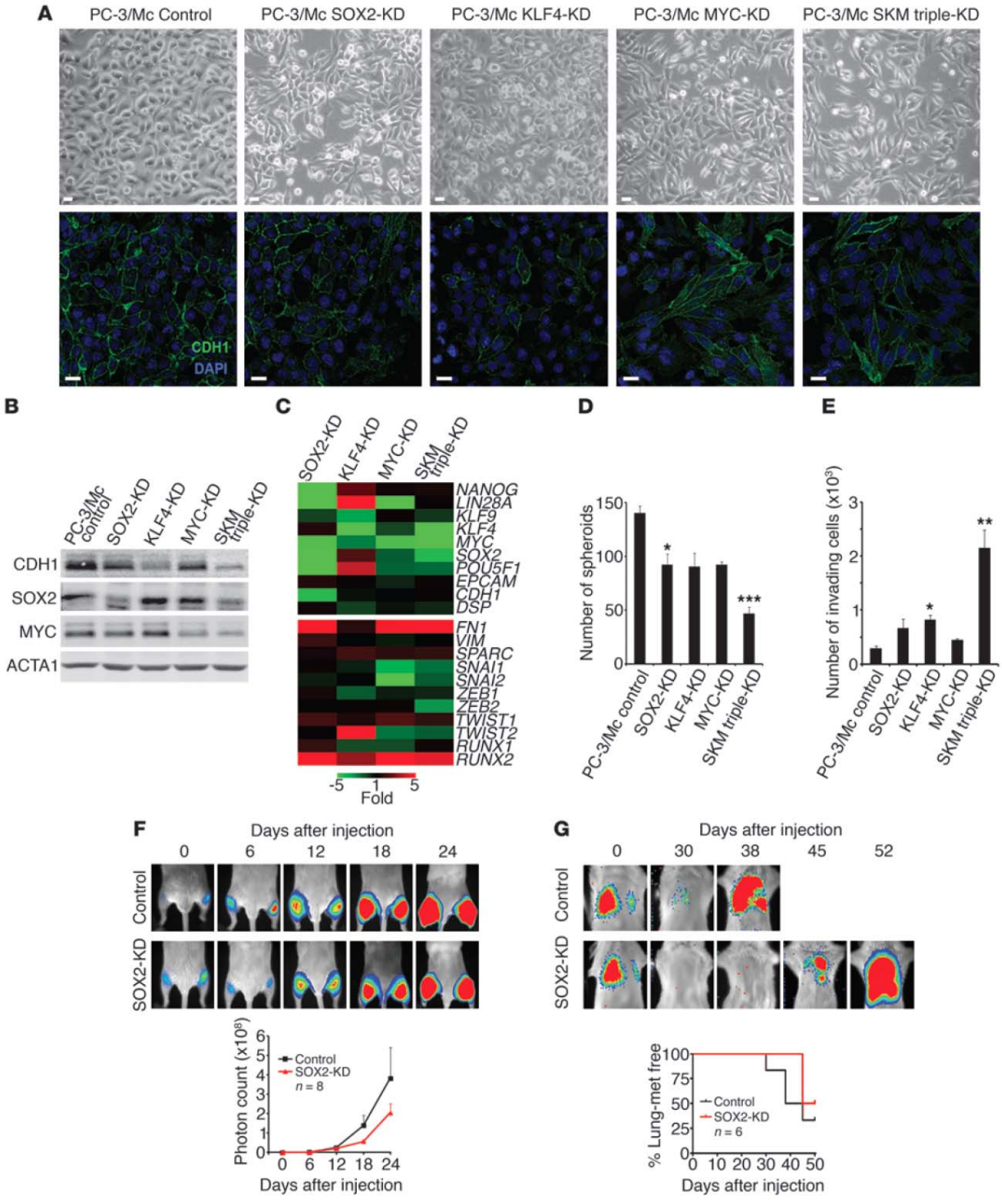


Figure 8
Self-renewal factors are required for a strong epithelial program, anchorage-independent growth, and lung colonization of PC-3/Mc cells. (**A**) Knockdown in PC-3/Mc cells of *SOX2*, *KLF4*, *MYC*, or a triple SKM knockdown induced a fibroblastoid morphology and downregulation of membrane-associated E-cadherin. Controls were puromycin-selected PC-3/Mc cells bearing pLK0-scrambled control vector. Scale bars: 20 μm. (**B**) Knockdown in PC-3/Mc cells of *SOX2*, *KLF4*, *MYC*, or a triple SKM knockdown caused a downregulation of E-cadherin, strongest for the triple knockdown and *KLF4*. (**C**) Knockdown in PC-3/Mc cells of *SOX2*, *KLF4*, *MYC*, or a triple SKM knockdown caused a downregulation of *CDH1* and an upregulation of *FN1* and *SPARC*. Real-time RT-PCR values, determined by the ΔΔCp method, are represented as a heat map (green, underexpressed relative to control PC-3/Mc cells; red, overexpressed). (**D**) Knockdown of *SOX2*, *KLF4*, *MYC*, or a triple SKM knockdown caused an inhibition of the capacity of PC-3/Mc cells to grow spheroids under anchorage-independent conditions, strongest for the triple knockdown. (**E**) Knockdown of *SOX2*, *KLF4*, *MYC*, or a triple SKM knockdown caused an enhanced invasiveness of PC-3/Mc cells, strongest for the triple knockdown. (**F**) Knockdown of *SOX2* was sufficient to inhibit the tumorigenic potential of PC-3/Mc cells. Cells (2.0×10^5) were implanted i.m. in male SCID mice. Bottom panel: graphical representation of photon counts at the indicated times. (**G**) Knockdown of *SOX2* was sufficient to inhibit lung colonization by PC-3/Mc cells. Cells (2.5×10^5) were inoculated i.v. in male SCID mice. Bottom: Kaplan-Meier actuarial plot for lung colonization–free mice. Results are expressed as mean ± SEM. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

cient at extravasation. Indeed, upon i.c. inoculation, PC-3/Mc cells coinjected with PC-3/S cells were not more efficient than PC-3/Mc cells injected alone at colonizing bones (Figure 11D), where the permeable sinusoidal capillary system of bone marrow represents a much weaker barrier to extravasation than the capillaries in the lungs or other organs (44). Additionally, PC-3/Mc cells coinjected i.c. with PC-3/S cells colonized adrenal glands (Supplemental Table 7), which they never colonized when injected alone. Coinjection of PC-3/Mc cells expressing green fluorescent protein together with PC-3/S cells expressing red fluorescent protein (RFP) and *Renilla* luciferase allowed us to identify the cells of origin of the tumors that developed in distant organs. Such tumors contained only green fluorescence but not red fluorescence — or *Renilla* luciferase–expressing cells, indicating that only PC-3/Mc cells contributed to distant organ colonization (Figure 11A and Figure 12).

These results suggest that, in vivo, while the mesenchymal-like PC-3/S tumor cells can escape local tumor sites but lack metastatic potential, their presence facilitates the escape of the more epithelial PC-3/Mc cells from local tumor sites in order to establish distant metastases.

Expression of a self-renewal gene network active in PC-3/Mc cells is associated with more advanced stages of prostate cancer. The above observations suggest that more aggressive tumors contain larger representations of epithelial tumor cells with high self-renewal potential. To determine whether the self-renewal gene network active in the more epithelial PC-3/Mc subpopulation is associated with aggressive prostate cancers, we extracted a subset of the ESC gene set (12) that most significantly discriminated PC-3/Mc from PC-3/S cells (designated M gene set; Supplemental Table 8) and interrogated it for its enrichment in an expression data set for 150 samples from prostate cancer patients. (45). This analysis showed that the M gene set is indeed significantly enriched in metastatic relative to primary prostate cancer samples and also in primary tumor samples from more advanced stages (T3 and T4 vs. T1 and T2) (Figure 13, A and B).

To determine whether these observations could be applied in in situ histopathological analyses, we studied the expression of SOX2 as a potential indicator of self-renewing populations by immunohistochemistry on samples from primary and metastatic prostate cancer (Supplemental Table 9). The results revealed a significantly higher frequency of SOX2-positive samples in stage T3 than in stage T2 tumors (Figure 13C). Strikingly, in several SOX2-positive metastases, all tumor cells expressed SOX2 with a strong nuclear staining. Such strong and homogeneous expression of SOX2 was not observed in any of the primary tumors studied. In addition, those metastatic samples with the strongest expression of SOX2 showed the strongest staining for E-cadherin (Figure 13D). This suggests that some prostate cancer metastases are enriched in tumor cells with active self-renewal programs expressing high levels of SOX2 and E-cadherin, consistent with the finding by Tsuji et al. that cells that colonize to organs are non-EMT cells (25) and also with studies describing a stronger expression of E-cadherin in metastatic tumors (46, 47).

Discussion

A major focus of study of the metastasis problem is understanding the mechanisms by which tumor cells escape the local environment and colonize distant organs (19, 20, 48). It has been proposed that the engagement of an EMT program simultaneously leads, through mechanisms not yet elucidated, to the acquisition of a self-renewal program (24, 48), endowing tumor cells not only with the capacity to invade and migrate through tissues, but also to survive in the circulation and form colonies in distant organs. The latter hypothesis is largely based on experiments in which the expression of transcription factors that direct the expression of EMT programs is manipulated for overexpression or silencing in relatively heterogeneous populations of tumor cells and on the study of their capacity to form tumors, invade local tissues, and establish metastases in immunocompromised mice.

Here, we have studied clonal populations derived from the PC-3 prostate cancer and the TSU-Pr1 bladder cancer cell lines displaying relatively stable and contrasting phenotypes, namely cells with a strong epithelial phenotype (PC-3/Mc and TSU-Pr1-B2) and a more mesenchymal phenotype (PC-3/S and TSU-Pr1), as determined by the expression of genes characteristic of either program. Our analysis shows that the subpopulations with the stronger epithelial phenotypes display clearly enhanced capacities to form spheroids in culture and to colonize lungs and bone, compared with the tumor subpopulations with stable mesenchymal-like phenotypes, despite the fact that the latter are more invasive through extracellular matrix in in vitro assays.

In order to further explore the hypothesis that a strong self-renewal and metastatic phenotype requires the maintenance of an epithelial program in our cell models, we have (a) induced constitutive EMT in epithelial tumor subpopulations through the transduction and overexpression of EMT-directing transcription factors, (b) knocked down the same factors in the mesenchymal-like PC-3/S cell subpopulation, (c) knocked down E-cadherin in the epithelial tumor subpopulations, (d) transduced and overexpressed E-cadherin in PC-3/S cells, (e) knocked down self-renewal/pluripotency factors in the strongly epithelial PC-3/Mc subpopulation, and (f) transduced and overexpressed the self-

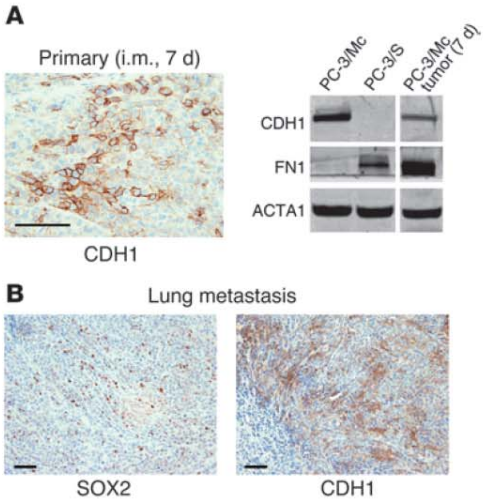


Figure 9
Downregulation of E-cadherin from PC-3/Mc cells at primary implantation sites and maintenance of its expression in lung metastasis. (**A**) Downregulation of E-cadherin and upregulation of fibronectin in PC-3/Mc cells after implantation in NOD-SCID mice. Seven days after i.m. implantation, PC-3/Mc cells, homogeneously positive for E-cadherin and negative for fibronectin in culture prior to implantation, become heterogeneous for expression of membrane-associated E-cadherin, as determined by immunohistochemistry (left panel), and downregulate E-cadherin and upregulate fibronectin, as determined by Western blotting (right panel). Lanes separated by the white lines were run on the same gel but were noncontiguous. Scale bar: 100 μm. (**B**) PC-3/Mc cells that had metastasized to lungs after i.v. injection were largely positive for nuclear SOX2 and membrane-associated E-cadherin, as determined by immunohistochemistry. Scale bars: 100 μm.

renewal factor *SOX2* and the epithelial gene E-cadherin in the mesenchymal-like PC-3/S tumor cells. The results from all these complementary approaches lead to the same conclusions, namely that the suppression of an epithelial program (through constitutive expression of EMT transcription factors or knockdown of E-cadherin) inhibits the self-renewal/pluripotency gene network of tumor cells, their capacity to grow under attachment-independent conditions, and their tumorigenic and metastatic potentials. The association between properties attributed to TICs and an epithelial phenotype is further supported by the fact that knockdown of 3 of the 4 canonical Yamanaka pluripotency transcription factors (*SOX2*, *KLF4*, and *MYC*) in PC-3/Mc cells reduced their epithelial phenotype and TIC attributes and induced an invasive and more mesenchymal phenotype, while the overexpression of *SOX2* in PC-3/S cells was sufficient to enhance the expression of epithelial markers and properties of TICs in these cells, including enhanced tumorigenicity, while inhibiting invasiveness and the expression of mesenchymal markers.

These results suggest that the self-renewal properties of these tumor cells depend on the same factors that endow normal cells with self renewal and pluripotency (31, 32), that this gene network sustains the expression of an epithelial gene program and, at the same time, opposes the expression of a mesenchymal gene program and the acquisition of a motile and invasive phenotype. Reciprocally, the induction of a mesenchymal gene program in our cells opposes not only their epithelial gene program, but also their self-renewal gene network and associated properties. This situation is reminiscent of the requirement for normal adult fibroblasts

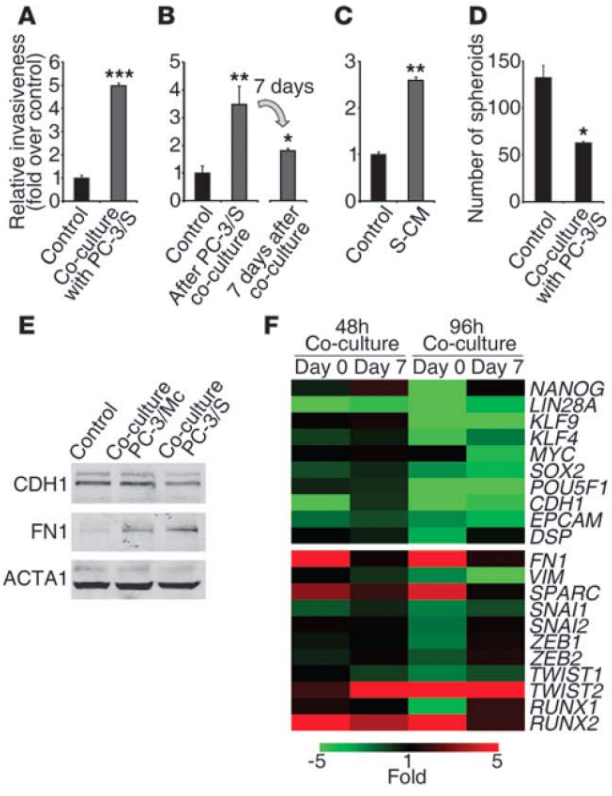


Figure 10
PC-3/S cells enhance the invasiveness of PC-3/Mc cells. (**A**) Coculture with PC-3/S cells induced the invasiveness of PC-3/Mc cells. Green-labeled PC-3/Mc cells were cocultured with red-labeled PC-3/S cells on Transwell units and green or red fluorescent invading cells scored by flow cytometry. Controls were green-labeled PC-3/Mc cells cocultured with unlabeled PC-3/Mc cells. (**B**) The enhanced invasiveness of PC-3/Mc cells was maintained for several days after coculture with PC-3/S cells. GFP-labeled PC-3/Mc cells were cocultured for 48 hours with red-labeled PC-3/S cells, sorted, and assayed for invasiveness either immediately or 7 days later. (**C**) Diffusible factors secreted by PC-3/S cells enhanced the invasiveness of PC-3/Mc cells. PC-3/Mc cells were exposed for 48 hours to CM from PC-3/S cells (S-CM) and assayed for invasiveness. (**D**) Coculture with PC-3/S cells inhibited the spheroid growth of PC-3/Mc cells. GFP-expressing PC-3/Mc cells and RFP-expressing PC-3/S cells were cocultured and scored for spheroids after 14 days. (**E**) Coculture of PC-3/Mc cells with PC-3/S cells induced a downregulation of E-cadherin and an upregulation of fibronectin. Green-labeled PC-3/Mc cells and red-labeled PC-3/S cells were cocultured for 48 hours, sorted, and analyzed by Western blotting. (**F**) PC-3/Mc cells cocultured with PC-3/S cells shifted their transcriptional programs following a time-dependent reversion after coculture. Green-labeled PC-3/Mc cells were cocultured with red-labeled PC-3/S cells for 48 or 96 hours, sorted, and analyzed either immediately (day 0) or 7 days after sorting (day 7). Relative qPCR transcript levels are represented as a heat map (green, underexpressed relative to control PC-3/Mc; red, overexpressed). Results are expressed as mean ± SEM. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

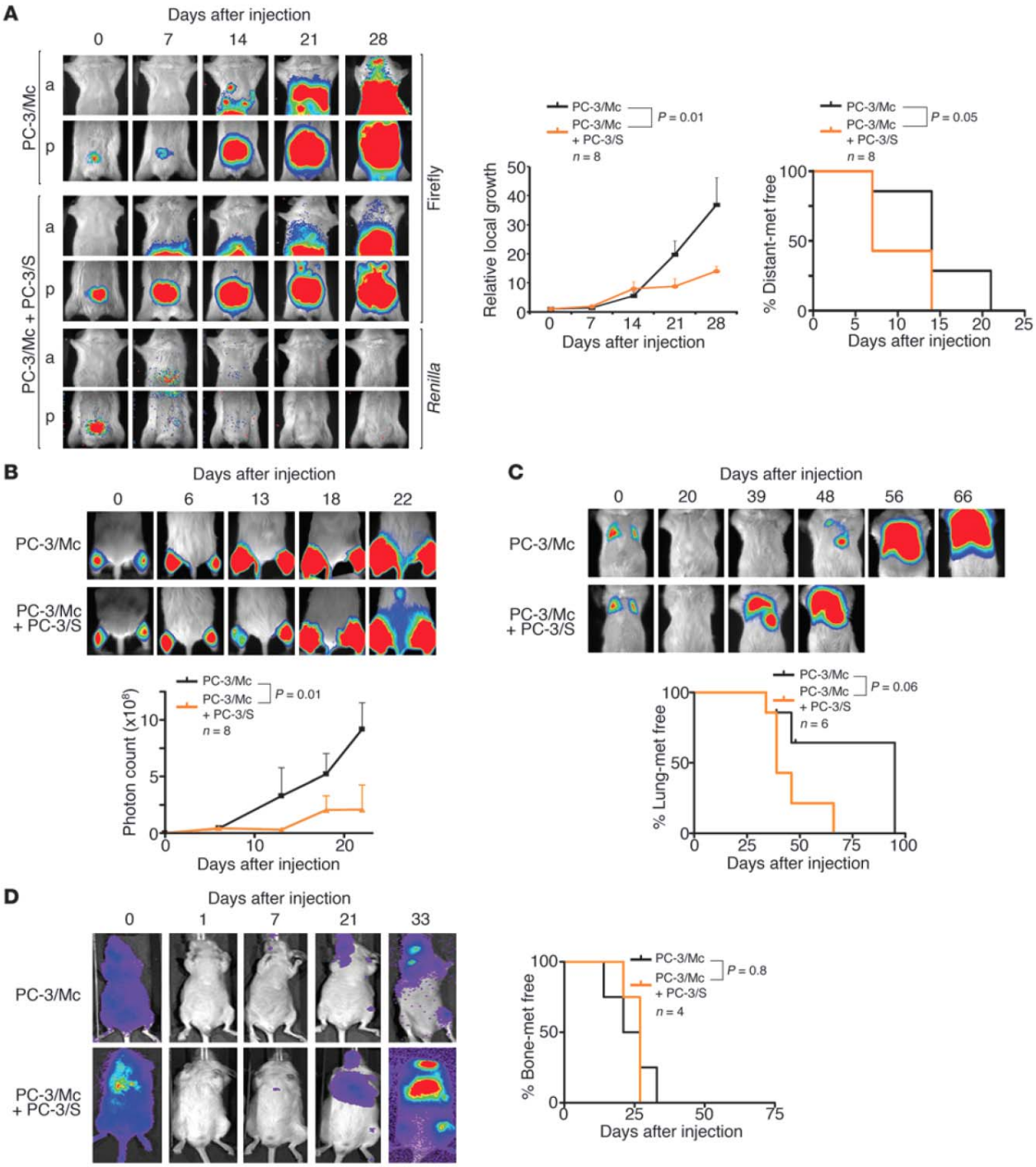


Figure 11

PC-3/S cells facilitate the spread and metastatic growth of PC-3/Mc cells. (**A**) Orthotopic coimplantation of GFP-PC-3/Mc cells with RFP- and *Renilla* luciferase–expressing PC-3/S cells in the ventral prostate of NOD-SCID mice diminished their growth rate at the implantation, while accelerating the appearance of metastatic growth. Bioluminescence monitoring was performed separately for the anterior and posterior halves of the mice, for improved resolution. Middle: growth curves of orthotopically implanted tumor cells, with photon counts normalized relative to values on day 0. Right: Kaplan-Meier actuarial plots for metastasis-free mice. (**B**) Coimplantation (i.m.) of GFP-PC-3/Mc cells with RFP-PC-3/S cells diminished their growth rate as compared with GFP-PC-3/Mc cells implanted alone. Bottom: graphical representation of growth at the implantation site. (**C**) Coinoculation (i.v.) of GFP-PC-3/Mc cells with RFP-PC-3/S cells accelerated their lung colonization. Bottom: Kaplan-Meier actuarial plots for lung colony–free mice. (**D**) Coinoculation (i.c.) of GFP-PC-3/Mc cells with RFP-PC-3/S cells did not significantly affect their bone colonization efficiency. Right panel: Kaplan-Meier actuarial plots for bone colony–free mice. Results are expressed as mean ± SEM.

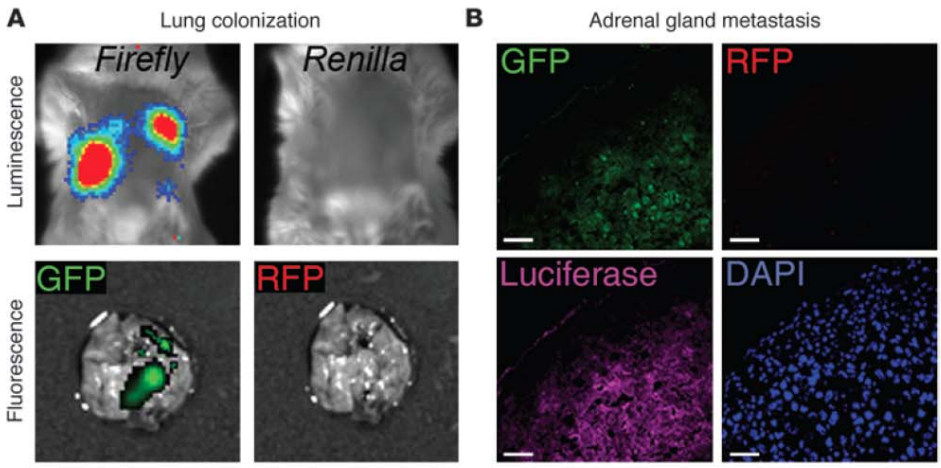


Figure 12
Metastases formed after joint injection of PC-3/Mc and PC-3/S cells contain exclusively epithelial PC-3/Mc cells. (**A**) Only PC-3/Mc cells, but not PC-3/S cells, colonized lungs after joint i.v. injection. PC-3/S, but not PC-3/Mc, cells also expressed *Renilla* luciferase. Firefly luciferase, but not *Renilla*, signal was detected in lung tumors. In parallel, GFP (expressed by GFP-PC-3/Mc cells) or RFP (expressed by RFP-PC-3/S cells) was visualized microscopically. Only GFP signal, but not RFP signal, was detected in lung tumors. (**B**) Only PC-3/Mc cells, but not PC-3/S cells, colonize adrenal glands after i.c. joint inoculation. Thirty-three days after inoculation, mice were sacrificed and adrenal metastatic tumors frozen and processed for fluorescent visualization of GFP or RFP and for immunofluorescent detection of firefly luciferase (as a marker common to both cell types). Samples were counterstained for nuclei with DAPI. Only GFP signal, but not RFP signal, was detected in adrenal metastases. Scale bars: 50 μm.

to undergo a mesenchymal-epithelial transition (MET) for their reprogramming into self-renewing pluripotent cells (49–51). In fact, it has been shown that the expression of E-cadherin by itself can facilitate the reprogramming of adult fibroblasts and the acquisition of pluripotency (52, 53). Our results suggest that the association between self renewal and epithelial gene programs also holds true for at least the 2 cellular models studied here, derived from prostate and bladder cancers.

The contraposition between a gene program that drives self renewal and an invasive program also has precedents in other biological settings. During normal vertebrate development, neural crest progenitor cell migration and specification require the activation, among other factors, of Snail/Slug and concomitant suppression of Sox2, events that are induced both by diffusible factors and cell-cell interactions (54, 55). On the other hand, epithelial reprogramming is required for the induction or maintenance of pluripotent states, which are facilitated by inhibition of EMT (49, 51), while induction of EMT by SNAI1 can be antiproliferative (56, 57). This evidence and our own observations suggest that mutual exclusion between progenitor/stem cell (or CSC in tumors) and EMT programs may be the prevalent mode in the differentiation of normal progenitor cells in some tissues and also in the evolution of some epithelial tumors.

The suppression of major attributes of TICs by constitutive EMT found in our study may seem to contradict other models, in which EMT induced by a number of factors potentiates self renewal together with enhanced tumorigenic and metastatic capacities (22, 24). However, our observations are supported by models in which local invasiveness is inversely correlated with metastatic or organ colonization potential (25, 28). Additional studies, including some in which EMT is proposed to enhance tumorigenic and metastatic potentials, have also shown that tumor cell subpopulations with clear epithelial phenotypes are endowed with the strongest metastatic potential (23, 28, 58, 59). Our observations led us to propose that tumor cells that depend on a self-renewal/pluripotency gene network for their aggressive properties may be susceptible to an inhibition of those properties by EMT, perhaps through direct or indirect downregulation of the self-renewal/pluripotency gene network by factors such as SNAI1. We speculate that tumor cell types that do not depend on the self-renewal/pluripotency gene network for their tumorigenic and metastatic potentials may use EMT factors to induce a self-renewal state, as described for several models (23, 60), through a variety of mechanisms. Tissue-specific regulatory networks could also account for different phenotypic consequences of EMT in different cell types. For example, it has been reported that EMT of primary prostate epithelial cells is not accompanied with enhanced anchorage-independent growth, (61), and as we have found in this study, EMT can suppress the self-renewal states of prostate and bladder cancer cells, while the induction of EMT in noncancerous MCF10 breast epithelial cells enhances their potential to form mammospheres (22, 24).

A second major finding of our study is that, during cooperation between epithelial and mesenchymal-like tumor subpopulations, the former transiently undergo an EMT. The cooperation between epithelial and mesenchymal-like tumor cell subpopulations for enhanced local invasiveness has been described in other models (25, 26). Our findings add a further level of regulatory complexity in these cell-cell interactions and suggest that the escape of epithelial TICs from their primary sites is facilitated by both passive and active mechanisms. Passive escape mechanisms include the breakdown of extracellular matrix and other tissue structures by tumor-associated stromal cells (8, 19, 48) and by tumor cells that have acquired relatively stable mesenchymal-like gene programs (25, 26). We consider here as an active mechanism the reversible acquisition of a mesenchymal-like invasive state by epithelial tumor cells with self-renewal potential, as observed in this study. In this scheme, it is important to distinguish between tumor cell subpopulations that have acquired relatively stable mesenchymal-like phenotypes and those subpopulations with strong epithelial phenotypes that can undergo transient EMT. In our prostate cancer model, the former subpopulation is unable to establish distant metastases, while the latter cells maintain their capacity to metastasize, suggesting that they undergo a reversion of the EMT, or MET, after they escape from their primary sites. The latter hypothesis is supported by our observations that many

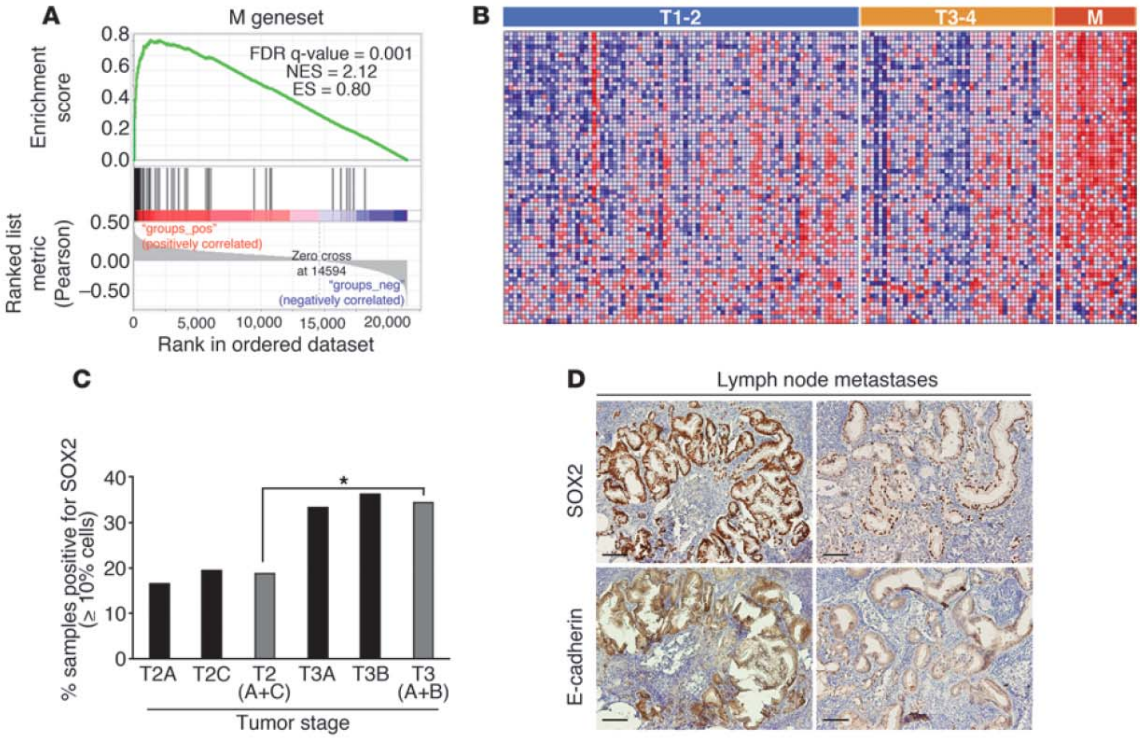


Figure 13

Expression of a self-renewal gene network active in PC-3/Mc cells is associated with more advanced stages of prostate cancer. (**A**) GSEA on an expression data set for 150 prostate cancer samples (45) showing a significant enrichment of the M geneset (genes of the ESC module [ref. 13] enriched in PC-3/Mc cells) in metastases relative to primary tumors, and in T3 and T4 stage primary tumors relative to T1 and T2 stage primary tumors. Pearson's correlation was applied to determine linear relationships between gene profiles and 3 phenotypes (class 1: metastatic; class 2: T3 and T4 stage primary; class 3: T1 and T2 stage primary) taken as continuous variables. (**B**). Heat map illustrating the relative expression levels for the 70 genes of the M gene set. Samples are ordered as primary tumors with stages T1 or T2, stages T3 or T4, or metastases (M). (**C**) Ninety-four cases of prostate cancer were analyzed for SOX2 expression by immunohistochemistry. Positive cases contained at least 10% of cells with nuclear SOX2 staining. $^*P < 0.05$, between the frequencies of SOX2-positive cases in stages T2A and T2C versus and stage T3A and T3B tumors. (**D**) In some lymph node metastases, but in none of the 94 primary tumors, all visible tumor cells were strongly positive for nuclear SOX2, and stronger SOX2 expression correlated with stronger E-cadherin expression. Right: a second metastatic sample with a more heterogeneous and weaker nuclear SOX2 staining of tumor cells displays weaker membrane E-cadherin staining. Scale bars: 100 μm.

experimental metastases express E-cadherin and that metastases from prostate cancer patients also frequently display a strong expression of E-cadherin.

Collectively, our observations suggest that, in some cancers, the acquisition of mesenchymal traits by tumor cells that leads to their loss of epithelial properties occurs at the expense of their self-renewal potential. When the induction of EMT is constitutive, as by forced overexpression of *Snai1*, the losses in self renewal and tumorigenic and metastatic potentials are also sustained. On the other hand, transient EMT, such as that induced by the cooperation between epithelial and mesenchymal tumor subpopulations described in this study, could enhance the local invasiveness of the epithelial subpopulation, thus contributing to the overall metastatic potential of a tumor in which heterogeneous epithelial and mesenchymal subpopulations coexist. This model is schematically summarized in Figure 14 and proposes that the more epithelial/self-renewal tumor populations that leave their primary site either passively, aided by stromal or mesenchymal-like tumor cells, or actively, through their own transient EMT, can form metastases because they have maintained their epithelial phenotypes or, if they have undergone EMT at the primary site, revert to an epithelial/self-renewal program at distant sites. We further propose that, once at their metastatic sites, epithelial TICs may follow a cycle similar to that occurring at the primary sites, with induction of EMT at varying degrees, depending on the environment of the metastatic sites. Formal proof of this model, in particular the demonstration of transient local EMT of epithelial TICs followed by MET at distant sites, would require additional experimental confirmation. Our model is compatible with a more direct participation of stromal or other nontumoral cells in promoting the local invasiveness of tumor cells (8).

Methods

Additional information appears in Supplemental Methods.

Cell lines and reagents. PC-3/Mc and PC-3/S were clonally derived from the human cell line PC-3, isolated from the bone metastasis of a prostate adenocarcinoma (62). Both sublines carry the integrated firefly luciferase gene coding region cloned in the Superluc pRC/CMV vector (Invitrogen). The PC-3/Mc clone was selected by limiting dilution from PC-3/M, isolated from liver metastases produced in nude mice subsequent to intrasplenic

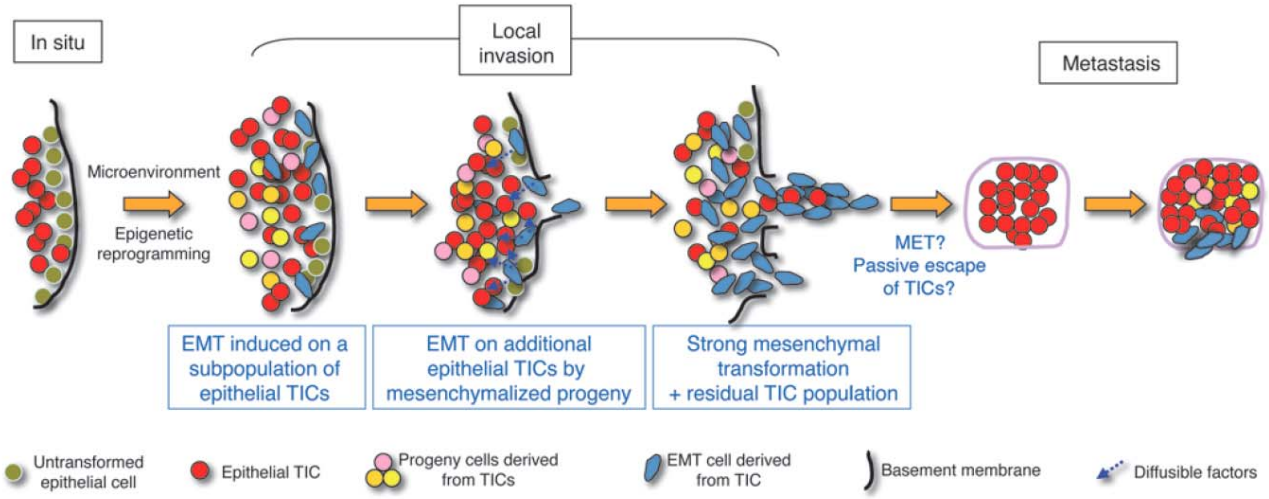


Figure 14

A model of metastasis potentiated by cooperation between tumor cell populations expressing either epithelial/TIC or mesenchymal programs. We propose a model in which some TICs, with properties of CSCs, undergo EMT under the influence of environmental factors. This results in epigenetic reprogramming, including a repression in those cells of pluripotency programs that sustain cell self renewal. These "mesenchymalized" cells, in turn, either through direct cell-cell interaction or through diffusible factors, drive the mesenchymal conversion of additional populations of TIC/CSCs that have not yet undergone EMT, resulting in a reinforcement of the mesenchymalization of the tumor. The predominantly mesenchymalized populations of tumor cells complete the breach of local barriers and thus the tumor becomes fully invasive. The tumor cells escaping from the local site would thus be a combination of stably mesenchymalized tumor cells, cells retaining TIC/CSC properties that leave the tumor following paths open by actively invading cells (passive escape), or TIC/CSCs that have undergone transient EMT (active escape). After hematogenous or lymphogenous spread, TIC/CSCs that have not undergone EMT or that have reverted to an epithelial program and phenotype from their transient EMT (MET) can establish distant metastases. This cycle may be repeated at the metastatic site. Tumor cells with stable mesenchymal-like phenotypes that have escaped from the local tumor site but that do not revert to an epithelial gene program and phenotype would not have the capacity to establish distant metastases.

injection of PC-3 cells (30). PC-3/S cells were selected by limiting dilution from parental PC-3 cells. Cells were grown at 37°C in a 5% CO_2 atmosphere in complete RMPI 1640 supplemented with 200 μg/ml Geneticin (Sigma-Aldrich) to maintain the chromosomal integration of the luciferase gene. TSU-Pr1 and B2 cells (36, 37) were maintained at 37°C in a 5% CO_2 atmosphere in complete DMEM. All media were supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 100 μg/ml streptomycin, and 10% FBS. Unless otherwise indicated, media and sera were from PAA.

Spheroid formation assay. Cells (10^3/well) were seeded on 24-well Ultra Low Attachment culture plates (Corning) in complete culture medium containing 0.5% methyl cellulose (Sigma-Aldrich) and allowed to grow for 14 days. For serial transfer experiments, single spheroids were picked with a pipette, disgregated, and processed as above. All experimental conditions were done in triplicate.

In vitro invasiveness assay. Transwell chambers (Costar) with 8-μm diameter pore membranes were coated with growth factor–reduced Matrigel (BD Biosciences) at 410 μg/ml and human umbilical cord hyaluronic acid (Sigma-Aldrich) at 100 μg/cm². Cells (1.5×10^5/well in 24-well plates; 7×10^4/well in 96-well plates) were serum deprived for 24 hours, detached, resuspended in medium supplemented with 0.1% BSA/0.5% FBS, and then seeded onto the precoated Transwell inserts, with the lower chamber containing medium supplemented with 0.5% FBS. After 24 hours, cells migrating to the lower chamber were collected by detachment with trypsin-EDTA, washed with PBS, and counted in a Coulter Multisizer II instrument (Coulter Electronics). Each experiment was done in triplicate.

Coculture and cell-sorting experiments. PC-3/Mc cells were labeled with Oregon Green 488 carboxy-DFFDA-SE (Invitrogen) with excitation maximum at 488 nm and emission at 524 nm. PC-3/S cells were labeled with Far Red DDAO-SE (Invitrogen), with excitation maximum at 600 nm and emission at 670 nm. Cells were labeled by adding 25 μM of fluorophore to the cell suspensions for 30 minutes, washed with PBS, and reseeded. Fluorophore-preloaded cells were cocultured at a 1:1 proportion for 48 or 96 hours and either assayed for invasiveness on Transwell-Matrigel chambers or sorted with a FACSAria SORP instrument (BD). After sorting, cells were either assayed for invasiveness in Transwell-Matrigel assays or processed for protein extraction (for Western blotting), RNA extraction (for quantitative PCR [qPCR]), or ChIP. As controls, PC-3/Mc cells were preloaded with Oregon Green, cocultured with unlabeled PC-3/Mc cells, and sorted by FACS.

Invasiveness assays with PC-3/S CM. PC-3/S cells were cultured at 70% confluence, at which time they were changed to fresh CD-CHO medium (Invitrogen) without FBS; CM was collected after 48 hours, centrifuged, and filtered through a 0.22-μm filter (Millipore). PC-3/Mc cells were cultured with PC-3/S CM (S-CM) mixed with fresh medium without FBS at a 1.5:1 proportion for 48 hours, and these cells were then analyzed for invasiveness and Western blotting.

Statistics. Results are expressed as mean ± SEM, illustrated as error bars. A 2-tailed Student's *t* test was applied for statistical analysis.

Study approval. All tests in this study employing human tissues were performed on postdiagnosis surplus samples obtained at the Hospital Clínic de Barcelona, following protocols previously approved by the Hospital Clínic Institutional Review Board. All animal studies were performed following protocols previously approved by the CID-CSIC Institutional Review Board. As independently certified by these 2 institutional review boards, all studies involving human or animal materials or subjects were performed in compliance with Spanish laws regulating ethics in research and patient data confidentiality (Ley de Investigación Biomédica 14/2007, de 3 de Julio de 2007).

Acknowledgments

We thank Dídac Domínguez and Marc Guiu (xenograft experiments), Mònica Marín (immunohistochemistry), Mònica Pons (confocal microscopy), and Jaume Comas and Ricard Álvarez (flow cytometry) for excellent technical assistance. T. Celià-Terrassa is a recipient of a doctoral fellowship from the CSIC. Ó. Meca-Cortés, A. Arnal-Estapé, and C. Estarás are recipients of doctoral fellowships and F. Mateo of a Juan de la Cierva postdoctoral fellowship from the Ministerio de Ciencia e Innovación (MICINN). R.R. Gomis is a researcher of the Institució Catalana de Recerca i Estudis Avançats. Financial support was provided to T.M. Thomson by MICINN (SAF2008-04136-C02-01 and SAF2011-24686), the Agència de Gestió d'Ajuts Universitaris i de Recerca (2009SGR1482), the Agencia Española de Cooperación Internacional y Desarrollo (A/023859/09), and the Xarxa de Referencia en Biotecnologia; to P.L. Fernández by MICINN (FIS-PI080274), the Fondo de Investigaciones de la Seguridad Social (PI080274), the Spanish National Biobank Network, the Instituto de Salud Carlos III (ISCIII-RETIC RD06/0020), the Xarxa de Bancs de Tumours de Catalunya-Pla Director d'Oncologia, and the Fondo Europeo de Desarrollo Regional (FEDER) — Unión Europea "Una manera de hacer Europa"; and to R.R. Gomis by the BBVA Foundation and MICINN (SAF2007-62691).

Received for publication May 31, 2011, and accepted in revised form February 29, 2012.

Address correspondence to: Timothy M. Thomson, Department of Cell Biology, Institute for Molecular Biology (IBMB), Science Research Council (CSIC), Parc Científic de Barcelona, Edifici Hèlix room 2A21, c. Baldiri Reixac 15-21, 08028 Barcelona, Spain. Phone: 34.93.4020199; Fax: 34.93.4034979; E-mail: titbmc@ibmb.csic.es.

1. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet*. 2006;7(1):21–33.
2. Axelrod R, Axelrod DE, Pienta KJ. Evolution of cooperation among tumor cells. *Proc Natl Acad Sci U S A*. 2006;103(36):13474–13479.
3. Balkwill F. Cancer and the chemokine network. *Nat Rev Cancer*. 2004;4(7):540–550.
4. Yang MH, et al. Direct regulation of TWIST by HIF-1alpha promotes metastasis. *Nat Cell Biol*. 2008;10(3):295–305.
5. Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. *Nat Rev Cancer*. 2006;6(12):924–935.
6. Moreno E. Is cell competition relevant to cancer? *Nat Rev Cancer*. 2008;8(2):141–147.
7. Wu M, Pastor-Pareja JC, Xu T. Interaction between Ras(V12) and scribbled clones induces tumour growth and invasion. *Nature*. 2010; 463(7280):545–548.
8. Karnoub AE, et al. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature*. 2007;449(7162):557–563.
9. Lapidot T, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*. 1994;367(6464):645–648.
10. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A*. 2003;100(7):3983–3988.
11. Singh SK, et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res*. 2003; 63(18):5821–5828.
12. Ben-Porath I, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet*. 2008; 40(5):499–507.
13. 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):333–344.
14. Floor S, van Staveren WC, Larsimont D, Dumont JE, Maenhaut C. Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating-cancer stem cells: distinct, overlapping or same populations. *Oncogene*. 2011;30(46):4609–4621.
15. Gupta PB, Chaffer CL, Weinberg RA. Cancer stem cells: mirage or reality? *Nat Med*. 2009;15(9):1010–1012.
16. Huber MA, Kraut N, Beug H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol*. 2005; 17(5):548–558.
17. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer*. 2007;7(6):415–428.
18. Acloque H, Adams MS, Fishwick K, Bronner-Fraser M, Nieto MA. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *J Clin Invest*. 2009;119(6):1438–1449.
19. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell*. 2009;139(5):871–890.
20. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. *Science*. 2011;331(6024):1559–1564.
21. Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell*. 2008;14(6):818–829.
22. Ansieau S, et al. Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer Cell*. 2008; 14(1):79–89.
23. Yang J, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*. 2004;117(7):927–939.
24. Mani SA, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133(4):704–715.
25. Tsuji T, et al. Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth. *Cancer Res*. 2008; 68(24):10377–10386.
26. Tsuji T, Ibaragi S, Hu GF. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res*. 2009;69(18):7135–7139.
27. Hermann PC, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell*. 2007;1(3):313–323.
28. Korpal M, et al. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nat Med*. 2011; 17(9):1101–1108.
29. El Hilali N, Rubio N, Martinez-Villacampa M, Blanco J. Combined noninvasive imaging and luminometric quantification of luciferase-labeled human prostate tumors and metastases. *Lab Invest*. 2002;82(11):1563–1571.
30. Kozlowski JM, Fidler IJ, Campbell D, Xu ZL, Kaighn ME, Hart IR. Metastatic behavior of human tumor cell lines grown in the nude mouse. *Cancer Res*. 1984; 44(8):3522–3529.
31. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. *Nat Protoc*. 2007;2(12):3081–3089.
32. Wernig M, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*. 2007;448(7151):318–324.
33. Kim J, et al. A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell*. 2010;143(2):313–324.
34. Coppe JP, et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol*. 2008;6(12):2853–2868.
35. Kuilman T, Peeper DS. Senescence-messaging secretome: SMS-ing cellular stress. *Nat Rev Cancer*. 2009; 9(2):81–94.
36. Chaffer CL, Brennan JP, Slavin JL, Blick T, Thompson EW, Williams ED. Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2. *Cancer Res*. 2006; 66(23):11271–11278.
37. Chaffer CL, et al. Upregulated MT1-MMP/TIMP-2 axis in the TSU-Pr1-B1/B2 model of metastatic progression in transitional cell carcinoma of the bladder. *Clin Exp Metastasis*. 2005;22(2):115–125.
38. Hurt EM, Kawasaki BT, Klarmann GJ, Thomas SB, Farrar WL. CD44+ CD24(-) prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis. *Br J Cancer*. 2008;98(4):756–765.
39. Shipitsin M, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell*. 2007;11(3):259–273.
40. Li C, et al. Identification of pancreatic cancer stem cells. *Cancer Res*. 2007;67(3):1030–1037.
41. Yeung TM, Gandhi SC, Wilding JL, Muschel R, Bodmer WF. Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci U S A*. 2010; 107(8):3722–3727.
42. Tani H, Morris RJ, Kaur P. Enrichment for murine keratinocyte stem cells based on cell surface phenotype. *Proc Natl Acad Sci U S A*. 2000; 97(20):10960–10965.
43. Peiro S, et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res*. 2006;34(7):2077–2084.
44. Kopp H-G, Avecilla ST, Hooper AT, Rafii S. The bone marrow vascular niche: home of HSC differentiation and mobilization. *Physiology (Bethesda)*. 2005; 20:349–356.
45. Taylor BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18(1):11–22.
46. Jeschke U, et al. Expression of E-cadherin in human ductal breast cancer carcinoma in situ, invasive carcinomas, their lymph node metastases, their distant metastases, carcinomas with recurrence and in recurrence. *Anticancer Res*. 2007;27(4A):1969–1974.
47. Park D, Karesen R, Axcrona U, Noren T, Sauer T. Expression pattern of adhesion molecules (E-cadherin, alpha-, beta-, gamma-catenin and claudin-7), their influence on survival in primary breast carcinoma, and their corresponding axillary lymph node metastasis. *APMIS*. 2007;115(1):52–65.
48. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009; 119(6):1420–1428.

49. Li R, et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*. 2010;7(1):51–63.
50. Polo JM, Hochedlinger K. When fibroblasts MET iPSCs. *Cell Stem Cell*. 2010;7(1):5–6.
51. Samavarchi-Tehrani P, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*. 2010;7(1):64–77.
52. Lowry WE. E-cadherin, a new mixer in the Yamanaka cocktail. *EMBO Rep*. 2011;12(7):613–614.
53. Redmer T, Diecke S, Grigoryan T, Quiroga-Negreira A, Birchmeier W, Besser D. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Rep*. 2011;12(7):720–726.
54. Nieto MA, Sargent MG, Wilkinson DG, Cooke J. Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science*. 1994; 264(5160):835–839.
55. Sauka-Spengler T, Bronner-Fraser M. A gene regulatory network orchestrates neural crest formation. *Nat Rev Mol Cell Biol*. 2008;9(7):557–568.
56. Liu J, et al. Slug inhibits proliferation of human prostate cancer cells via downregulation of cyclin D1 expression. *Prostate*. 2010;70(16):1768–1777.
57. Vega S, Morales AV, Ocana OH, Valdes F, Fabregat I, Nieto MA. Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev*. 2004; 18(10):1131–1143.
58. Dykxhoorn DM, et al. miR-200 enhances mouse breast cancer cell colonization to form distant metastases. *PLoS One*. 2009;4(9):e7181.
59. Lou Y, et al. Epithelial-mesenchymal transition (EMT) is not sufficient for spontaneous murine breast cancer metastasis. *Dev Dyn*. 2008; 237(10):2755–2768.
60. Olmeda D, Moreno-Bueno G, Flores JM, Fabra A, Portillo F, Cano A. SNAI1 is required for tumor growth and lymph node metastasis of human breast carcinoma MDA-MB-231 cells. *Cancer Res*. 2007; 67(24):11721–11731.
61. Ke XS, et al. Epithelial to mesenchymal transition of a primary prostate cell line with switches of cell adhesion modules but without malignant transformation. *PLoS One*. 2008;3(10):e3368.
62. Kaighn ME, Narayan KS, Ohnuki Y, Lechner JF, Jones LW. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest Urol*. 1979;17(1):16–23.

Supplemental data

Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells

Toni Celià-Terrassa, Óscar Meca-Cortés, Francesca Mateo, Alexia Martínez de Paz, Nuria Rubio, Anna Arnal-Estapé, Brian J. Ell, Raquel Bermudo, Alba Díaz, Marta Guerra-Rebollo, Juan José Lozano, Conchi Estarás, Catalina Ulloa, Daniel Álvarez-Simón, Jordi Milà, Ramón Vilella, Rosanna Paciucci, Marian Martínez-Balbás, Antonio García de Herreros, Roger R. Gomis, Yibin Kang, Jerónimo Blanco, Pedro L. Fernández, and Timothy M. Thomson

Contents:

Supplemental Methods

Sources of primary antibodies to: E-cadherin intracytoplasmic domain (clone 36); E-cadherin extracellular domain (clone HECD-1), EpCAM (B29.1/VU-ID9), trimethyl histone H3K4 (Abcam, Cambridge, UK); fibronectin, ZEB1 (Sigma); vimentin (clone V9, Labvision); SOX2 (clone D6D9, Cell Signaling); human SNAI1 (SN9H2), human TWIST1 (Cell Signaling); actin (I-19), CD44 (DF1485, or produced and purified in-house from hybridoma 33-3B3), SPARC (H-90), CD24 (ML5), CD40 (LOB-11), luciferase (251-550) (Santa Cruz Biotechnology, Santa Cruz, CA); CD71 (produced and purified in-house from hybridoma 120-2A3); acetyl histone H3, acetyl histone H4, trimethyl histone H3K27 (Upstate Millipore, Billerica, MA); hemagglutinin (Clone 3F10, Roche, Diagnostics, Mannheim, Germany); MYC (Epitomics, Burlingame, CA). LiCl and MG132 were from Sigma.

Growth curves. Cells (5 x 10^3/well) were seeded in sextuplicate in 96-well plates (Costar, Corning, NY). All cells used have the firefly luciferase gene stably integrated in their genomes. Standard curves were generated to correlate cell numbers to luminescence levels yielded by cell lysates, using the Luciferase Reporter Assay Kit (Promega, Madison, WI). For growth curves, cells were harvested daily for 6 or 7 consecutive days, luminescence quantified, and cell numbers extrapolated on the basis of standard curves. Luminescence was quantified either on a Victor 3 instrument (Perkin-Elmer, Wellesley, MA) or an OrionII Microplate Luminometer (Berthold Detection Systems, Pforzheim, Germany).

Wound healing assays. Cells (2 x 10^5/well) were seeded in 24-well plates, allowed to reach confluent monolayers and serum-starved for 24 h. Wounds were created with a 0.5 mm plastic pipette tip. Afterwards, cells were fed with medium with 0.5% or 10% FBS, to create differential healing conditions. Images were captured and surface areas between leading edges of the monolayers at predetermined wound sites (3 sites per condition, performed in triplicate) were measured for the following 54 h and analyzed with the ImageJ software.

Cell cycle analysis. Cells were seeded in 6-well plates (Costar), detached with Trypsin/EDTA/1% BSA, washed twice and resuspended in PBS, followed by dropwise addition of 70% ethanol and fixation at 4 ºC for 2 h. Subsequently, fixed

cells were washed twice with PBS/50 mM EDTA/1% BSA, incubated with 1 mg/mL RNase A (Sigma) at 37ºC for 1 h, and 0.1 mg/mL propidium iodide (Sigma). DNA content was determined with a Cytomics FC500 instrument (Coulter, Hialeah, FL), and cell cycle distribution analyzed with the Multicycle program coupled to the instrument. All cell cycle determinations were done in triplicate.

***In vivo* tumorigenic assays**. For orthotopic implantation, 10^5 cells in a volume of 25 μL were inoculated in the ventral lobe of anesthetized 6-week-old male NOD-SCID mice. Eight mice were implanted for each cell type analyzed. To assess localized growth, cells were xenografted by intramuscular injection in 6-week-old male NOD-SCID mice, 4 mice for each cell line. The injections were 2.5 x 10^5 cells in a volume of 50 μL RPMI 1640 (without FBS) in each hind limb (2 injection sites per mouse). Tumor growth was monitored once or twice a week by luminometry on an ORCA-2BT instrument (Hamamatsu Photonics, Hamamatsu, Japan). Images and light counts were initiated 5 min after intraperitoneal injection of luciferine (100 mg/Kg in 150 μL). Animals were allowed to form tumors up to 1.5 cm in diameter, at which point they were euthanized. To monitor metastatic growth after orthotopic implantation experiments, separate images and photon counts were obtained for the thoracic (anterior) and abdominal (posterior) areas of mice. The images obtained were analyzed by the Hokawo 2.1 software (Hamamatsu Photonics). To assess lung colony formation, 2.5 x 10^5 cells in 150 μL RPMI 1640 (without FBS) were injected through the dorsal caudal vein of anesthetized 6-week-old male NOD-SCID mice (6 mice per cell line). The images obtained were analyzed with the Hokawo 2.1 software. To assess colonization to other organs, 2 x 10^5 cells in 100 μL PBS were injected in the left ventricle of anesthetized 6-week-old Balb c/nude, NOD-SCID or athymic Ncr-nu/nu mice (6 to 10 mice per cell line). Mice were imaged immediately after injection and thereafter tumor development was monitored by weekly imaging on an IVIS-200 instrument (Xenogen-Caliper Life Sciences, Hopkinton, MA) after retroorbital injection of 1.5 mg luciferine (15 mg/mL in PBS). For bioluminescence plots, photon flux was calculated for each mouse by using a circular region of interest of the mouse in a supine position. This value was scaled to a comparable background value (from a luciferin-injected mouse with no tumor cells), and then normalized to the value obtained immediately after xenografting at the same area (day 0), so that all mice had an arbitrary starting Bioluminiscent I signal of 10,000. Lesions were localized by *ex*

vivo bioluminescence imaging and resected under sterile conditions. Some of the lesions were fixed with formalin and processed for histological analysis. Statistics: in lung colonization and bone metastasis free survival analysis, lesions that had an increased photon flux value above day 0 were counted as events. Statistics were performed by Log Rank (Mantel-Cox) test using SPSS software. For PC-3/S-CDH1 and PC-3/S-SOX2 tumor growth, cells (2 x 10^5) were implanted i.m. into 6-week old male Swiss-nude mice and tumor size monitored with a caliper. For serial xenotransplantation experiments, 10^3, 10^4 or 10^5 cells were implanted i.m. in the hind limbs of anesthetized 6-week-old male Swiss-Nude mice. Tumors formed by implantation of 10^5 cells were extracted, subjected to mechanical disgregation and digestion with 300 U/mL collagenase A (Sigma) in Hank's Balanced Salt Solution (PAA) for 30 min at 37ºC, selected for neomycin resistance for 7 d in order to remove contaminating mouse cells, and reimplanted i.m. in a second series of animals, as above. This process was repeated in a third series of animals. Tumor growth was monitored with a caliper, and tumor volumes calculated as described in (1).

RNA isolation, reverse transcription and transcriptomic analysis. Cells were grown to 70-80% confluence, lysed and RNA isolated with the RNeasy Kit (Qiagen, Hilden, Germany), including a DNase digestion step. For microarray analysis, RNAs were amplified, labeled and hybridized to Affymetrix U133 2.0 Plus arrays (Affymetrix, Santa Clara, CA). Microarray data were normalized using the robust multi-array (RMA) algorithm (2). Next, we employed a conservative probe-filtering step, eliminating those probes with a maximum expression value lower than 5. To identify differentially expressed genes, we applied Significance Analysis of Microarrays (SAM-R) (3), selecting those genes with a False Discovery Rate (FDR) below 10% (Q < 10). Comparative transcriptomic analysis was performed on independent triplicate samples.

Gene set enrichment analysis. For comparative pathway and gene set correlations, Gene Set Enrichment Analysis (GSEA) (4, 5) was applied in order to identify overrepresented predefined gene sets using C2 and C5 genesets from MSigDB v3. Additionally, 4 reported gene sets (ESC-like module (6), MYC module (7), ES1 and ES2 modules (10)) were analyzed for their representation in PC-3/Mc and PC-3/S cells using normalized microarray expression data. Gene-set permutation type with 1,000

random permutations was run to obtain a rank gene list, enrichment plots and heat maps. For GSEA of a previously published prostate cancer expression dataset (11), a gene set was generated consisting of a subset of genes of the ESC-like module most significantly enriched in PC-3/Mc cells (M geneset) and used to run GSEA on a continuous phenotype modality, in which a numerical variable was generated for three values corresponding to metastatic samples (class 1), primary tumor samples with pathological stages T3 or T4 (class 2) and primary tumor samples with pathological stages T1 or T2 (class 3), and a Pearson correlation was applied as the metric to determine correlations the representation of the M geneset and these three classes of samples taken as continuous variables (http://www.broadinstitute.org/gsea/doc/GSEAUserGuideFrame.html?_Metrics_for_Ranking).

Real-time RT-PCR (qPCR). Complementary DNAs were synthesized with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Carlsbad, CA). Real-time quantitative PCR assays were performed in a LightCycler 480 instrument (Roche Diagnostics) and analyzed with the LightCycler 480 Software release 1.5.0. Either gene-specific TaqMan assays (Applied Biosystems) or the Universal Probe Library system (UPL; Roche) were used, following the specific running conditions recommended in each case. The following transcripts were quantified with TaqMan assays (Supplemental Table 10): *SOX2*, *RUNX2* and *TWIST2*. The following transcripts were quantified with the UPL system (Supplemental Table 11): *NANOG*, *LIN28A*, *KLF9*, *KLF4*, *MYC*, *POU5F1* (*OCT4*), *SNAI1*, *SNAI2*, *ZEB1*, *ZEB2*, *TWIST1*, *RUNX1*, *CDH1*, *DSP*, *TACSTD1*, *CDH2*, *SPARC*, *VIM* and *FN1*. The amplification levels of *RN18S1* and HMBS were used as an internal reference to estimate the relative levels of specific transcripts, and relative quantification was determined by the ΔΔCp method. All determinations were done in triplicate and values represented as $\log_{10}$ of $2^{-\Delta\Delta Cp}$ (12). In some cases in which transcriptional profiles from multiple samples were compared, relative expression values were represented as a heatmap with pseudocoloring ranging from green (underexpressed relative to values in control cells) to red (overexpressed relative to control cells).

Cell surface immunophenotyping. Cells were detached with PBS/2 mM EDTA for 20-30 min, washed, incubated with primary antibody diluted 1:20 in PBS/3% normal

goat serum for 30 min in a shaker at 4 °C, washed, incubated with the secondary antibody (Alexa Fluor 488, anti-mouse, 1:200 dilution; Invitrogen) for 30 min, washed and analyzed by flow cytometry on a Cytomics FC500 instrument (Coulter, Hialeah, FL)

Immunocytochemistry. Sterile coverslips placed at the bottom of 24-well plates were seeded with 3 x 10^4 cells, allowed to attach for 72 h, washed with PBS and fixed for 1 h with methanol at -20 ºC. After fixation, samples were washed with acetone, then 5 times with PBS, blocked for 30 min with blocking buffer (5% normal goat serum, 0.5% Triton X-100 in PBS) and incubated with anti-E-cadherin (1:100; Clone 36; Becton-Dickinson), anti-SNAI1 (1:20) or anti-SOX2 (1:50) for 2 h at room temperature. This was followed by PBS washes, a 1 h incubation with Alexa Fluor 488-conjugated goat-anti-mouse IgG (1:1,000), and final washes followed by incubation with DAPI (1:1,000; Sigma) for 10 min. Coverslips were mounted on slides with Mowiol 4-88 and images captured with a Leica SP5 confocal microscope.

Immunoblotting. Cell lysates were boiled in Laemmli sample buffer, electrophoresed by SDS-PAGE, transferred to PVDF membranes (Immobilon-FL; Millipore, Billerica, MA) and blotted with antibodies to human E-cadherin (1:8,000), vimentin (1:2,000), fibronectin (1:500), SNAI1 (1:500), TWIST1 (1:1,000), SOX2 (1:1,000), MYC (1:1,000), SPARC (1:500), EpCAM (1:500), ZEB1 (1:1,000), actin (1:2,000). Reactions were detected with fluorescent dye-conjugated secondary antibodies (IRDye 800CW Goat Anti-Mouse IgG; IRDye 680 Goat Anti-Rabbit IgG; IRDye 680 Donkey Anti-Goat IgG) on an Odyssey infrared imaging system (Li-COR Biosciences, Lincoln, NE). Sample loadings were normalized by detection of actin or β-tubulin levels.

Chromatin immunoprecipitation (ChIP). The procedure followed that described by Kimura *et al*. (13). Briefly, 7 x 10^5 cells were collected and fixed with 1% formaldehyde in PBS, centrifuged and washed twice with PBS. Then, cells were resuspended in 300 µL lysis buffer (10% SDS, 10 mM EDTA, 1 M Tris pH 8, 100 mM PMSF, 1µg/µL leupeptin, 1µg/µL aprotinin, 1 M Na_5VO_4). Cell lysates were sonicated in a BioRuptor sonicator at maximum potency for 13 min. Chromatin was pre-cleared with 5% salmon sperm DNA (2 µg) / Protein A Agarose (Millipore) for 2

h at 4 ºC. Immunoprecipitation was performed by incubation with 2 µg of primary antibodies in 5% salmon sperm DNA (2 µg) /Protein A agarose overnight at 4ºC. As a negative control, 2 µg of normal rabbit IgG was used. Specific DNA segments from the immunoprecipitated material were amplified with primers for selected gene promoters, and the PCR products were quantified by SYBRGreen incorporation under real-time conditions in a LightCycler 480 Instrument (Roche). The primers used for amplification (sequences as shown below) were designed within the 1 Kb region upstream of the transcript start site for each gene, with the help of the LightCycler Probe Design2 Software (Roche), and were tested for specificity and efficiency of amplification prior to their use in ChIP (Supplemental Table 12). The results are expressed as specific amplification levels on immunoprecipitated DNA relative to the amplification levels yielded by input DNA. All determinations were done in triplicate.

Production and transduction of retroviral particles. Construct pBABEpuro-Twist1 (mouse) was kindly provided by Dr. Gabriel Gil (IMIM, Barcelona). pBABEpuro-Snai1-HA (mouse), pMSCV-Flag-SOX2 (human) and pWZL-Blast-E-cadherin (mouse) have been described (14-16). pBABEpuro-HA-TWIST2 was constructed inserting full-length human TWIST2 cDNA into pCMV-HA in frame with the hemagglutinin epitope and then into pBABEpuro. The retrovirus packaging cell line PG13 was co-transfected with these DNAs and pVSV-G (Clontech, Mountain View, CA) for 12 h using Fugene HD (Roche). Supernatants were collected during the following 48 h and filtered through 0.45 µm methylcellulose filters (Millipore). Retroviral particles were concentrated by ultracentrifugation at 27,000 rpm for 90 minutes on 20% sucrose density gradients. Viral particles were resupended with medium and added to the cells together with 8 µg/mL polybrene (Sigma). Plates were centrifuged at 1,800 rpm for 60 minutes and cells allowed to recover in fresh medium for 24-48 h. Cells with integrated retroviral sequences were selected for 5 days in medium supplemented with 3 µg/mL puromycin (Biomol) for PC-3/Mc cells or 1 µg/mL puromycin for TSU-Pr1-B2 cells. PC-3/S cells transduced with pWZL-Blast-E-cadherin were selected with 25 µg/mL blastocydin for 10 d.

Production and transduction of lentiviral particles. Constructs pLKOpuro-shRNA-CDH1 (TRCN0000039665 and TRCN0000039666), pLKOpuro-shRNA-SOX2 (TRCN0000003252, TRCN0000003253 and TRCN0000010772), pLKOpuro-shRNA-

KLF4 (TRCN0000005314, TRCN0000005313, TRCN0000005315 and TRCN0000005316), pLKOpuro-shRNA-MYC (TRCN0000039640, TRCN0000039641 and TRCN0000039642), pLKOpuro-shRNA-SNAI1 (TRCN0000063819, TRCN0000063818, TRCN0000063820, TRCN0000063821 and TRCN0000063822), pLKOpuro-shRNA-TWIST2 (TRCN0000020869, TRCN0000020870, TRCN0000020871 and TRCN0000020872) and pLKOpuro-shRNA-ZEB1 (TRCN0000017565, TRCN0000017563, TRCN0000017564 and TRCN0000017566) and the non-target control vector Shc002 were purchased from Sigma-Aldrich. The bicistronic expression vectors pRRL-Luc-IRES-EGFP and pRRL-Renilla-IRES-RFP were as described (17). The lentivirus packaging cell line HEK293T was co-transfected with these DNAs, pCMVdeltaR8.91 and pVSV-G (Clontech, Mountain View, CA) for 12 h using Fugene HD (Roche). Supernatants were collected for the following 48 h and filtered through 0.45 μm methylcellulose filters (Millipore). Lentiviral particles were concentrated by ultracentrifugation at 27,000 rpm for 90 minutes on 20% sucrose density gradients. Viral particles were resupended with medium and added to the cells together with 8 μg/mL polybrene (Sigma). Cells were infected for 24 hours and allowed to recover in fresh medium for 24-48 h. Selection for cells with integrated lentiviral sequences was performed as above.

Immunohistochemistry. Two μm thick sections were obtained for immunohistochemistry either from formalin-fixed and paraffin-embedded tissue blocks or from tissue microarrays (TMA), built with a Manual Tissue Arrayer 1 (Beecher Instruments, Sun Prairie, WI). A total of 99 tumors, 37 normal samples from human specimens were analyzed, as well as tumors from immunodeficient mouse xenografts. Tissue sections were mounted on xylaned glass slides (DAKO, Glostrup, Denmark) and used for immunohistochemical staining using the Bond Polymer Refine Detection System (Leica, Wetzlar, Germany). Samples were deparaffinized, antigen retrieval performed at pH 6 for 20 minutes and primary antibodies incubated for 1 hour. Antibody dilutions used were 1:50 for SOX2 (Clone D6D9, Cell Signaling) and 1:50 or 1:150 for E-cadherin (NCL-E, Novocastra, Newcastle, UK) on mouse and human samples, respectively. SOX2 staining was scored as the percentage of cells with nuclear positivity and the predominant staining intensity. E-cadherin staining was assessed as a percentage of membrane or cytoplasmic pattern. Images were captured

with an Olympus BX-51 microscope equipped with an Olympus DP70 camera. Human tissues were procured from patients at the Hospital Clínic of Barcelona after informed consent by the patients and approval by the Institutional Ethics Committee.

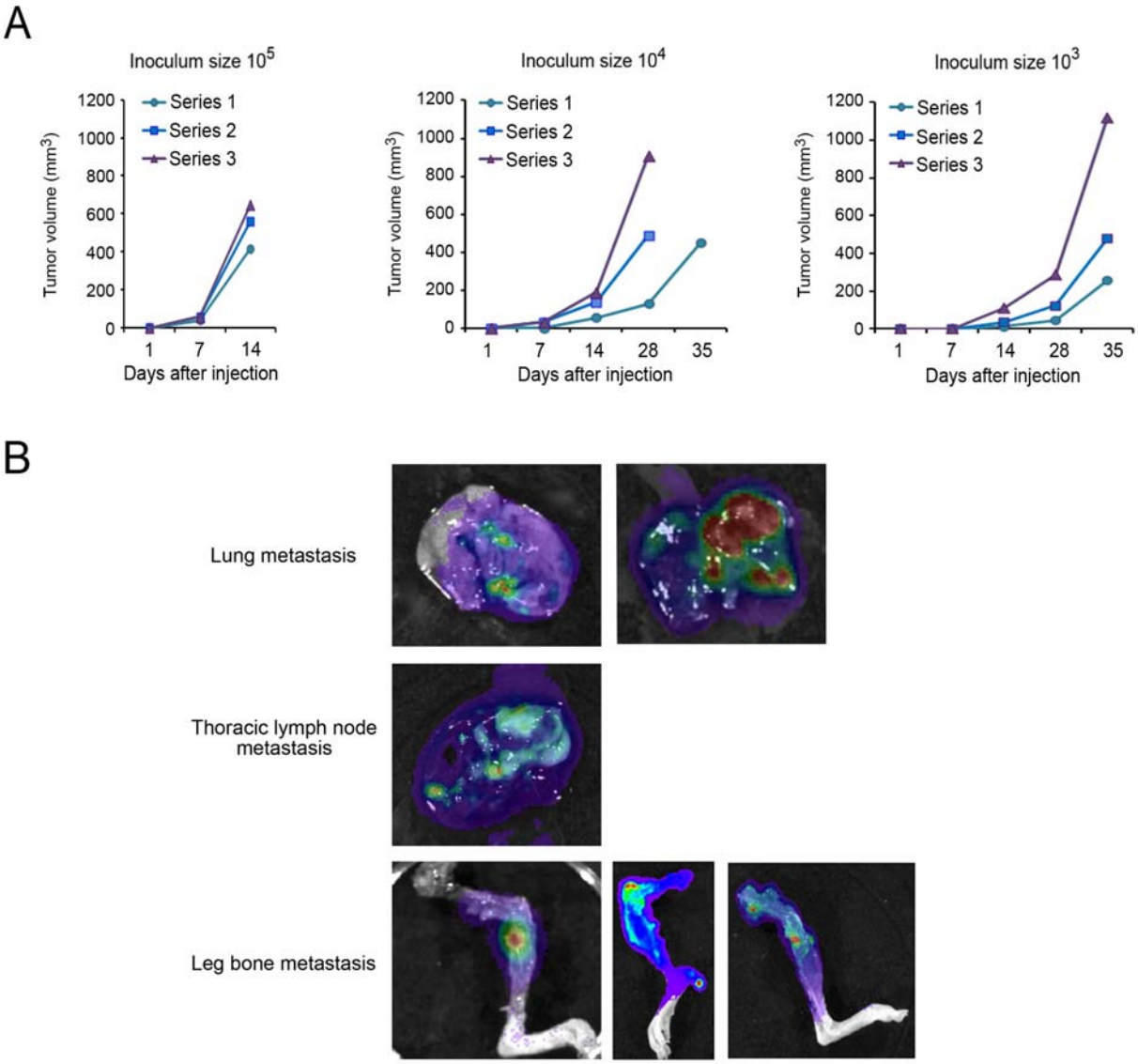
Immunofluorescence of tumors from xenograft experiments. Tumors from xenografted mice were snap-frozen in OCT and 2-µm sections fixed in 4% paraformaldehyde for 15 min, washed twice with PBS and incubated with blocking solution (0.2% saponin, 5% normal goat serum in PBS). Primary antibodies were incubated for 2 h, washed with PBS, incubated with fluorescent-conjugated secondary antibodies, washed and counterstained with DAPI (1:1,000; Sigma). Slides were mounted with Mowiol 4-88 and images captured with a Leica SP5 confocal microscope. Rabbit anti-luciferase (Santa Cruz) was used at 1:100. As a secondary antibody, Alexa Fluor 647-conjugated goat-anti-rabbit IgG (Invitrogen) was used at 1:200.

Supplemental References

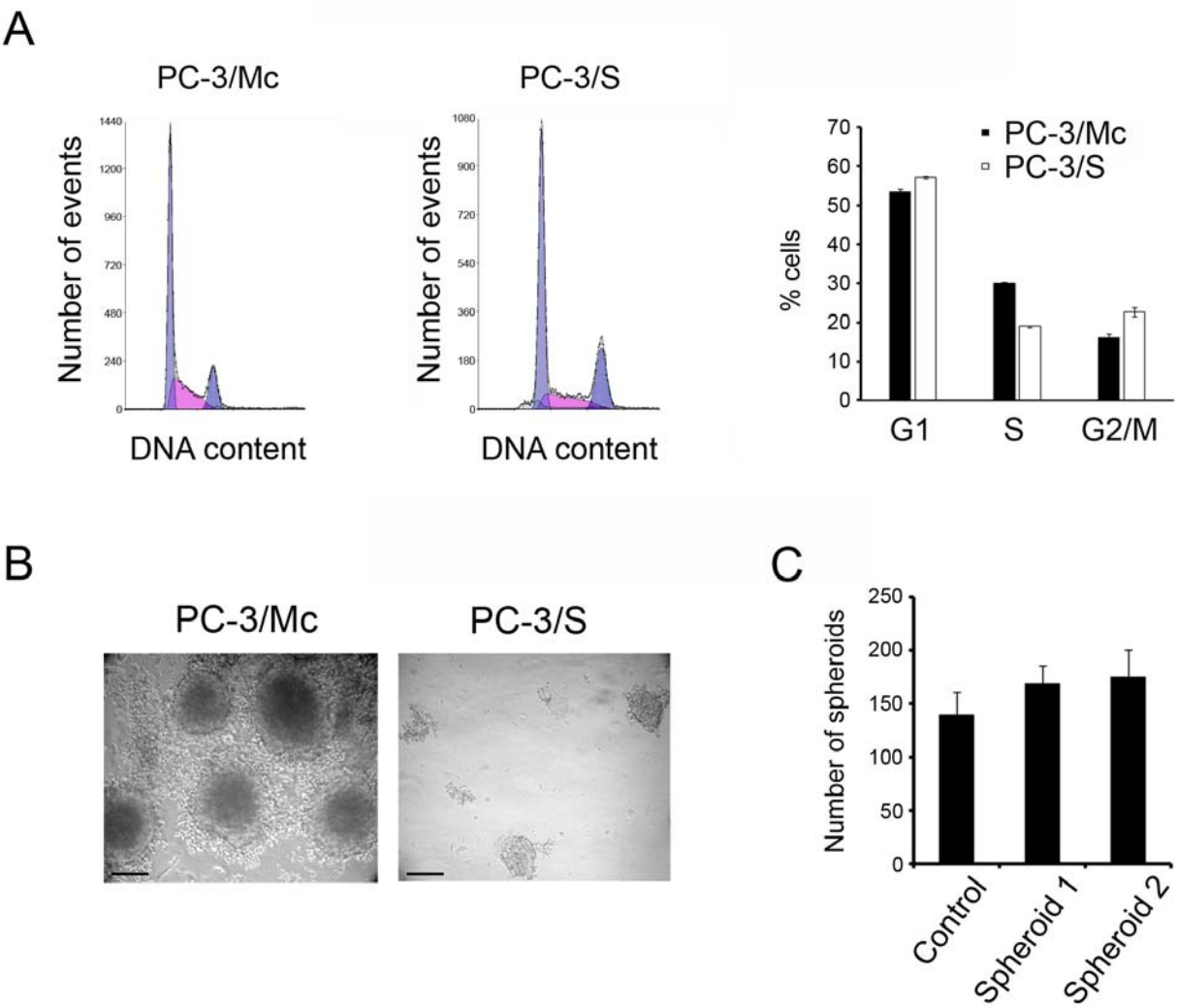
1. Feldman, J.P., Goldwasser, R., Mark, S., Schwartz, J., and Orion, I. 2009. A mathematical model for tumor volume evaluation using two-dimensions. *J Appl Quant Meth* 4:455-462.
2. Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31:e15.
3. Tusher, V.G., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98:5116-5121.
4. Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34:267-273.
5. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545-15550.
6. Wong, D.J., Liu, H., Ridky, T.W., Cassarino, D., Segal, E., and Chang, H.Y. 2008. Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* 2:333-344.
7. Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. 2010. A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* 143:313-324.
8. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947-956.
9. Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. 2006. A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444:364-368.
10. Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40:499-507.
11. Taylor, B.S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B.S., Arora, V.K., Kaushik, P., Cerami, E., Reva, B., et al. 2010. Integrative genomic profiling of human prostate cancer. *Cancer Cell* 18:11-22.
12. http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_040980.pdf.
13. Kimura, H., Hayashi-Takanaka, Y., Goto, Y., Takizawa, N., and Nozaki, N. 2008. The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies. *Cell Struct Funct* 33:61-73.
14. Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., et al. 2008. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol* 26:1276-1284.

15. Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia De Herreros, A. 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumor cells. *Nat Cell Biol* 2:84-89.
16. Onder, T.T., Gupta, P.B., Mani, S.A., Yang, J., Lander, E.S., and Weinberg, R.A. 2008. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* 68:3645-3654.
17. Degano, I.R., Vilalta, M., Bago, J.R., Matthies, A.M., Hubbell, J.A., Dimitriou, H., Bianco, P., Rubio, N., and Blanco, J. 2008. Bioluminescence imaging of calvarial bone repair using bone marrow and adipose tissue-derived mesenchymal stem cells. *Biomaterials* 29:427-437.
18. Peiro, S., Escriva, M., Puig, I., Barbera, M.J., Dave, N., Herranz, N., Larriba, M.J., Takkunen, M., Franci, C., Munoz, A., et al. 2006. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res* 34:2077-2084.

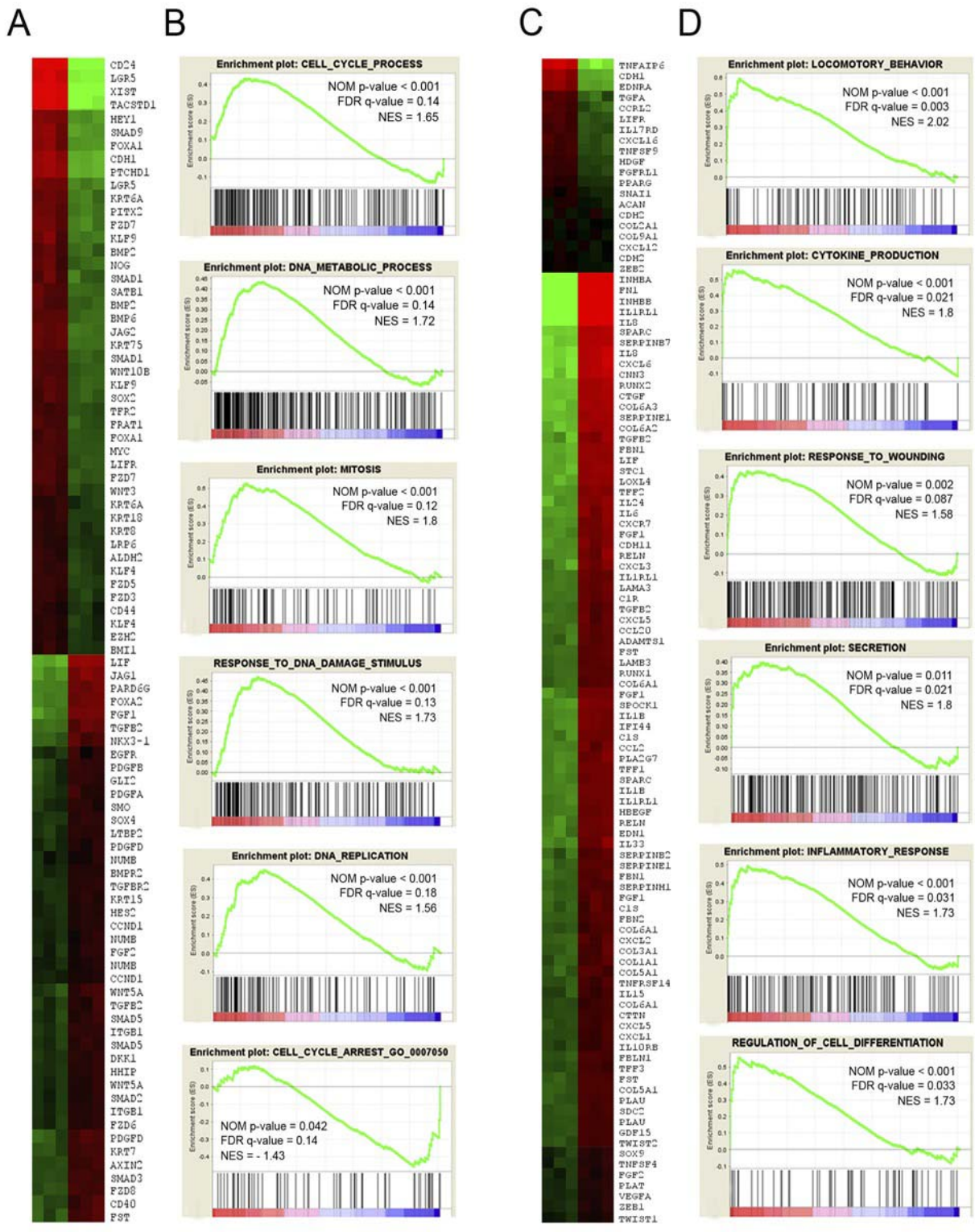
Supplemental Figures



Supplemental Figure 1. Serial transplantation and metastatic growth of PC-3/Mc cells. (**A**) Serial transplantation of PC-3/Mc cells in Swiss-Nude mice. PC-3/Mc cells (10^5, 10^4 or 10^3) were implanted intramuscularly in the hind limbs of 5-week-old male Swiss-Nude mice (Series 1). Tumors formed by implantation of 10^5 cells were explanted, disgregated and reimplanted i.m. in additional mice (10^5, 10^4 or 10^3 cells) (Series 2). Newly formed tumors were equally processed and implanted in a third series of mice (Series 3). Four mice were transplanted in each series for each inoculum size. Tumors were measured with a caliper and volumes estimated as described in (1). (**B**) Bioluminescent images of metastatic tumors grown in lungs, lymph nodes and bone after intracardiac inoculation of PC-3/Mc cells. Cells (2 x 10^5) were inoculated and their growth monitored by bioluminescence (Fig. 1). At the end of the monitoring period, mice were anesthetized, administered luciferine (1.5 mg retroorbitally), sacrificed and metastatic growth detected visually with the aid of a IVIS-200 instrument. Visible metastases were removed and imaged *ex vivo*.

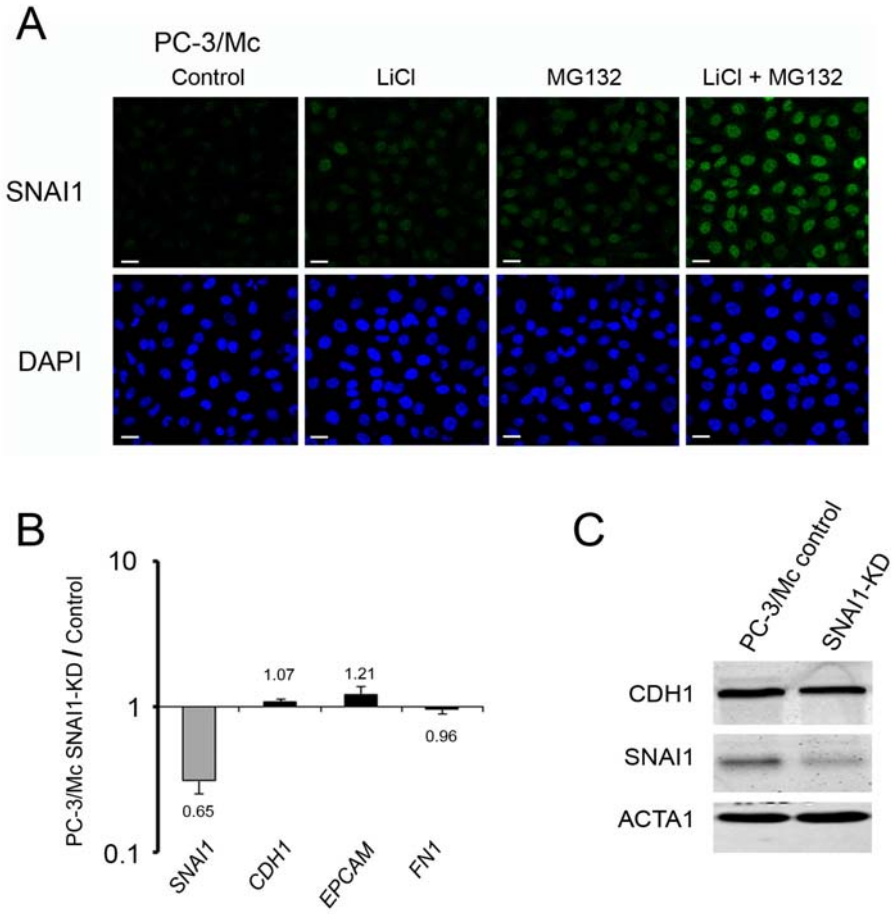


Supplemental Figure 2. Additional growth properties of PC-3/Mc and PC-3/S prostate cancer subpopulations. (**A**) Cell cycle analysis of PC-3/Mc and PC-3/S cells, with distribution histograms relative to DNA content and (right panel) distribution of populations in G1, S and M phases. (**B**) Spheroid formation assays of PC-3/Mc compared to PC-3/S cells. The figure is representative of phase-contrast images for cells grown for 14 days in low-attachment plates in medium containing 0.5% methyl cellulose Scale bar: 150 μm. (**C**) *In vitro* serial transplantation of spheroids formed by PC-3/Mc cells. Spheroids formed after 14 days of culture were dispersed, replated (10^3 cells) and grown as in (B) (Spheroid 1). The processed was repeated a third time (Spheroid 2). Graphs represent mean values and standard deviations of triplicate experiments.

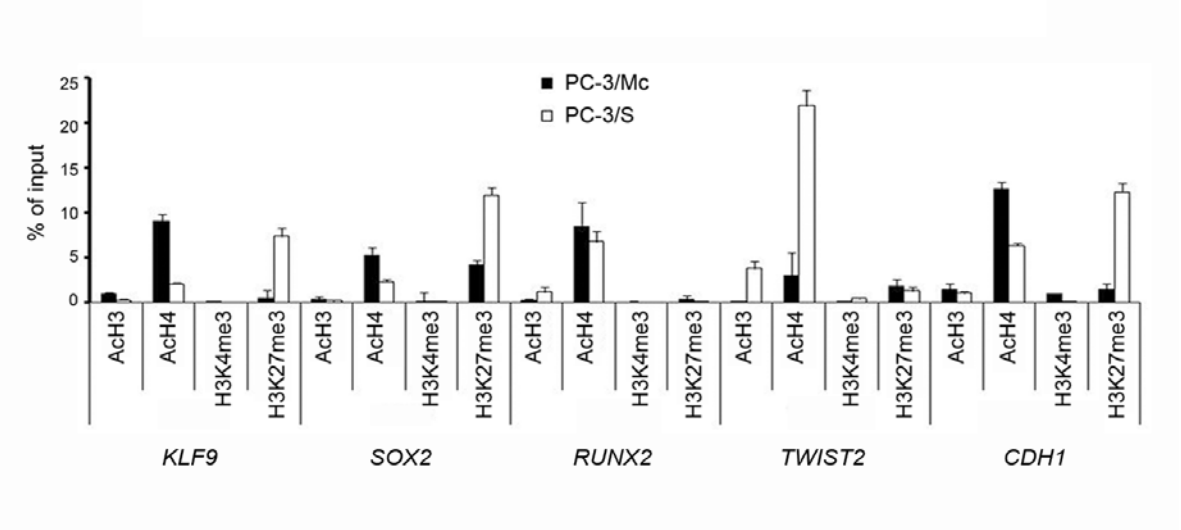


Supplemental Figure 3. PC-3/Mc cells express epithelial and self-renewal genes, while PC-3/S cells express mesenchymal genes and secretory and inflammatory gene networks. (**A**) Higher levels of expression by PC-3/Mc cells, as compared to PC-3/S cells, of epithelial genes, including E-cadherin (*CDH1*) or EpCAM (*TACSTD1*). Heatmap of normalized expression values for genes selected from a microarray analysis using Affymetrix U133A, comparing transcriptomes for PC-3/Mc and PC-3/S cells. (**B**) Enrichment in PC-3/Mc cells of

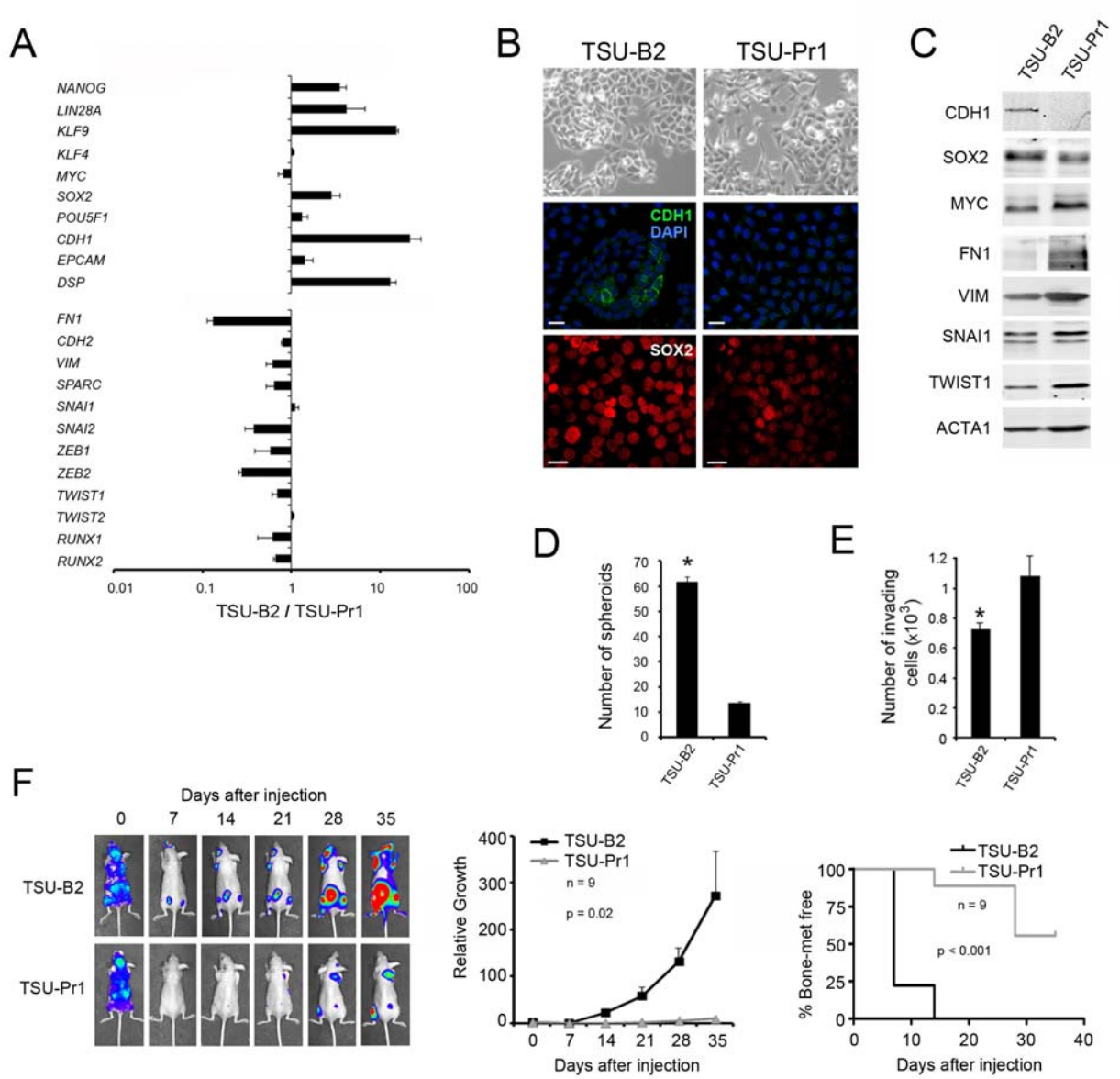
gene sets for cell cycle, mitosis, DNA damage response and DNA replication. GSEA analysis was performed for microarray data comparing transcriptomes for PC-3/Mc *vs*. PC-3/S cells. (**C**) Higher levels of expression by PC-3/S cells, relative to PC-3/Mc cells, of mesenchymal genes, including fibronectin (*FN1*), *SPARC*, *TWIST2* or *ZEB1*. Heatmap of normalized expression values for genes selected from a microarray analysis using Affymetrix U133A, comparing transcriptomes for PC-3/Mc and PC-3/S cells. (**D**) Enrichment in PC-3/S cells of gene sets for motile phenotypes, cytokines, secretion or inflammation. GSEA analysis was performed for microarray data comparing transcriptomes for PC-3/Mc *vs*. PC-3/S cells.



Supplemental Figure 4. Knockdown of *SNAI1* in the epithelial PC-3/Mc cells does not significantly affect the expression of E-cadherin, EpCAM or fibronectin. (**A**) Enhanced expression and nuclear localization of SNAI1 in PC-3/Mc cells by treatment for 5 hours with the GS3K inhibitor LiCl (50 mM), the proteasome inhibitor MG132 (5 μM), or both, determined by indirect immunofluorescence. Scale bar: 20 μm. (**B**) The levels of E-cadherin transcript levels in PC-3/Mc did not vary significantly after knockdown of *SNAI1*. Real-time determination of relative transcript levels of *SNAI1*, E-cadherin (*CDH1*), EpCAM (*TACSTD1*) and fibronectin (*FN1*) after knockdown of *SNAI1*. PC-3/Mc cells were lentivirally transduced with a *SNAI1*-specific shRNA, selected for puromycin resistance for over 5 days and processed for expression analyzes. Real-time Cp values were normalized to values for *RN18S1* (reference transcript), and represented as the $\log_{10}$ of ratios between knockdown and control cells. Controls were PC-3/Mc transduced with a pLK0-scrambled vector and selected for puromycin resistance. (**C**) The levels of E-cadherin protein in PC-3/Mc did not vary significantly after knockdown of *SNAI1*, as determined by Western blotting. Actin signal was used as an indicator of protein loading and transfer.

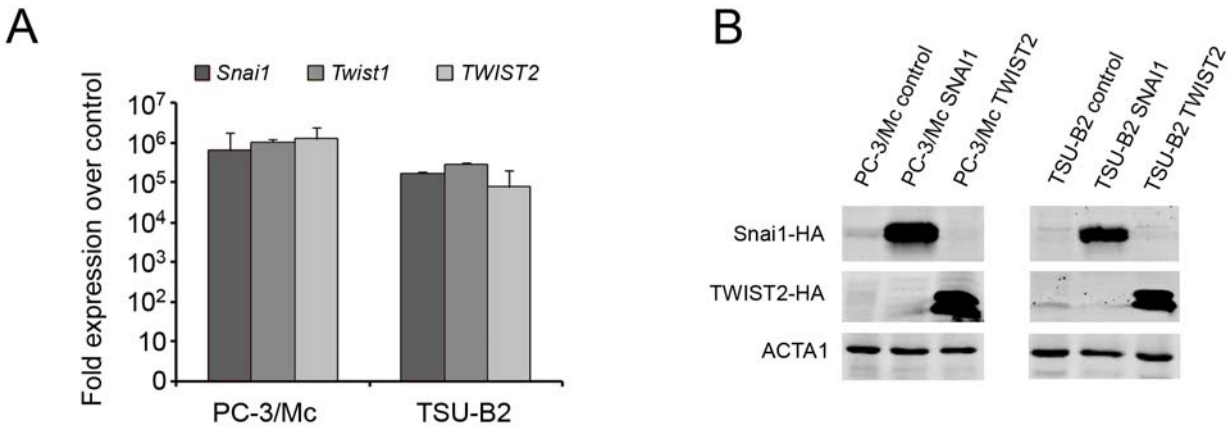


Supplemental Figure 5. Chromatin immunoprecipitation in PC-3/Mc cells showed enrichment for histone marks associated with active transcription at promoters of genes of self-renewal/pluripotency (*SOX2*, *KLF9*) and E-cadherin (*CDH1*), but not the mesenchymal genes *TWIST2* and *RUNX2*. Conversely, in PC-3/S cells, the promoters of *TWIST2* and *RUNX2* were enriched for histone marks associated with active transcription (acetyl-histone H3 and acetyl-histone H4), while the promoters of *SOX2*, *KLF9* and E-cadherin were enriched for the repressive histone mark H3K27me$_3$. The enrichment of promoter sequences after immunoprecipitation for specific histone marks was determined by real-time PCR with primers specific for each promoter, and values expressed as the percentage of the amplification values obtained for input DNA.

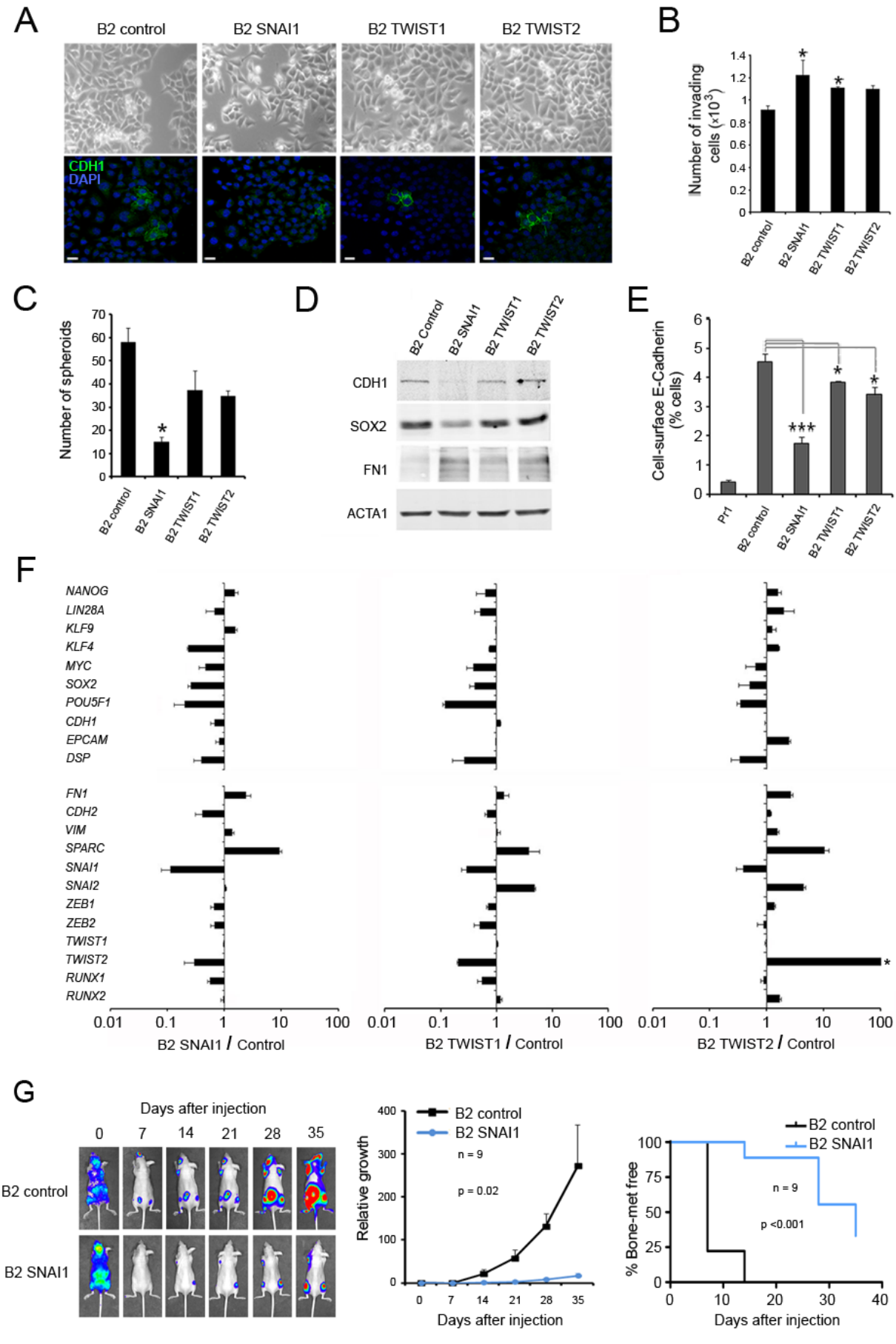


Supplemental Figure 6. Divergent phentoypes and gene programs of TSU-Pr1 *vs*. TSU-Pr1-B2 bladder cancer cell subpopulations. (**A**) TSU-Pr1-B2 cells expressed higher levels than TSU-Pr1 cells of epithelial (e.g., *CDH1*, *TACSTD1*, *DSP*) and self-renewal (e.g., *NANOG*, *LIN28*, *KLF9*, *SOX2*) genes, while TSU-Pr1 cells expressed higher levels than TSU-Pr1-B2 cells of mesenchymal genes (e.g., *FN1*, *SNAI2*, *ZEB1*, *ZEB2*, *TWIST1*). Transcripts were quantified by real-time RT-PCR and values, normalized to reference genes and expressed as the $\log_{10}$ of ratios of qPCR values between the two cell lines. (**B**) TSU-Pr1-B2 cells displayed an epithelioid morphology with a proportion of cells expressing membrane-associated E-cadherin, and TSU-Pr1 displayed a more fibroblastoid morphology, without detectable expression of E-cadherin by indirect immunofluorescence. TSU-Pr1-B2 cells also express higher levels of nuclear SOX2 than TSU-Pr-1 cells. Scale bar: 20 μm. (**C**) Comparative expression levels of epithelial, mesenchymal and self-renewal proteins in TSU-Pr1-B2 vs. TSU-Pr1 cells, determined by Western blotting. Actin signals were used as indicators of protein loading and transfer. (**D**) The more epithelial TSU-Pr1-B2 cells displayed a significantly stronger capacity to form spheroids than the more mesenchymal TSU-Pr1 cells. Cells (10^3) were seeded in low-attachment plates in medium containing 0.5% methyl cellulose.

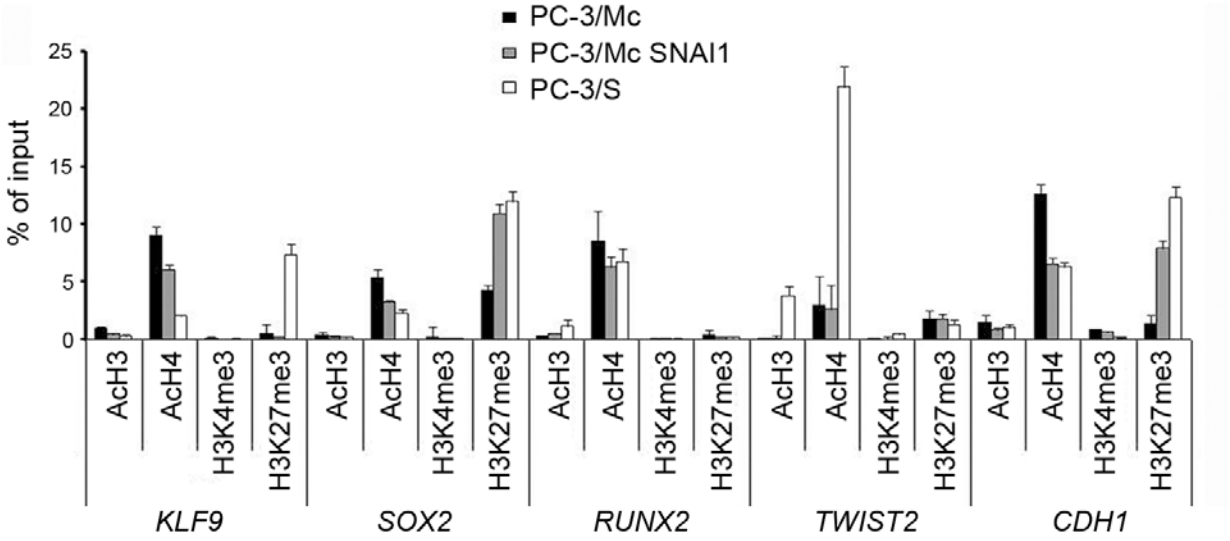
Three independent experiments were performed, each in triplicate. Student's *t*-test: * $p < 0.05$. (**E**) The more mesenchymal TSU-Pr1 cells were significantly more invasive than the more epithelial TSU-Pr1-B2 cells in Transwell-Matrigel assays. Three independent experiments were performed, each in triplicate. Student's *t*-test: * $p < 0.05$. (**F**) Upon intracardiac injection in Ncr-nu/nu mice, the more epithelial TSU-Pr1-B2 cells grow significantly faster at different sites (middle panel) and colonize bones much more efficiently (right panel) than the more mesenchymal TSU-Pr1 cells. Mice (n = 9) were injected i.c. with 2.0×10^5 cells bearing a stably integrated firefly luciferase gene. The total tumor burden (middle panel) was estimated as the sum of photon counts for each mouse. The Kaplan-Maier plot (right panel) reflects the number of animals free of detectable bone colonization at each time point.



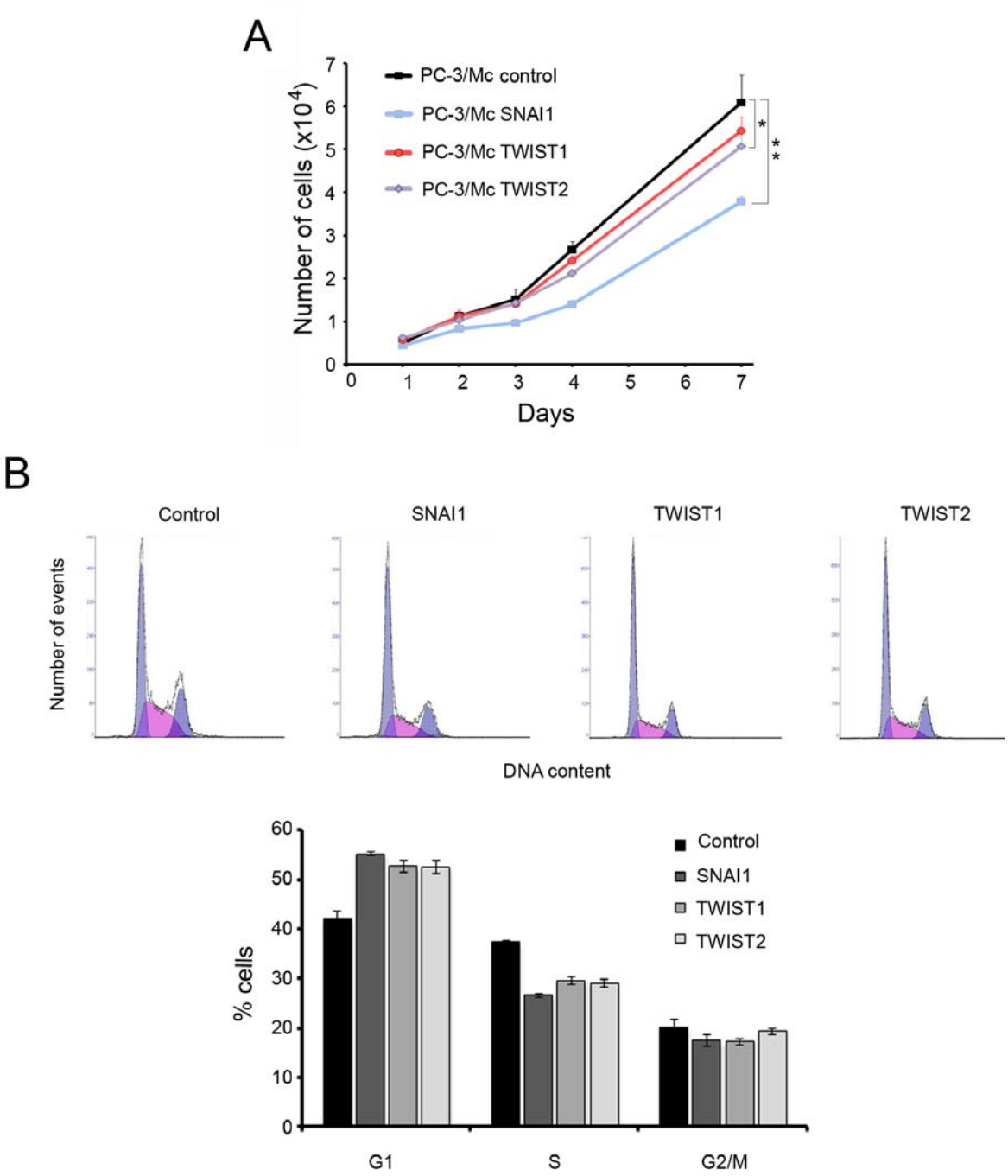
Supplemental Figure 7. Expression levels of EMT factors after retroviral transduction into PC-3/Mc and TSU-Pr1-B2 cells. (**A**) Determination by real-time RT-PCR of relative transcript levels for mouse *Snai1*, mouse *Twist1* and human *TWIST2* after retroviral transduction of the corresponding genes. Real-time Cp values were normalized to values for *RN18S1* (reference transcript) and represented as ratios between cells transduced with the experimental vectors and control cells transduced with the same empty vectors. Transcript levels for the transduced mouse *Snai1* and *Twist1* were quantified with primers and probes specific for the murine sequences. (**B**) Determination by Western blotting of protein levels of Snai1 and TWIST2 transduced into PC-3/Mc and TSU-Pr1-B2 cells, detected as proteins bearing the hemagglutinin (HA) epitope. Cells were transduced and processed as in (A). Actin signal was used as an indicator of protein loading and transfer.


A
B2 control
B2 SNAI1
B2 TWIST1
B2 TWIST2
CDH1
DAPI
B
Number of invading cells (x10^3)
C
Number of spheroids
D
B2 Control
B2 SNAI1
B2 TWIST1
B2 TWIST2
CDH1
SOX2
FN1
ACTA1
E
Cell-surface E-Cadherin (% cells)
Pr1
F
NANOG
LIN28A
KLF9
KLF4
MYC
SOX2
POU5F1
CDH1
EPCAM
DSP
FN1
CDH2
VIM
SPARC
SNAI1
SNAI2
ZEB1
ZEB2
TWIST1
TWIST2
RUNX1
RUNX2
B2 SNAI1 / Control
B2 TWIST1 / Control
B2 TWIST2 / Control
G
Days after injection
0
7
14
21
28
35
B2 control
B2 SNAI1
Relative growth
n = 9
p = 0.02
% Bone-met free
p <0.001

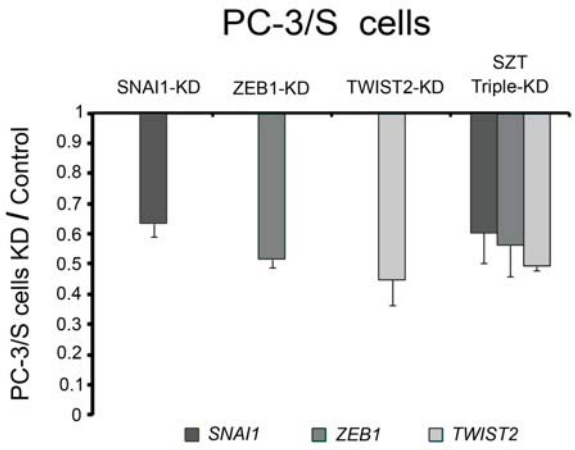
Supplemental Figure 8. Inhibition of TIC attributes of TSU-Pr1-B2 bladder cancer cells by constitutive overexpression of *Snai1*. (**A**) Overexpression of *Snai1* in TSU-Pr1-B2 cells induced a more fibroblastoid morphology and a decreased expression of membrane-associated E-cadherin. Controls were TSU-Pr1-B2 cells transduced with empty pBABE retroviral vector and selected for puromycin resistance. Scale bar: 20 μm. (**B**) Overexpression of *Snai1* in TSU-Pr1-B2 cells induced a significant enhancement of invasiveness in Transwell-Matrigel assays. Overexpression of *Twist1* or *TWIST2* caused a more modest effect. (**C**) Overexpression of *Snai1* in TSU-Pr1-B2 cells significantly inhibited their capacity to form spheroids in anchorage-independent growth conditions. Overexpression of *Twist1* or *TWIST2* caused a more modest effect. (**D**) Overexpression of *Snai1* in TSU-Pr1-B2 cells induced a downregulation of E-cadherin and SOX2 and an upregulation of fibronectin, as determined by Western blotting. Actin signal was used as an indicator of protein loading and transfer. (**E**) Overexpression of *Snai1* in TSU-Pr1-B2 cells caused a downregulation of cell-surface E-cadherin, as determined by flow cytometry. Overexpression of *Twist1* and *TWIST2* caused a more modest effect. (**F**) Overexpression of exogenous *Snai1* caused a downregulation of epithelial (*CDH1*, *DSP*) and of self-renewal (*KLF4*, *MYC*, *SOX2*, *POU5F1*) genes, along with an upregulation of mesenchymal genes (*FN1*, *SPARC*). The endogenous human *SNAI1* gene was strongly downregulated upon overexpression of exogenous mouse *Snai1*, as expected from its known auto-inhibitory feedback loop (18). Overexpression of *Twist1* and *TWIST2* did not significantly affect the transcript levels of E-cadherin (*CDH1*), but induced changes in the levels of transcripts for some self-renewal and mesenchymal genes. The levels of human *TWIST2* in the cells transduced with this EMT factor are off-scale. See Supplemental Figure 7 for quantification of the transduced exogenous murine *Snai1*, and murine *Twist1*. Real-time Cp values were normalized to values for *RN18S1* (reference transcript) and represented as ratios between cells transduced with the experimental vectors and control cells transduced with the same empty vector (pBABE). (**G**) Overexpression of *Snai1* strongly suppressed the growth of TSU-Pr1-B2 cells in different organs (middle panel) and their bone colonization (right panel) after intracardiac injection. Ncr-nu/nu mice (n = 9) were injected i.c. with 2.0 x 10^5 cells bearing a stably integrated firefly luciferase gene. The total tumor burden (middle panel) was estimated as the sum of photon counts for each mouse. The Kaplan-Maier plot (right panel) reflects the number of animals free of detectable bone colonization at each time point. For graphs with statistical analysis (Student's *t*-test): * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.



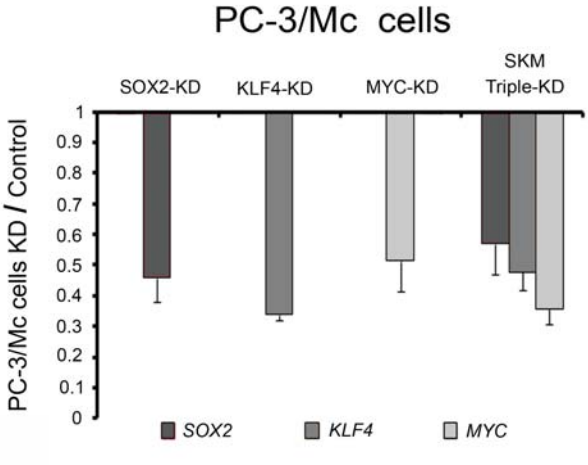
Supplemental Figure 9. Transduction and overexpression of *Snai1* in PC-3/Mc cells induced changes in histone marks associated with the promoters of *SOX2* and E-cadherin (*CDH1*), notably an enrichment of histone $H3K27_{me3}$ and depletion of acetylated histone H4, as analyzed by chromatin immunoprecipitation. For comparison, PC-3/S cells also had enriched repressive histone marks at the promoters of *SOX2* and E-cadherin.



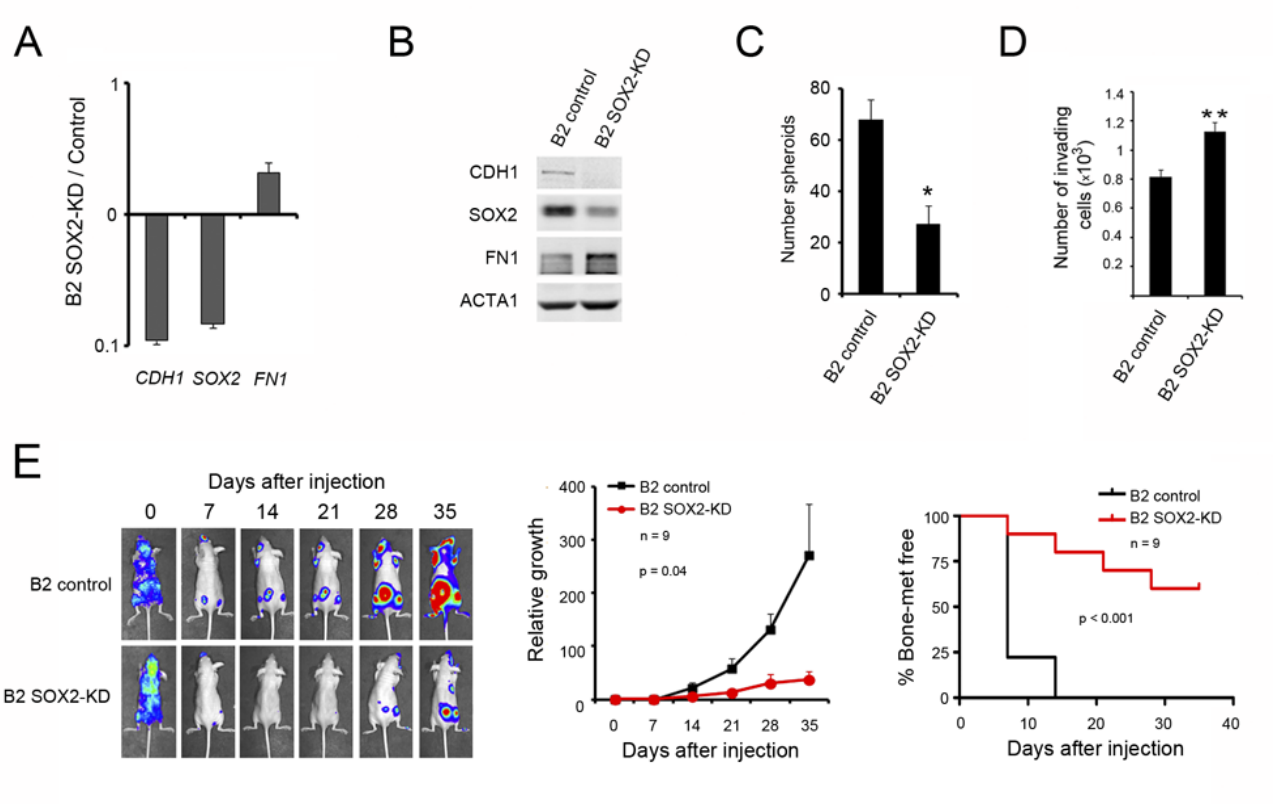
Supplemental Figure 10. Effects of the transduction and overexpression of the EMT factors *Snai1*, *Twist1* and *TWIST2* on the cell cycle profile of PC-3/Mc cells. (**A**) Overexpression of *Snai1*, *Twist1* or *TWIST2* caused a reduced growth rate on plastic, the strongest effect caused by *Snai1* overexpression. Cells (5 x 10^3) were seeded in triplicate in 96-well plates and their numbers determined 1, 2, 3, 4 and 7 days after plating. * $p < 0.05$, ** $p < 0.01$ (Student's *t*-test). (**B**) Overexpression of *Snai1*, *Twist1* or *TWIST2* caused a reduction in the proportion of cells in the S phase of the cell cycle with a concomitant increase in cells in the G phase, with the strongest effects caused by *Snai1* overexpression. Upper panels, representative DNA content histograms from flow cytometry determinations. Lower panel, histograms illustrating the quantification of cell populations in G1, S and G2-M. Experiments were done in triplicate.



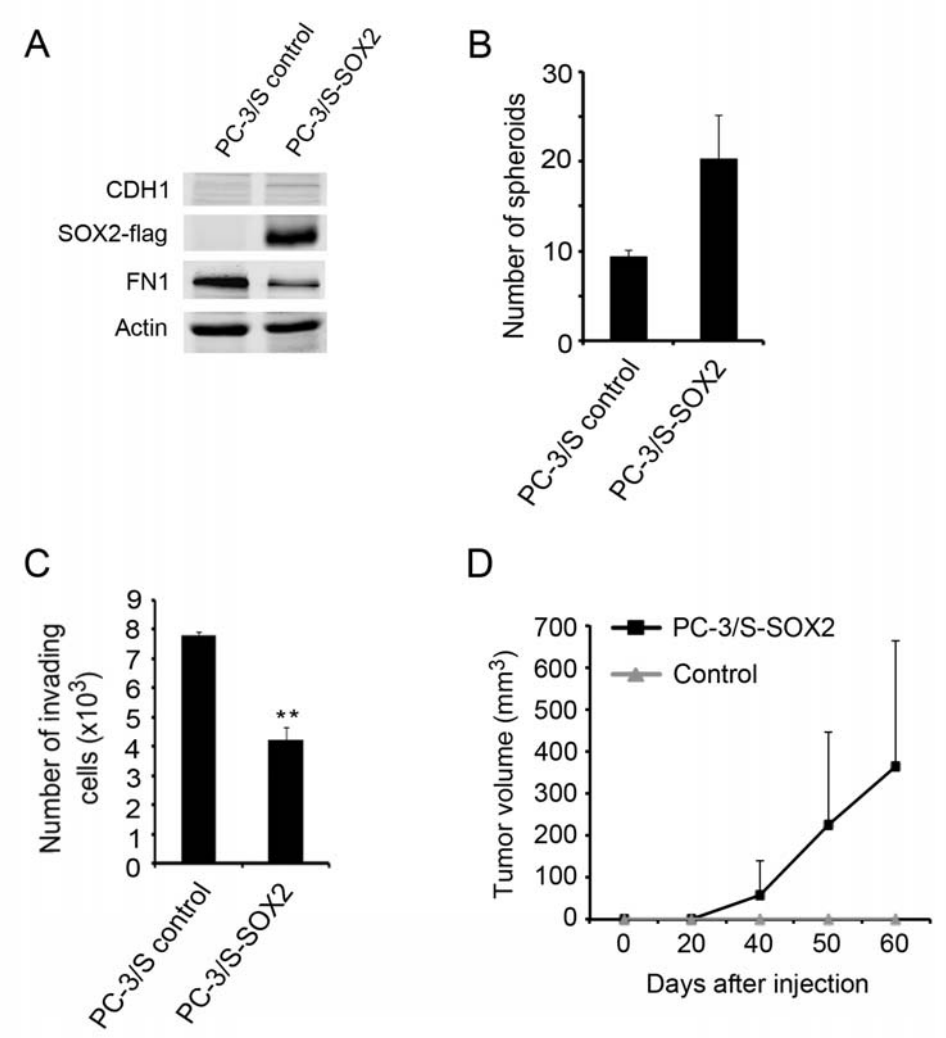
Supplemental Figure 11. Real-time RT-PCR quantification of *SNAI1*, *ZEB1* and *TWIST2* transcripts levels in PC-3/S cells knocked down for these EMT factors or a triple knockdown (*SNAI1*, *ZEB1* and *TWIST2*, SZT triple KD). Cells were transduced with lentiviral particles carrying shRNAs specific for each transcript, or for all three transcripts, selected for puromycin resistance and processed. Reference gene-normalized Cp values are represented as the ratio between experimental and control cells. Controls were PC-3/S cells transduced with a control pLK0-scrambled vector and puromycin selected. Means and standard deviations are from triplicate experiments.



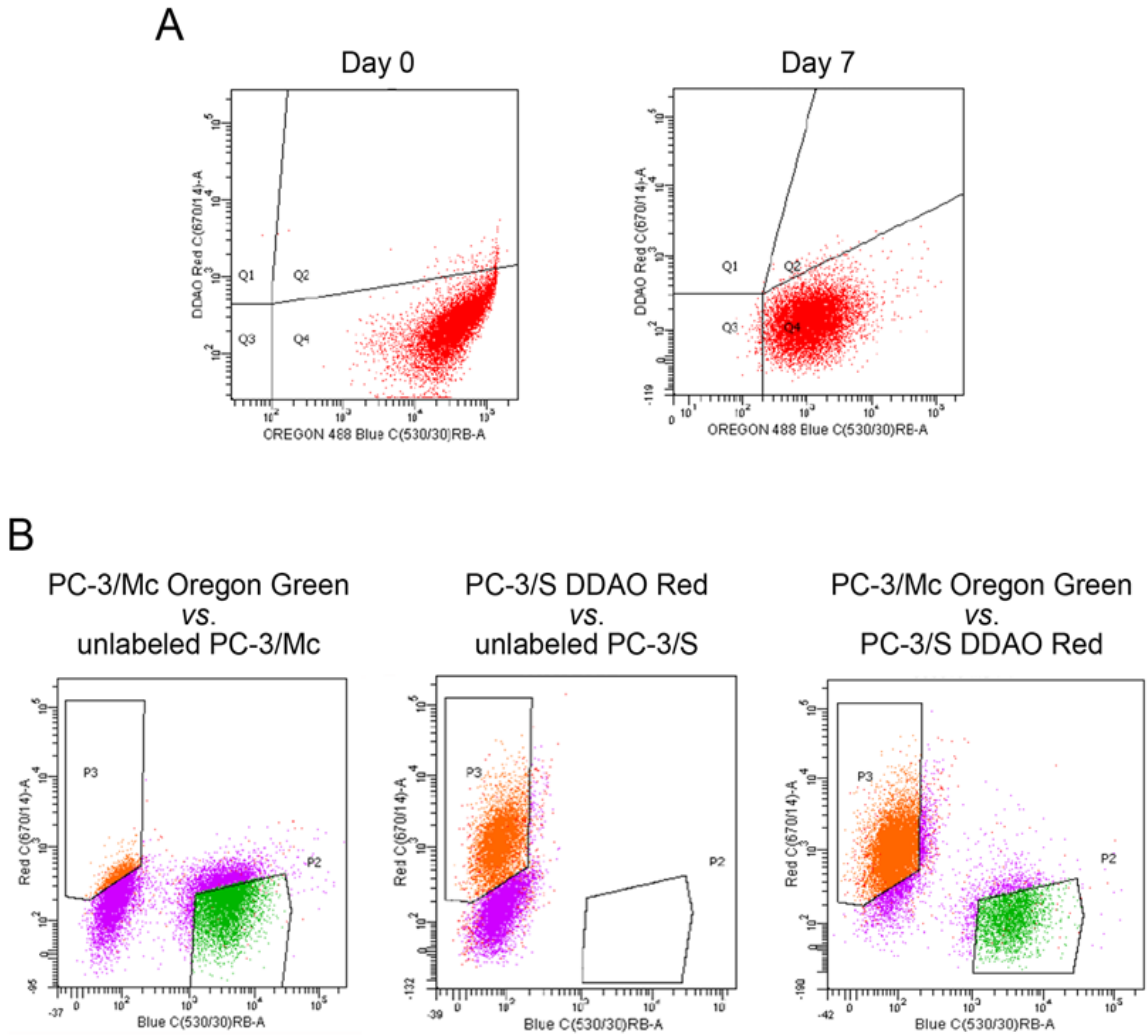
Supplemental Figure 12. Real-time RT-PCR quantification of *SOX2*, *KLF4* and *MYC* transcript levels in PC-3/Mc cells knocked down for these self-renewal/pluripotency factors or a triple knockdown (*SOX2*, *KLF4* and *MYC*, SKM triple KD). Cells were transduced with lentiviral particles carrying shRNAs specific for each transcript, or for all three transcripts, selected for puromycin resistance and processed. Reference gene-normalized Cp values are represented as the ratio between experimental and control cells. Controls were PC-3/S cells transduced with a control pLK0-scrambled vector and puromycin selected. Means and standard deviations are from triplicate experiments.



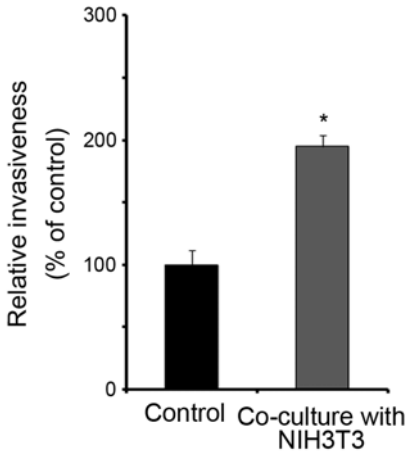
Supplemental Figure 13. Self-renewal transcription factors are required to maintain a strong epithelial program, anchorage-independent growth and metastatic potential of TSU-Pr1-B2 cells. (**A**) Knockdown of *SOX2* in TSU-Pr1-B2 cells caused a downregulation of E-cadherin (*CDH1*) and an upregulation of fibronectin (*FN1*), as determined by real-time RT-PCR. (**B**) Knockdown of *SOX2* in TSU-Pr1-B2 cells caused a downregulation of E-cadherin and an upregulation of fibronectin, as determined by Western blotting. (**C**) Knockdown of *SOX2* in TSU-Pr1-B2 cells inhibited the formation of spheroids in anchorage-independent growth conditions. Cells (10^3) were seeded in low-attachment plates in the presence of 0.5% methyl cellulose. Assays were performed in triplicate. (**D**) Knockdown of *SOX2* in TSU-Pr1-B2 cells enhanced their invasiveness in Transwell-Matrigel assays. Assays were performed in triplicate. (**E**) Knockdown of *SOX2* in TSU-Pr1-B2 cells inhibited the growth in different organs (middle panel) and bone colonization (right panel) after intracardiac injection. Ncr-nu/nu mice (n = 9) were injected i.c. with 2.5 x 10^5 cells bearing a stably integrated firefly luciferase gene. The total tumor burden (middle panel) was estimated as the sum of photon counts for each mouse. The Kaplan-Maier plot (right panel) reflects the number of animals free of detectable bone colonization at each time point. For graphs with statistical analysis (Student's *t*-test): * $p < 0.05$; ** $p < 0.01$.



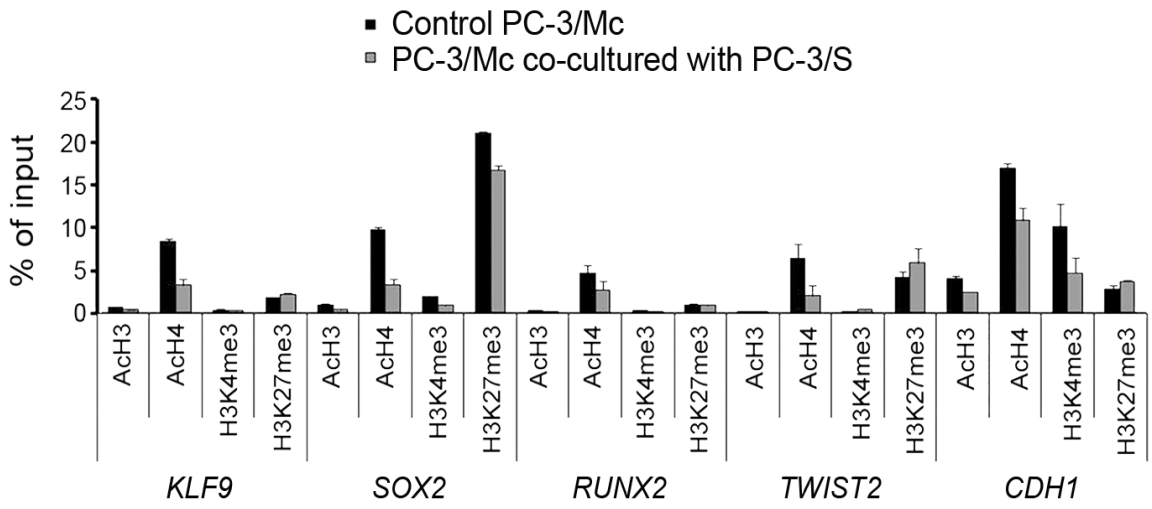
Supplemental Figure 14. Transduction and overexpression of *SOX2* in the mesenchymal-like PC-3/S cells enhances their anchorage-independent growth and tumorigenicity. (**A**) Overexpression of *SOX2* in PC-3/S cells caused an upregulation of E-cadherin and a downregulation of fibronectin, as determined by Western blotting. Controls were PC-3/S cells transduced with retroviral particles containing the empty pBABE-puro vector and selected for puromycin resistance. (**B**) Overexpression of *SOX2* in PC-3/S cells caused a gain in their capacity to form spheroids in low-attachment plates in the presence of 0.5% methyl cellulose. Assays were done in triplicate. (**C**) Overexpression of *SOX2* in PC-3/S cells inhibited their invasiveness, as determined in Tranwell-Matrigel assays. Assays were done in triplicate. Student's *t*-test: ** $p < 0.01$. (**D**) Overexpression of *SOX2* strongly enhanced the tumorigenicity of PC-3/S cells. Cells (5×10^5) were implanted i.m. in the hind limbs of male Swiss-nude mice and tumor growth monitored with a caliper (n = 4).



Supplemental Figure 15. Fluorophore loading and separation by FACS in cell co-culture experiments. (**A**) Loading with fluorophores is maintained for a prolonged period of time. PC-3/Mc cells were loaded with Oregon Green and examined for fluorescent levels on the same day of the loading and after 7 days in culture. (**B**) Loading, co-culture and FACS separation of PC-3/Mc cells (loaded with Oregon Green) and PC-3/S cells (loaded with DDAO Red). After loading with their respective fluorophores, PC-3/Mc and PC-3/S cells were either co-cultured or not for 48 h, and submitted to two-channel FACS (488 nm and 670 nm). The selection windows were designed so as to avoid any overlapping between green- and red-labeled cells.



Supplemental Figure 16. Co-culture with NIH3T3 mouse fibroblasts enhanced the invasiveness of PC-3/Mc cells in Transwell-Matrigel assays. Equal numbers of PC-3/Mc cells pre-loaded with Oregon Green and NIH3T3 cells pre-loaded with DDAO Red were seeded together and assayed for invasiveness on Transwell-Matrigel assays. After 24 h, cells that had invaded into the lower chamber were quantified by flow cytometry set at 488 nm (for green fluorescent cells) and 670 nm (for red fluorescent cells). Experiments were done in triplicate. Student's *t*-test: * $p < 0.05$.



Supplemental Figure 17. Co-culture of PC-3/Mc cells with PC-3/S cells induces changes in histone marks associated with the promoters of *SOX2* and E-cadherin (*CDH1*). Most notably, the *SOX2* promoter in PC-3/Mc cells was depleted of the active transcription mark acetylated histone H4 and the *CDH1* promoter was depleted of the active transcription marks acetylated histones H3 and H4 and histone $H3K4_{me3}$. PC-3/Mc cells pre-loaded with Oregon Green and PC-3/S cells pre-loaded with DDAO Red were co-cultured for 48 h, separated by FACS and PC-3/Mc cells analyzed for histone marks by chromatin immunoprecipitation with specific antibodies, and enrichment of promoter sequences quantified by real-time PCR. Results are expressed as specific amplification levels on immunoprecipitated DNA relative to the amplification levels yielded by input DNA. All determinations were done in triplicate.

Annex data tables

Supplemental Table 1

GENES DIFFERENTIALLY EXPRESSED IN PC-3/Mc vs. PC-3/S

SAM q-value = 0
3 ≤ Log2[M/S] ≤ -3

Affy ID	Gene symbol	Log2[M/S]
1554766_s_at	PVT1	8,311991081
201387_s_at	UCHL1	7,862055964
209771_x_at	CD24	6,984133159
226517_at	BCAT1	6,982382677
216379_x_at	CD24	6,793668238
208650_s_at	CD24	6,331612482
200606_at	DSP	6,316180697
266_s_at	CD24	6,126306155
225285_at	BCAT1	6,094628611
243175_at	UTS2D	5,876860856
219778_at	ZFPM2	5,837800488
214240_at	GAL	5,743542351
232381_s_at	DNAH5	5,619312455
224588_at	XIST	5,518296231
214452_at	BCAT1	5,468429783
201839_s_at	TACSTD1	5,431873579
210295_at	MAGEA10	5,392731908
225846_at	RBM35A	5,356507871
206172_at	IL13RA2	5,341960689
228010_at	PPP2R2C	5,328039033
215733_x_at	CTAG2	5,250179931
225645_at	EHF	5,249531814
213913_s_at	KIAA0984	5,244769821
230896_at	CCDC4	5,13608918
213880_at	LGR5	5,135330894
218332_at	BEX1	5,110053856
212328_at	LIMCH1	5,096526849
206953_s_at	LPHN2	5,052356292
219985_at	HS3ST3A1	5,030347099
225792_at	HOOK1	5,012996142
230323_s_at	TMEM45B	4,992302095
203917_at	CXADR	4,984719781
218035_s_at	FLJ20273	4,981733089
206026_s_at	TNFAIP6	4,94663762
1553798_a_at	FBXL13	4,905739191
203130_s_at	KIF5C	4,887006599
206070_s_at	EPHA3	4,861404056
1560527_at	NA	4,846105101
207016_s_at	ALDH1A2	4,799382107
205625_s_at	CALB1	4,790273864
227623_at	NA	4,777184128
206291_at	NTS	4,745025676
219983_at	HRASLS	4,706485069
219263_at	RNF128	4,62030812
209772_s_at	CD24	4,581071293
227929_at	NA	4,555171049
205626_s_at	CALB1	4,510964225
237449_at	SP8	4,505631335
232361_s_at	EHF	4,503556962
221731_x_at	VCAN	4,496968429
213912_at	KIAA0984	4,480457419
220062_s_at	MAGEC2	4,454356735
239370_at	NA	4,449726822
222496_s_at	FLJ20273	4,428198224
1554614_a_at	PTBP2	4,42431084
206025_s_at	TNFAIP6	4,39505039
203824_at	TSPAN8	4,388327204
208651_x_at	CD24	4,373184147
204466_s_at	SNCA	4,332987481
219734_at	SIDT1	4,330307599
221728_x_at	XIST	4,324020632
219850_s_at	EHF	4,305978709
206204_at	GRB14	4,29987495

219984_s_at	HRASLS	4,28898763
204620_s_at	VCAN	4,272671604
202437_s_at	CYP1B1	4,265753166
202202_s_at	LAMA4	4,255784733
213134_x_at	BTG3	4,229241011
1553830_s_at	MAGEA2	4,214742967
231930_at	ELMOD1	4,177956925
227671_at	XIST	4,163252405
205548_s_at	BTG3	4,149561546
1553972_a_at	CBS	4,145829512
202403_s_at	COL1A2	4,114548734
212325_at	LIMCH1	4,100161622
214603_at	MAGEA2B	4,094605795
236442_at	DPF3	4,063765316
219121_s_at	RBM35A	4,051521306
202436_s_at	CYP1B1	4,040359243
202404_s_at	COL1A2	4,03991138
205523_at	HAPLN1	4,008732643
210229_s_at	CSF2	4,005193049
206884_s_at	SCEL	4,001694195
219789_at	NPR3	3,9553395
215646_s_at	VCAN	3,948003698
212327_at	LIMCH1	3,936091004
225911_at	NPNT	3,934779102
229139_at	JPH1	3,933648662
226226_at	TMEM45B	3,922233066
206440_at	LIN7A	3,899979565
230272_at	LOC645323	3,892959119
205899_at	CCNA1	3,86406436
211674_x_at	CTAG1B	3,843730366
228737_at	TOX2	3,84228484
201998_at	ST6GAL1	3,794980839
217764_s_at	RAB31	3,7479559
227230_s_at	KIAA1211	3,74513122
230204_at	NA	3,743343413
1558871_at	EPGN	3,737593478
213056_at	FRMD4B	3,723116714
1552848_a_at	PTCHD1	3,7230457
201131_s_at	CDH1	3,681358166
224189_x_at	EHF	3,662791324
1552897_a_at	KCNG3	3,653199913
203021_at	SLPI	3,630676084
214218_s_at	XIST	3,627625536
244623_at	NA	3,626490673
205421_at	SLC22A3	3,616730825
229800_at	DCLK1	3,614391757
217763_s_at	RAB31	3,603672763
243802_at	DNHD2	3,603093858
236129_at	GALNT5	3,560803668
227176_at	SLC2A13	3,558385828
211571_s_at	VCAN	3,528533094
225809_at	DKFZP564O0823	3,509292108
238850_at	LOC645323	3,509120885
227492_at	LOC647859	3,503368128
204798_at	MYB	3,501196933
213050_at	COBL	3,475548471
204464_s_at	EDNRA	3,469816845
212816_s_at	CBS	3,467013743
206002_at	GPR64	3,44934852
223423_at	GPR160	3,437004411
230895_at	NA	3,434099389
223075_s_at	C9orf58	3,401746215
217762_s_at	RAB31	3,39814416
243110_x_at	NPW	3,386192897
218683_at	PTBP2	3,338924978
210546_x_at	CTAG1B	3,334577718
234472_at	GALNT13	3,323838705
218839_at	HEY1	3,316067153
218829_s_at	CHD7	3,31096006
213285_at	TMEM30B	3,303108778
209170_s_at	GPM6B	3,27262105
217339_x_at	CTAG1B	3,264288917
230085_at	PCYT1B	3,248737949
227719_at	NA	3,248029898
226374_at	CXADR	3,237354865

227688_at	LRCH2	3,221780344
207076_s_at	ASS1	3,201798376
219355_at	CXorf57	3,182120994
203680_at	PRKAR2B	3,17330932
237086_at	FOXA1	3,15363767
212942_s_at	KIAA1199	3,14892415
219572_at	CADPS2	3,140087394
224589_at	XIST	3,132494879
44783_s_at	HEY1	3,119046683
235144_at	NA	3,105100974
204351_at	S100P	3,094863197
214930_at	SLITRK5	3,0935354
224856_at	FKBP5	3,071320377
216604_s_at	SLC7A8	3,050043046
226459_at	PIK3AP1	3,0455874
205524_s_at	HAPLN1	3,044051061
214390_s_at	BCAT1	3,033515227
232760_at	TEX15	3,030989401
240027_at	LIN7A	3,023217009
213107_at	TNIK	3,021476122
204720_s_at	DNAJC6	3,001401032
219855_at	NUDT11	-3,038217609
244406_at	ZNF20	-3,039596806
1557599_a_at	RELN	-3,041093169
226702_at	LOC129607	-3,05565616
209098_s_at	JAG1	-3,057502174
219368_at	NAP1L2	-3,057788034
212845_at	SAMD4A	-3,057893463
241780_at	CDH11	-3,063898622
219848_s_at	ZNF432	-3,075123891
206240_s_at	ZNF136	-3,075973608
212977_at	CXCR7	-3,080546593
222449_at	TMEPAI	-3,094704718
241827_at	ZNF615	-3,102022577
211796_s_at	TRBC1	-3,106124657
202363_at	SPOCK1	-3,116798879
221261_x_at	MAGED4B	-3,125223904
216869_at	PDE1C	-3,126881652
226425_at	CLIP4	-3,134762136
216268_s_at	JAG1	-3,139698855
230258_at	GLIS3	-3,144742382
214995_s_at	APOBEC3F	-3,157468404
201649_at	UBE2L6	-3,158644224
208370_s_at	RCAN1	-3,168189118
211343_s_at	COL13A1	-3,17638103
223618_at	FMN2	-3,198932033
228442_at	NA	-3,203380396
219489_s_at	RHBDL2	-3,208328491
231098_at	NA	-3,211924298
209457_at	DUSP5	-3,216047831
226757_at	IFIT2	-3,240029704
202842_s_at	DNAJB9	-3,24812455
223784_at	TMEM27	-3,263717808
1553995_a_at	NT5E	-3,266102256
214476_at	TFF2	-3,268446163
222802_at	EDN1	-3,276644686
234331_s_at	FAM84A	-3,279458691
204926_at	INHBA	-3,282536752
206969_at	KRT34	-3,286132528
1555471_a_at	FMN2	-3,298523917
226743_at	SLFN11	-3,302536139
204160_s_at	ENPP4	-3,304864443
203939_at	NT5E	-3,305857315
205207_at	IL6	-3,314996931
223313_s_at	MAGED4B	-3,320205243
225481_at	FRMD6	-3,324919156
228462_at	IRX2	-3,3309018
207417_s_at	ZNF177	-3,334050359
218312_s_at	ZSCAN18	-3,337274278
226612_at	FLJ25076	-3,343618127
229576_s_at	TBX3	-3,344715434
226731_at	PELO	-3,353220251
206825_at	OXTR	-3,361837933
225464_at	FRMD6	-3,369765496
206569_at	IL24	-3,375442607

220987_s_at	NUAK2	-3,392669654
205514_at	ZNF415	-3,401980796
214822_at	FAM5B	-3,402268494
205081_at	CRIP1	-3,430044068
227045_at	ZNF614	-3,431838815
225273_at	WWC3	-3,438473586
242761_s_at	ZNF420	-3,444732991
225667_s_at	FAM84A	-3,447978772
218573_at	MAGEH1	-3,453911006
206858_s_at	HOXC6	-3,460920727
215253_s_at	RCAN1	-3,462670698
209909_s_at	TGFB2	-3,465691579
244740_at	MGC9913	-3,468443605
202765_s_at	FBN1	-3,469101927
236635_at	ZNF667	-3,516602563
226869_at	MEGF6	-3,578465117
236201_at	NA	-3,586502433
205266_at	LIF	-3,587359748
227145_at	LOXL4	-3,601296013
203066_at	GALNAC4S-6ST	-3,651997951
206300_s_at	PTHLH	-3,657063715
231725_at	PCDHB2	-3,66157221
202766_s_at	FBN1	-3,664569094
238725_at	NA	-3,669798908
230746_s_at	STC1	-3,673908206
204684_at	NPTX1	-3,676616313
223170_at	TMEM98	-3,676618984
226991_at	NFATC2	-3,678734206
229963_at	NGFRAP1L1	-3,694620907
235648_at	ZNF567	-3,715397541
204279_at	PSMB9	-3,73148778
227847_at	EPM2AIP1	-3,736797262
239218_at	PDE1C	-3,781149397
214022_s_at	IFITM2	-3,788882568
1558700_s_at	ZNF260	-3,796376784
202843_at	DNAJB9	-3,808833931
223714_at	ZNF256	-3,815998204
202086_at	MX1	-3,824862321
203908_at	SLC4A4	-3,831921847
202149_at	NEDD9	-3,841831581
204719_at	ABCA8	-3,853257099
212372_at	MYH10	-3,855525131
202627_s_at	SERPINE1	-3,877193173
202957_at	HCLS1	-3,879744792
205286_at	TFAP2C	-3,892929847
206584_at	LY96	-3,912360467
228988_at	ZNF711	-3,914556264
209156_s_at	COL6A2	-3,927604347
238532_at	DPF3	-3,93276397
228920_at	ZNF260	-3,941641348
239671_at	NA	-3,945439897
210538_s_at	BIRC3	-3,951751811
209074_s_at	FAM107A	-3,968429087
202628_s_at	SERPINE1	-3,976310617
219682_s_at	TBX3	-4,012223008
1560562_a_at	ZNF677	-4,015611967
201438_at	COL6A3	-4,085604541
209101_at	CTGF	-4,089598239
204205_at	APOBEC3G	-4,096899598
232231_at	RUNX2	-4,175614256
228805_at	C5orf25	-4,190325493
224819_at	TCEAL8	-4,198164345
219288_at	C3orf14	-4,203733392
202909_at	EPM2AIP1	-4,231887602
229441_at	PRSS23	-4,240821945
1553605_a_at	ABCA13	-4,244254676
201445_at	CNN3	-4,283025812
205498_at	GHR	-4,294003993
228004_at	C20orf56	-4,316195104
211959_at	IGFBP5	-4,320354564
211756_at	PTHLH	-4,346785265
236344_at	NA	-4,354925892
206421_s_at	SERPINB7	-4,392492645
202888_s_at	ANPEP	-4,396860418
200665_s_at	SPARC	-4,406643686

210560_at	GBX2	-4,422967501
221024_s_at	SLC2A10	-4,429365873
226109_at	C21orf91	-4,452426882
230030_at	HS6ST2	-4,489178786
218804_at	TMEM16A	-4,506509845
228974_at	ZNF677	-4,513907781
204749_at	NAP1L3	-4,541757619
1552767_a_at	HS6ST2	-4,558476491
228297_at	CNN3	-4,571314977
203315_at	NCK2	-4,583988571
209656_s_at	TMEM47	-4,605089222
201289_at	CYR61	-4,605342583
218309_at	CAMK2N1	-4,61161678
202859_x_at	IL8	-4,636088656
206336_at	CXCL6	-4,660168846
219140_s_at	RBP4	-4,66676543
209990_s_at	GABBR2	-4,739295308
229435_at	GLIS3	-4,739929779
203373_at	SOCS2	-4,792174341
228176_at	C9orf47	-5,0240463
1555673_at	KRTAP2-1	-5,069577277
210764_s_at	CYR61	-5,076277019
222444_at	ARMCX3	-5,081191671
244741_s_at	MGC9913	-5,115775288
202458_at	PRSS23	-5,1439897
227752_at	SPTLC3	-5,239789728
203372_s_at	SOCS2	-5,333767267
206157_at	PTX3	-5,357600958
213711_at	KRT81	-5,416639318
211506_s_at	IL8	-5,58929713
207526_s_at	IL1RL1	-5,66468703
226279_at	PRSS23	-5,689132329
213122_at	TSPYL5	-5,722694481
217858_s_at	ARMCX3	-6,037459788
205258_at	INHBB	-6,156621741
210495_x_at	FN1	-6,286068735
216442_x_at	FN1	-6,295885648
201859_at	SRGN	-6,40660428
205681_at	BCL2A1	-6,436395023
202391_at	BASP1	-6,456138853
231265_at	COX7B2	-6,59185963
212464_s_at	FN1	-6,775782597
212143_s_at	IGFBP3	-7,020526183
211719_x_at	FN1	-7,151036736
210095_s_at	IGFBP3	-7,191149715
227140_at	NA	-7,220756805
201858_s_at	SRGN	-8,042725632
210511_s_at	INHBA	-8,059804356

Supplemental Table 2

GSEA SUMMARY
UPREGULATED IN PC-3/Mc (FDR q-val ≤ 0.1)

NAME	SIZE	NOM p-val	FDR q-val
IDX_TSA_UP_CLUSTER3	88	0	0
SERUM_FIBROBLAST_CELLCYCLE	136	0	0
MANALO_HYPOXIA_DN	78	0	0
DOX_RESIST_GASTRIC_UP	44	0	0
CMV_IE86_UP	49	0	0
DNA_REPLICATION_REACTOME	44	0	0
TESTIS_EXPRESSED_GENES	61	0	0
CELL_CYCLE	78	0	0
LI_FETAL_VS_WT_KIDNEY_DN	159	0	0
AGUIRRE_PANCREAS_CHR12	59	0	0
LEE_TCELLS3_UP	106	0	0,001356417
CROONQUIST_IL6_STARVE_UP	33	0	0,002807356
TARTE_PLASMA_BLASTIC	307	0	0,002591406
PENG_GLUTAMINE_DN	248	0	0,002689092
GOLDRATH_CELLCYCLE	28	0	0,003002351
YU_CMYC_UP	30	0	0,003995809

VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	36	0	0,004204131
SCHUMACHER_MYC_UP	51	0	0,004792422
CELL_CYCLE_KEGG	86	0	0,00592988
BETA_ALANINE_METABOLISM	27	0	0,006016438
LYSINE_DEGRADATION	30	0	0,007883537
MOREAUX_TACI_HI_IN_PPC_UP	73	0	0,008568765
PROPANOATE_METABOLISM	31	0	0,010492585
CANCER_UNDIFFERENTIATED_META_UP	67	0	0,010216346
HESS_HOXAANMEIS1_DN	60	0	0,011774324
HESS_HOXAANMEIS1_UP	60	0	0,012654977
PENG_RAPAMYCIN_DN	188	0	0,012186275
SERUM_FIBROBLAST_CORE_UP	202	0	0,012689641
BUTANOATE_METABOLISM	27	0	0,013537449
GREENBAUM_E2A_UP	33	0	0,014602549
MIDDLEAGE_DN	14	0	0,018159244
MITOCHONDRIAL_FATTY_ACID_BETAOXIDATION	16	0,015267176	0,033398524
BHATTACHARYA_ESC_UP	62	0	0,038976043
PANTOTHENATE_AND_COA_BIOSYNTHESIS	12	0	0,042713948
ADIP_DIFF_CLUSTER5	39	0	0,043301117
PENG_LEUCINE_DN	139	0	0,04305184
SA_REG_CASCADE_OF_CYCLIN_EXPR	12	0	0,04332542
CHANG_SERUM_RESPONSE_UP	148	0	0,044560842
G1_TO_S_CELL_CYCLE_REACTOME	68	0,00877193	0,05133676
BLEO_MOUSE_LYMPH_HIGH_24HRS_DN	34	0	0,053444702
COLLER_MYC_UP	18	0,007751938	0,05623854
BRCA1_OVEREXP_DN	110	0	0,062010065
GPCRDB_CLASS_B_SECRETIN_LIKE	23	0,008695652	0,06136114
PRMT5_KD_UP	183	0	0,06183698
ZHAN_MULTIPLE_MYELOMA_VS_NORMAL_DN	39	0	0,06054734
ADIP_DIFF_CLUSTER4	34	0,01923077	0,05955304
ERM_KO_TESTES_DN	21	0	0,06236282
VERNELL_PRB_CLSTR1	67	0	0,06305951
PARP_KO_DN	13	0,016949153	0,06185107
CMV_HCMV_TIMECOURSE_48HRS_UP	66	0	0,0613558
MITOCHONDRIA	391	0	0,060373165
BRENTANI_CELL_CYCLE	81	0	0,060155127
GPCRS_CLASS_B_SECRETIN_LIKE	23	0,008196721	0,05923299
CANTHARIDIN_DN	50	0,024193548	0,066550404
5FU_RESIST_GASTRIC_DN	16	0,008130081	0,066701874
MRNA_PROCESSING_REACTOME	108	0	0,07201631
E2F1_DNA_UP	11	0,033613447	0,071073286
HDACI_COLON_CLUSTER4	16	0,007936508	0,074437924
PYRUVATE_METABOLISM	37	0,017391304	0,0812955
ZHAN_MM_CD1_VS_CD2_DN	54	0,0078125	0,08062623
AGUIRRE_PANCREAS_CHR6	32	0,008695652	0,0799248
SHEPARD_GENES_COMMON_BW_CB_MO	68	0	0,08379386
MUNSHI_MM_VS_PCS_DN	11	0,022900764	0,084383786
CANCERDRUGS_PROBCELL_DN	15	0,014925373	0,08670663
IGFR_IR_UP	18	0,016	0,08901257
BENZOATE_DEGRADATION_VIA_COA_LIGATION	10	0,008333334	0,08812336
ELONGINA_KO_DN	177	0	0,09299067
CROONQUIST_IL6_RAS_DN	24	0,01754386	0,09609271

Supplemental Table 3

GSEA SUMMARY
UPREGULATED IN PC-3/S (FDR q-val ≤ 0.1)

NAME	SIZE	NOM p-val	FDR q-val
SANA_TNFA_ENDOTHELIAL_UP	80	0	0
HINATA_NFKB_UP	106	0	0
NAKAJIMA_MCS_UP	92	0	0
GILDEA_BLADDER_UP	30	0	0
ZUCCHI_EPITHELIAL_DN	43	0	0
CORDERO_KRAS_KD_VS_CONTROL_UP	74	0	7,01E+03
TNFA_NFKB_DEP_UP	18	0	6,01E+03
LINDSTEDT_DEND_8H_VS_48H_UP	64	0	5,26E+02
VERHAAK_AML_NPM1_MUT_VS_WT_UP	189	0	4,68E+02
BENNETT_SLE_UP	28	0	4,21E+02

IFNALPHA_NL_UP	27	0	3,83E+03
IFNA_HCMV_6HRS_UP	53	0	6,84E+02
KNUDSEN_PMNS_UP	74	0	6,31E+03
PASSERINI_GROWTH	33	0	5,86E+02
RADAEVA_IFNA_UP	50	0	5,47E+03
IL1_CORNEA_UP	62	0	5,13E+03
DER_IFNA_UP	66	0	4,83E+02
SANA_IFNG_ENDOTHELIAL_UP	73	0	4,56E+03
CARIES_PULP_HIGH_UP	91	0	6,62E+02
ESR_FIBROBLAST_UP	50	0	6,29E+02
CHEN_HOXA5_TARGETS_UP	229	0	5,99E+02
GRANDVAUX_IFN_NOT_IRF3_UP	15	0	0,001336657
TAKEDA_NUP8_HOXA9_3D_UP	187	0	0,001452867
CMV_8HRS_UP	32	0	0,001933808
DER_IFNB_UP	93	0	0,002033628
GRANDVAUX_IRF3_UP	13	0	0,002927756
JNK_UP	30	0	0,003289401
VERNELL_PRB_CLSTR2	22	0	0,003484217
YANG_OSTECLASTS_SIG	38	0	0,004366924
CMV_UV-CMV_COMMON_HCMV_6HRS_UP	20	0	0,00521013
XPB_TTD-CS_DN	24	0	0,005184943
IFNALPHA_NL_HCC_UP	18	0	0,005022914
CMV-CHX_HCMV_6HRS_DN	13	0	0,005899884
TFF2_KO_UP	23	0	0,005968231
DORSEY_DOXYCYCLINE_UP	29	0	0,00615452
AGUIRRE_PANCREAS_CHR7	48	0	0,007036564
CHANG_SERUM_RESPONSE_DN	124	0	0,008074184
RUTELLA_HEPATGFSNDCS_UP	158	0	0,009080959
CMV_HCMV_TIMECOURSE_12HRS_UP	26	0	0,009720216
ROTH_HTERT_UP	14	0	0,009792994
UVB_NHEK3_C3	17	0	0,011612952
HYPOPHYSECTOMY_RAT_DN	49	0	0,011942437
LINDSTEDT_DEND_UP	50	0	0,012549402
CYTOKINEPATHWAY	20	0	0,012264189
LAIRPATHWAY	15	0	0,012279492
HEDVAT_ELF_UP	11	0	0,012195339
VEGF_HUVEC_30MIN_UP	24	0	0,011935864
DAC_FIBRO_UP	17	0	0,011774023
INFLAMPATHWAY	29	0	0,011533736
UVC_HIGH_ALL_UP	19	0,00862069	0,01316197
BASSO_GERMINAL_CENTER_CD40_UP	97	0	0,012903892
IFNA_UV-CMV_COMMON_HCMV_6HRS_UP	29	0,008264462	0,012893538
AGEING_KIDNEY_UP	406	0	0,012884542
IFN_ALPHA_UP	40	0	0,013283294
MUNSHI_MM_UP	65	0	0,013349483
ET743_SARCOMA_72HRS_UP	66	0	0,013642795
STRESS_TPA_SPECIFIC_UP	42	0	0,013546075
HYPERTROPHY_MODEL	20	0,015037594	0,013885926
MUNSHI_MM_VS_PCS_UP	77	0	0,013798627
ADIP_HUMAN_DN	27	0	0,013776199
TSA_HEPATOMA_UP	36	0	0,013550359
DER_IFNG_UP	62	0	0,013677579
NI2_MOUSE_UP	39	0	0,013666228
HASLINGER_B_CLL_11Q23	20	0	0,013982471
BLEO_HUMAN_LYMPH_HIGH_24HRS_UP	92	0	0,0146018
CMV_HCMV_TIMECOURSE_48HRS_DN	110	0	0,014761495
ADIPOGENESIS_HMSC_CLASS8_DN	32	0,007692308	0,015350604
CELL_SURFACE_RECEPTOR_LINKED_SIGNAL_TRANSDUCTION	130	0	0,015312407
HALMOS_CEBP_UP	50	0	0,016364615
HDACI_COLON_BUT16HRS_UP	42	0	0,016676474
CARIES_PULP_UP	205	0	0,016804406
REOVIRUS_HEK293_UP	236	0	0,017220277
CROONQUIST_IL6_STROMA_UP	37	0	0,017162452
IDX_TSA_DN_CLUSTER3	78	0	0,01813489
LIZUKA_L1_GR_G1	20	0	0,017947914
ERYTHPATHWAY	14	0	0,018380348
EMT_UP	61	0	0,018521752
AD12_24HRS_DN	18	0	0,019753436
SHEPARD_POS_REG_OF_CELL_PROLIFERATION	93	0	0,021697486
DSRNA_UP	38	0,007936508	0,022729788
BOQUEST_CD31PLUS_VS_CD31MINUS_DN	265	0	0,022758491
FERRARI_4HPR_UP	22	0	0,02248095
TGFBETA_EARLY_UP	47	0,015503876	0,024247209
LEI_MYB_REGULATED_GENES	317	0	0,024707815
IFN_ANY_UP	81	0	0,024466466

IL6_FIBRO_UP	47	0	0,02447332
AGEING_KIDNEY_SPECIFIC_UP	183	0	0,02443541
CMV_HCMV_TIMECOURSE_ALL_DN	420	0	0,025697103
MYOD_NIH3T3_DN	56	0,007633588	0,027683567
IDX_TSA_UP_CLUSTER1	25	0	0,027563699
CMV_ALL_UP	93	0	0,028426403
CMV_HCMV_6HRS_UP	25	0	0,028299272
FERRANDO_TAL1_NEIGHBORS	15	0,00862069	0,028223107
ET743_HELA_UP	56	0	0,030668546
HDACI_COLON_BUT48HRS_UP	96	0	0,03087268
CMV-UV_HCMV_6HRS_UP	120	0	0,030595317
INSULIN_ADIP_SENS_UP	13	0,008064516	0,03032406
IRITANI_ADPROX_VASC	151	0	0,03169152
IRITANI_ADPROX_DN	60	0	0,032476414
SERUM_FIBROBLAST_CORE_DN	196	0	0,032151647
GNATENKO_PLATELET	44	0	0,034101162
FLECHNER_KIDNEY_TRANSPLANT_REJECTION_UP	86	0,006711409	0,038110312
STEMPATHWAY	15	0	0,038064986
AD12_48HRS_DN	13	0,024793388	0,038875837
BAF57_BT549_UP	238	0	0,04102632
GALINDO_ACT_UP	76	0,008333334	0,041229043
AD12_ANY_DN	25	0,007407407	0,042335518
YAMA_RECURRENT_HCC_UP	12	0,014814815	0,04249489
ROSS_MLL_FUSION	83	0	0,042640064
CMV_HCMV_TIMECOURSE_24HRS_DN	42	0,015037594	0,042443175
TAKEDA_NUP8_HOXA9_16D_UP	172	0	0,04320617
O6BG_RESIST_MEDULLOBLASTOMA_UP	23	0	0,04405446
ABBUD_LIF_UP	40	0	0,044147205
NF90_UP	25	0	0,04427938
HTERT_DN	70	0,02	0,043894343
CALRES_RHESUS_UP	67	0,007462686	0,043551825
IFN_GAMMA_UP	38	0	0,044692803
UVB_NHEK1_C2	21	0	0,045417644
GNATENKO_PLATELET_UP	44	0	0,045143023
IL1RPATHWAY	31	0,006756757	0,047090124
IFNALPHA_HCC_UP	29	0	0,04980294
RANKLPATHWAY	13	0,024390243	0,050466835
CROONQUIST_RAS_STROMA_DN	19	0,007936508	0,052343667
LEE_E2F1_UP	60	0	0,05265764
TAKEDA_NUP8_HOXA9_10D_UP	190	0	0,053419836
TSADAC_RKOEXP_UP	16	0,022727273	0,053503282
SARCOMAS_HISTIOCYTOMA_UP	13	0,015503876	0,0538131
MKK6EE_UP	10	0,016666668	0,05900869
PASSERINI_SIGNAL	338	0	0,059466176
TPA_SENS_MIDDLE_UP	65	0	0,059656978
CROMER_HYPOPHARYNGEAL_MET_VS_NON_UP	72	0	0,059361473
JNK_DN	31	0,009090909	0,06050751
IFN_BETA_UP	65	0	0,060272537
DAC_BLADDER_UP	28	0,007692308	0,06032948
JECHLINGER_EMT_UP	54	0	0,06016439
IL5PATHWAY	10	0,007462686	0,060459528
NING_COPD_DN	119	0	0,06069287
IFN_ALL_UP	18	0,007692308	0,06280084
ADIP_DIFF_CLUSTER2	40	0,014084507	0,063224
VEGF_MMMEC_12HRS_UP	28	0,023076924	0,063946106
TGF_BETA_SIGNALING_PATHWAY	49	0,008064516	0,06480047
UVB_NHEK3_C1	57	0,007751938	0,0647855
GLUTATHIONE_METABOLISM	31	0,006944445	0,07009193
FSH_OVARY_MCV152_UP	61	0,007407407	0,07044436
KRETZSCHMAR_IL6_DIFF	145	0	0,07082966
HYPERME_COLONCA_SW48	18	0,02238806	0,07288069
PASSERINI_PROLIFERATION	64	0	0,07502339
GALE_FLT3ANDAPL_DN	18	0,015384615	0,07624459
AD12_32HRS_DN	14	0,007246377	0,07736945
TAKEDA_NUP8_HOXA9_6H_DN	40	0,006944445	0,07727484
NAKAJIMA_MCSMBP_MAST	46	0,014285714	0,07792875
PASSERINI_EM	35	0,01459854	0,07808091
WIELAND_HEPATITIS_B_INDUCED	106	0	0,08116941
ET743_SARCOMA_6HRS_UP	30	0,015625	0,08152316
DAC_PANC_UP	365	0	0,08197479
BROCKE_IL6	145	0	0,08193863
ARAPPATHWAY	20	0,020134227	0,08807551
NAB_LUNG_DN	54	0,02631579	0,09223271
IL17PATHWAY	14	0,03649635	0,09287429
TAKEDA_NUP8_HOXA9_8D_UP	151	0	0,09692045

SARCOMAS_LIPOSARCOMA_DN	11	0,026548672	0,09720942
UVC_LOW_ALL_DN	58	0,014925373	0,09666061
ET743_SARCOMA_UP	69	0	0,096955165
CHIARETTI_T_ALL_DIFF	278	0,006756757	0,09757595
NTHIPATHWAY	21	0,02919708	0,0983403

Supplemental Table 4

ESC-LIKE MODULE SIGNIFICANTLY ENRICHED IN PC-3/Mc OVER PC-3/S (GSEA FDR q-value < 0.001)

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
BCAT1	24	3,899618387	0,024511253
GLDC	153	2,208153725	0,03262873
ELOVL6	249	1,73195231	0,039267384
PRMT3	300	1,604063749	0,047328863
ORC1L	315	1,581285715	0,057053167
CDCA7	316	1,580438733	0,06747696
CCNA2	442	1,333389044	0,06997601
PRIM1	457	1,31324625	0,077932455
GJA1	459	1,311803579	0,086534105
LMNB1	482	1,277396679	0,0938512
PLK1	483	1,273542047	0,10225086
CRABP2	545	1,179282665	0,106956705
SLC2A1	562	1,158363104	0,11379089
AK3L1	571	1,144245148	0,12093487
BLM	589	1,127632618	0,12751602
BAX	637	1,085546017	0,13230869
TTK	655	1,072149873	0,13852389
MRPL15	675	1,064553618	0,14458828
CCNF	690	1,049240232	0,15080348
NDC80	694	1,047491074	0,15756112
RUVBL1	697	1,046208262	0,16436067
CDCA3	701	1,042423725	0,1710849
TCOF1	717	1,033164501	0,1771437
PA2G4	729	1,025664568	0,18335448
WEE1	748	1,01253593	0,18912613
LYPLA1	758	1,009093523	0,19532835
NCAPD2	762	1,007425308	0,20182174
AURKB	772	0,999852598	0,20796302
HELLS	795	0,9842031	0,21334636
MCM2	797	0,983908951	0,21978538
CDC7	815	0,975096107	0,22536047
CDC20	920	0,910025477	0,22612482
RCC1	938	0,904319108	0,2312331
VRK1	976	0,881955922	0,23518664
DARS2	1017	0,863799453	0,23886932
POLR3K	1029	0,855638862	0,24395871
MYC	1052	0,84865129	0,24844801
COQ3	1090	0,83165437	0,25206977
NCAPH	1107	0,827557206	0,25672215
FAM60A	1131	0,814595163	0,26093647
GMNN	1143	0,810519516	0,26572827
GSPT2	1159	0,802991271	0,27026895
KIF22	1161	0,802636445	0,2755124
CHEK2	1163	0,80181396	0,2807504
KIF11	1186	0,796971023	0,28489885
KRAS	1208	0,790333569	0,2890539
MRPS2	1273	0,767228007	0,29089093
HSPE1	1298	0,760527909	0,2946983
PLK4	1307	0,757284641	0,29929006
NTHL1	1308	0,756931126	0,3042824
HNRNPL	1321	0,752166629	0,308639
ERCC6L	1368	0,740334034	0,31120518
CKAP2	1373	0,737542927	0,3158682
SOX2	1381	0,73558563	0,32036722
KIF23	1389	0,732352912	0,32484493
MYBL2	1391	0,732085645	0,329623
EXOSC7	1404	0,728375554	0,33382267
SMC4	1420	0,722967386	0,33783558
WDR57	1431	0,719718933	0,34207886
SIP1	1454	0,712792337	0,3456721

CDK4	1456	0,712302268	0,35031974
BUB1	1489	0,703863204	0,35335046
ACAD8	1495	0,702489436	0,35773194
IARS	1515	0,697458684	0,36137512
MID1IP1	1527	0,693232715	0,36539337
PDHA1	1545	0,689281404	0,36908337
RAD18	1563	0,683480561	0,37273508
NASP	1585	0,677246749	0,37614426
MAPK13	1593	0,675589919	0,3802476
CDCA8	1596	0,674183011	0,38459343
CHEK1	1600	0,672198892	0,38887584
CCNB2	1610	0,669424951	0,39283776
MAD2L1	1639	0,661989689	0,39579377
RNPS1	1643	0,660863757	0,40000144
KIF4A	1673	0,653722167	0,40285254
WDR77	1708	0,646371245	0,40540335
ETFA	1709	0,646137655	0,40966496
RACGAP1	1784	0,625950873	0,41006657
RCN2	1798	0,623038948	0,41352114
ANP32E	1804	0,621429801	0,41736796
MCM5	1855	0,608136475	0,4188608
MRPL12	1869	0,604957402	0,4221961
RAD23B	1904	0,596578598	0,42441848
RFC3	1947	0,588908195	0,4261874
PPM1G	1965	0,58620894	0,42919758
KIF20A	1983	0,5806337	0,432171
MSH2	1990	0,57921797	0,43568906
SMC2	2018	0,573566973	0,43811223
LSM4	2028	0,571519554	0,44142842
CKS2	2029	0,571243048	0,44519606
NOL1	2036	0,56962496	0,44865087
AURKA	2055	0,565522671	0,45147425
WBP11	2090	0,558767676	0,45344725
MCM7	2098	0,557399094	0,45677105
EXO1	2151	0,547469914	0,45776305
PRPS1	2156	0,546717286	0,46116745
RCC2	2181	0,541604519	0,46353093
TIMM8A	2208	0,536585689	0,46576056
PSME3	2211	0,53593576	0,4691946
TRIP13	2266	0,523357689	0,4699268
MRTO4	2276	0,521741867	0,4729147
TOP2A	2305	0,517573416	0,47491822
CCNC	2407	0,500260413	0,47313106
NEK2	2446	0,492967606	0,47446865
GEMIN6	2454	0,492182165	0,4773623
CHAF1A	2456	0,492103606	0,48055762
EBNA1BP2	2500	0,484164745	0,48158532
MCM3	2503	0,483869523	0,48467597
MRPL11	2578	0,471075058	0,48405612
NDUFA9	2632	0,462631792	0,48443818
SNRPA	2639	0,461376041	0,487179
RPP40	2684	0,455854833	0,48796967
SLC16A1	2732	0,449270457	0,48856577
DLAT	2743	0,44733125	0,49101254
MRPL4	2782	0,441830605	0,49201286
NCL	2790	0,441043735	0,4945692
NDUFAB1	2795	0,440249294	0,49727142
HADH	2800	0,439697057	0,49997002
DTL	2818	0,438132793	0,50200355
DHX9	2852	0,430988461	0,50318414
MRPL16	2862	0,429589182	0,5055643
HNRPK	2897	0,425457001	0,506658
LSM2	2929	0,42145735	0,5078765
STOML2	2950	0,418758392	0,5096312
NUP107	2963	0,417551965	0,5117808
CDCA5	2965	0,417491674	0,514484
STIP1	3002	0,411153436	0,5153827
PHF5A	3003	0,411020458	0,5180936
NDUFB10	3025	0,408638507	0,51973116
NT5DC2	3062	0,403542876	0,5205797
POLE2	3115	0,396017581	0,5205728
SS18	3137	0,392607808	0,5221046
HSPA14	3221	0,384500265	0,5204605
SSB	3240	0,382236511	0,522075
MRPL37	3256	0,381094784	0,5238331

PUS1	3270	0,379269958	0,5256798
HDAC1	3293	0,37704882	0,52705866
PARP1	3368	0,369399488	0,5257682
APEX1	3374	0,369081467	0,5279507
SEPHS2	3387	0,367627114	0,52977103
EIF4EBP1	3402	0,364599973	0,53147066
C2ORF47	3426	0,362476915	0,53270304
CKS1B	3463	0,358507395	0,53325456
NIP7	3535	0,350783765	0,5319924
RRM1	3545	0,349941283	0,53384715
RUVBL2	3564	0,347790807	0,5352345
GTSE1	3572	0,347215861	0,537172
NDUFB7	3574	0,347083002	0,5394108
ADSL	3578	0,346514255	0,5415452
SNRPD1	3650	0,337978065	0,54019856
SET	3663	0,336766362	0,5418154
PDCD2	3710	0,332380891	0,54169095
FBL	3722	0,33170104	0,54332465
BANF1	3753	0,328677744	0,5439816
DEK	3789	0,325253308	0,5443641
PPP4C	3824	0,321879745	0,5447747
DBF4	3841	0,320631176	0,5460836
NME4	3867	0,317296803	0,5469173
FH	3907	0,31362012	0,5470217
EIF4B	3909	0,313282311	0,5490376
HNRPAB	3925	0,311922163	0,5503394
HMGB2	3941	0,310527384	0,55163205
CDC2	3950	0,309771448	0,55327225
POLR2F	3998	0,305880994	0,55292267
PRIM2	4060	0,300804436	0,55183446
DLG7	4096	0,298359066	0,5520396
GNL3	4107	0,297401696	0,5534975
BXDC2	4137	0,295413077	0,5539854
SPAG5	4169	0,29224503	0,5543516
COX4NB	4171	0,291865647	0,5562263
NDUFB8	4187	0,290363997	0,557386
RPA2	4190	0,290097117	0,55919856
ABCB7	4247	0,286600977	0,55826855
CBX3	4286	0,282994628	0,5582212
SLC25A5	4310	0,28066954	0,55891407
SNRPA1	4376	0,275494069	0,5574575
EIF2S3	4413	0,272629827	0,5574426
VBP1	4425	0,271159291	0,558677
ECHS1	4426	0,271158695	0,56046546
CCT5	4489	0,265886426	0,55909663
XPO1	4520	0,262975156	0,5593202
PHB	4655	0,251748234	0,55423206
TIMM44	4666	0,251202285	0,55538523
PSMA5	4684	0,249446422	0,5561743
PCNA	4748	0,243616909	0,5546082
E2F3	4758	0,243004829	0,5557577
EIF3A	4763	0,242438942	0,55715525
RRM2	4857	0,234221309	0,55401635
BIRC5	4892	0,232178226	0,55383533
GART	4938	0,228193328	0,55307406
BUB1B	5011	0,222909316	0,55091816
FARSA	5016	0,222621366	0,552185
TIMM13	5063	0,218337998	0,5513084
YY1	5102	0,215833783	0,55081815
PRDX1	5131	0,214457929	0,55082244
PSMD14	5139	0,214090288	0,5518819
LSM10	5180	0,212055907	0,551266
HAT1	5183	0,212006688	0,5525636
CDC6	5267	0,206298381	0,5497441
FUSIP1	5288	0,204966754	0,55008876
PSMB5	5291	0,204689428	0,5513381
G3BP1	5292	0,20468533	0,55268806
NUDCD2	5376	0,199192911	0,54982173
COX5B	5456	0,194310099	0,5471247
UQCRH	5553	0,188049734	0,54353017
SDHD	5591	0,185857505	0,5428926
GNPDA1	5616	0,184014201	0,5428975
UGDH	5697	0,177977592	0,54004234
KPNA2	5725	0,17651622	0,5398468
DNMT1	5827	0,169651449	0,5358791

EIF3K	5833	0,169260636	0,53674364
TRIP6	5934	0,164014176	0,5327892
MKI67IP	5960	0,162491783	0,5326018
MTHFD2	5977	0,161590725	0,53286177
CLPP	6000	0,160029337	0,53280926
HSPA9	6088	0,154952526	0,5294497
NONO	6099	0,154089466	0,52996236
HN1	6137	0,151880205	0,5291007
POP7	6145	0,151683912	0,52974856
MTF2	6156	0,151023716	0,530241
HNRNPA1	6268	0,14486964	0,5256063
DTYMK	6341	0,140549019	0,52290714
UBE2G1	6346	0,140204892	0,52363044
DDX18	6357	0,139639989	0,5240478
NDUFS2	6403	0,137281835	0,5226869
FDPS	6457	0,134413287	0,52090424
ENO1	6485	0,1330982	0,5204223
NOLA2	6548	0,130358726	0,51815957
MRPL13	6675	0,124312826	0,5126338
PIPOX	6797	0,11854133	0,5073218
GLO1	6887	0,113809578	0,5035901
PROM1	6889	0,113685608	0,50428957
SUMO1	6905	0,112926453	0,50427896
YAP1	6988	0,108949862	0,5008678
C11ORF47	6991	0,10888356	0,5014852
EEF1E1	7065	0,105234623	0,49850282
TGIF2	7165	0,101080589	0,4941836
CCDC5	7209	0,098995738	0,49267092
CSE1L	7332	0,092902146	0,48713943
PSMB6	7363	0,091252521	0,4862304
XRCC5	7392	0,090024307	0,485414
PRMT1	7394	0,089996532	0,4859572
MRPS18B	7405	0,08921133	0,486042
EIF4A1	7513	0,084386095	0,48120975
EEF2	7671	0,076700591	0,4738087
CYCS	7833	0,069111079	0,46615615
RPSA	7880	0,067212172	0,46428278
TIMM8B	7944	0,064575806	0,46153584
U2AF1	8154	0,055781722	0,45137796
UTP18	8295	0,048784971	0,44464895
NLN	8372	0,044910282	0,44111758
CISD1	8386	0,04427832	0,44075492
RPL10A	8405	0,042954925	0,4401317
BTF3	8505	0,038467545	0,43539953
DAP3	8550	0,036765601	0,43342605
POLD1	8613	0,034197357	0,43052912
MCM4	8636	0,033148795	0,4296398
IPO9	8794	0,026349641	0,42190662
RPL27A	8978	0,019887455	0,41282144
SDHC	9074	0,016446242	0,40814546
BUB3	9075	0,016397966	0,4082536
EIF2S2	9088	0,015929315	0,40775433
CTSC	9162	0,012833091	0,4041625
TGIF1	9372	0,004258764	0,3936648
NME2	9471	4,16E-04	0,38873202
ERP29	9528	-0,001733117	0,38592315
RPS8	9636	-0,00627123	0,38057572
THOC3	9819	-0,013617452	0,3714995
RPS3	9856	-0,015123445	0,3697862
PSMA7	9933	-0,01847915	0,36608052
NCBP2	9998	-0,021485306	0,36299902
EIF3I	10066	-0,024087554	0,35978362
NIPSNAP1	10196	-0,029125676	0,35347894
TCF19	10198	-0,029251795	0,3536215
RPS19	10277	-0,032288346	0,34990618
MRPS28	10336	-0,03420496	0,34721074
LSM5	10562	-0,043191701	0,33616403
RPS16	10615	-0,045566306	0,3338457
PABPC1	10623	-0,045966893	0,33379635
NAP1L1	10861	-0,054676116	0,322221
RPA3	10899	-0,056132015	0,32072783
RPS23	11063	-0,062256008	0,31292933
ADH5	11089	-0,063065417	0,3120862
FGFR1	11178	-0,066621527	0,3080937
SARS	11200	-0,067360051	0,30748037

OTX2	11453	-0,076988436	0,29529676
MRPS30	11481	-0,078453653	0,2944544
WDHD1	11487	-0,078637779	0,29472125
ZIC3	11691	-0,087173618	0,2850726
EIF3EIP	11984	-0,097988792	0,271013
RPS5	12089	-0,101972304	0,26644784
RPL13	12769	-0,131055951	0,23311602
RPL22	13103	-0,14517042	0,21730274
CCND2	13205	-0,150399059	0,21320806
KPNA6	13277	-0,153675541	0,2106459
NDUFA11	13419	-0,159458071	0,20459647
EIF6	13514	-0,163347855	0,20093974
NSBP1	13667	-0,169261634	0,194401
CSRP2	14077	-0,187520698	0,17503949
UBE2V2	14309	-0,198781744	0,16471678
RPS27	14350	-0,200828969	0,16402686
MRPL39	14425	-0,204573408	0,16164929
GARS	14502	-0,207676828	0,15919146
RPS12	14542	-0,20947367	0,1586089
NUSAP1	14637	-0,213867009	0,15528537
RAB34	14675	-0,215915367	0,15484603
CDKN3	14692	-0,216512904	0,15546824
NOLA3	14700	-0,216881111	0,15654615
MRPS17	14846	-0,224187165	0,15072219
MRPS36	14886	-0,227123171	0,15025604
CTNNA1	14979	-0,231953025	0,14715253
PHC1	15336	-0,252516806	0,13088892
TP53	15339	-0,252603412	0,13245425
DPP3	15404	-0,256631523	0,13092364
ATP5O	15428	-0,258410275	0,13146965
PDPN	15684	-0,274250001	0,120436005
TMEFF1	15697	-0,275346458	0,1216477
ATP5J	16414	-0,325332403	0,087733805
TEAD2	16473	-0,33045876	0,086992316
SNX5	16836	-0,364256412	0,071163505
ALDH7A1	17276	-0,411136746	0,051765975
UQCR	17597	-0,452683091	0,038635615
CCND1	17699	-0,468081534	0,036636226
PDIA4	18315	-0,596978724	0,009600597
TERF1	18521	-0,653074086	0,003583615
SQLE	18724	-0,72413224	-0,001813614
CDC34	18938	-0,815627038	-0,007161377
GNA14	19303	-1,009731531	-0,018833676
ALDOC	19405	-1,093921065	-0,016705338
CDKN1C	19516	-1,187802315	-0,014411068
AMOTL2	19760	-1,506880641	-0,016710542
ZNF22	19853	-1,716729164	-0,010021204
SERPINH1	19882	-1,79171896	3,86E-04
TCF7L1	20000	-2,278077126	0,009518581

MYC MODULE SIGNIFICANTLY ENRICHED IN PC-3/Mc OVER PC-3/S (GSEA FDR q-value = 0.004)

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
KDELC1	307	1,590066791	-0,002195116
PASK	404	1,384562373	0,004530856
CBX5	491	1,264579058	0,01075871
INCENP	492	1,261291265	0,02129487
B3GNT7	497	1,248222232	0,031520195
SLC25A10	561	1,160498381	0,03803818
BLM	589	1,127632618	0,046096593
PPP1R7	632	1,088734388	0,053073835
DLST	684	1,054434896	0,059310813
TACC3	726	1,026229858	0,06581634
UNG	744	1,017814517	0,073461555
WEE1	748	1,01253593	0,0817685
BIVM	759	1,008894086	0,0896921
AURKB	772	0,999852598	0,09743935
HELLS	795	0,9842031	0,10455171
TCEA1	919	0,910576761	0,10595703
RCC1	938	0,904319108	0,11260375
RNF41	939	0,903352857	0,12014987
MTHFD1L	1084	0,834251583	0,119858876
AGPAT5	1162	0,802391052	0,1226796
DBR1	1291	0,762768507	0,12259813

HSPE1	1298	0,760527909	0,12864868
PLK4	1307	0,757284641	0,1345713
ELK4	1422	0,722518384	0,13485943
CDK4	1456	0,712302268	0,1391459
SLBP	1500	0,700855732	0,14283259
IARS	1515	0,697458684	0,14795296
LRRC40	1559	0,684642553	0,15150422
MRPL50	1605	0,670510828	0,1548366
CACYBP	1626	0,664805114	0,1593817
FNBP4	1651	0,65948987	0,16368076
SEH1L	1678	0,652569473	0,16782115
SUMF1	1724	0,641619802	0,17091219
SGOL2	1726	0,640817761	0,17621483
NOLC1	1746	0,635382175	0,18056458
RANBP1	1778	0,62728703	0,18424171
KIFC1	1783	0,626474857	0,18927328
ANP32B	1833	0,613984764	0,1919318
PRPF38B	1851	0,609197259	0,19616364
MCM5	1855	0,608136475	0,20109245
MRPL12	1869	0,604957402	0,20549053
TCP1	1887	0,601010203	0,20965399
TOE1	1901	0,596905053	0,2139848
KIF20A	1983	0,5806337	0,21475142
SUPV3L1	1993	0,578305721	0,21912853
ENO3	2001	0,576447427	0,22359097
NUTF2	2025	0,572189331	0,22721116
NOL1	2036	0,56962496	0,23146535
CDT1	2065	0,563379645	0,23475988
TARBP2	2079	0,561234891	0,23879273
LUC7L	2205	0,537463903	0,23698042
PSME3	2211	0,53593576	0,24120528
TRIP13	2266	0,523357689	0,24285467
ATF7	2268	0,523110449	0,24717404
PSAT1	2302	0,518388689	0,24984066
TIAL1	2386	0,503932714	0,24986573
CIRH1A	2474	0,488385588	0,24955925
EBNA1BP2	2500	0,484164745	0,2523433
FUBP1	2550	0,476551414	0,2538538
IKBKAP	2570	0,471962154	0,2568384
POLH	2587	0,469142258	0,25995073
SPHK2	2602	0,466991186	0,2631459
SNRPA	2639	0,461376041	0,265185
SLC16A1	2732	0,449270457	0,26429972
ABCE1	2769	0,442988038	0,26618522
NCL	2790	0,441043735	0,26886114
DTL	2818	0,438132793	0,27115986
NOLA1	2830	0,435137898	0,2742402
ARMC8	2832	0,434812516	0,27782196
RAD1	2834	0,434570014	0,28140172
GTF3C4	2837	0,433470011	0,28492185
DHX9	2852	0,430988461	0,2878163
UBE2I	2873	0,428126246	0,2903843
TAF15	2895	0,425604254	0,29288083
CLNS1A	2901	0,424763054	0,296177
LRSAM1	2913	0,423086762	0,29915667
RNF4	2924	0,421816707	0,30217615
LSM2	2929	0,42145735	0,3054951
TAF12	2942	0,420149535	0,30839983
MRPL18	2967	0,416884929	0,31067228
MMAA	3013	0,409922749	0,31182784
MTHFD1	3058	0,403819501	0,31298286
PPID	3082	0,400933266	0,31517246
COX5A	3093	0,398849785	0,31800008
RG9MTD2	3103	0,397981226	0,32087088
KCTD9	3106	0,396869272	0,32408527
HSPD1	3123	0,395320803	0,3265809
SS18	3137	0,392607808	0,32920516
IPO7	3173	0,38890931	0,33068934
HSPA14	3221	0,384500265	0,3315317
DDX31	3345	0,371696591	0,32843548
SP1	3420	0,362949908	0,3277366
SOS1	3479	0,357127756	0,32779574
ANKRD10	3499	0,354453385	0,32979876
WDR34	3555	0,348705113	0,32993877
ADSL	3578	0,346514255	0,33172423

NFYC	3608	0,34239623	0,33312234
EIF5A	3613	0,342000216	0,33577758
EXOSC1	3617	0,341256589	0,33847702
MRPS12	3624	0,340464175	0,34101856
CCT3	3634	0,339827091	0,34340355
TRAPPC6B	3665	0,336595595	0,3447028
RTCD1	3672	0,335629523	0,347204
TUBA3C	3682	0,334444523	0,34954402
SARS2	3688	0,334116906	0,35208297
WDR23	3698	0,333459884	0,3544148
HNRPM	3746	0,3294321	0,35479712
SFRS1	3755	0,328354478	0,3571367
NDUFS1	3787	0,325529039	0,35829312
CCT4	3855	0,318490922	0,35757574
XPO5	3895	0,314768374	0,35823894
PIGW	3916	0,312535107	0,35984138
ISG20L1	3935	0,310768068	0,3615299
HMMR	3949	0,309806764	0,36346242
ILF3	3994	0,305987388	0,3638002
SF3A2	4066	0,300473541	0,36273065
UBQLN4	4103	0,297722995	0,3634027
SF3B3	4129	0,295968056	0,36461467
BXDC2	4137	0,295413077	0,36672947
QTRTD1	4150	0,294006646	0,36858046
CCT7	4203	0,289602906	0,368378
GMPS	4213	0,288628638	0,37033534
ZDHHC16	4270	0,284742087	0,36989063
ZCCHC7	4272	0,28453052	0,37221703
CBX3	4286	0,282994628	0,3739256
BCLAF1	4296	0,282078236	0,37582818
NDUFA13	4371	0,275890291	0,37440205
SNRPA1	4376	0,275494069	0,3765017
PRPF19	4378	0,275437117	0,37875217
AAAS	4440	0,269995511	0,3779322
LDHA	4482	0,266263992	0,37808937
MCM6	4567	0,258907706	0,3760172
ADPGK	4573	0,258199304	0,37792197
STC2	4594	0,256993711	0,37906045
PAICS	4645	0,25280425	0,37865144
CSDE1	4649	0,2522493	0,38060737
FCHSD2	4653	0,251947016	0,38256073
EIF3A	4763	0,242438942	0,3790906
RRM2	4857	0,234221309	0,37635848
SUZ12	4887	0,232702017	0,3768403
HNRNPA2B1	4895	0,232003987	0,37842542
NUP50	4916	0,230546519	0,37934297
EPRS	4923	0,229799926	0,3809601
ST13	4984	0,225303218	0,37981722
AMD1	4987	0,224784553	0,38159412
PRDX1	5131	0,214457929	0,3761761
MARS	5133	0,2143545	0,37791628
ATPIF1	5155	0,213179111	0,37863833
EEF1B2	5192	0,211244315	0,378588
MRP63	5198	0,211029559	0,38009873
PRKDC	5200	0,210932985	0,38181034
TRNT1	5209	0,21056518	0,38316596
ATIC	5237	0,209014848	0,38355073
G3BP1	5292	0,20468533	0,3825381
MAPBPIP	5338	0,201544225	0,38195297
TOMM20	5439	0,195436254	0,37854397
AHI1	5563	0,187401548	0,37390825
RPL14	5606	0,184684023	0,37333354
MRPL17	5638	0,182327613	0,37329373
MAT2B	5786	0,172561616	0,36732405
RAD50	5788	0,172426045	0,368714
ZNHIT2	5808	0,170953795	0,36918417
NACA	5882	0,166376516	0,36689362
TOPORS	5990	0,160409972	0,36283907
CLPP	6000	0,160029337	0,36372215
RANBP2	6027	0,158805728	0,3637379
POLR3B	6047	0,157691002	0,36409727
PCBP1	6063	0,15681015	0,36465093
POP7	6145	0,151683912	0,36183435
NDUFS6	6181	0,14958328	0,36131933
MAN1B1	6211	0,147834137	0,36109218

IMP4	6254	0,145503536	0,36019018
HNRNPA1	6268	0,14486964	0,36074492
NOL5A	6286	0,14371796	0,3610884
DBT	6351	0,139878914	0,35903028
NME1	6367	0,138956249	0,35943478
RAB25	6387	0,137895599	0,3596288
ESPL1	6450	0,134971663	0,3576305
DNAJB4	6654	0,125126764	0,3484413
ATP5G2	6660	0,124952123	0,349233
NVL	6762	0,120141149	0,3451446
MRPL36	6775	0,119418137	0,34553716
PPAN	6995	0,108655654	0,3354037
LMLN	7011	0,107967012	0,33554938
PFN1	7101	0,103437401	0,3319264
PPP2R4	7123	0,102448903	0,33172348
NT5C3	7387	0,090183243	0,31921744
DDX54	7412	0,08888077	0,31874993
RPL18	7478	0,086226232	0,31619316
RPLP2	7517	0,08430177	0,31498158
NPM1	7609	0,07935483	0,3110566
EEF2	7671	0,076700591	0,30862197
HNRNPR	7715	0,074719451	0,30707824
RRS1	7744	0,073510401	0,30628067
CSTF1	7787	0,071475945	0,30476028
HNRPH1	7881	0,067197368	0,30063292
SF3A1	7910	0,066330485	0,29977536
AASDHPPT	7927	0,06540598	0,29951507
DDB1	7935	0,065048084	0,29970554
M6PRBP1	7989	0,062761992	0,29755777
ATP5C1	8123	0,057176076	0,29133007
TNPO3	8221	0,052461717	0,28687796
RPS13	8250	0,05108938	0,28589308
RAF1	8340	0,046442736	0,28179404
EWSR1	8408	0,042858548	0,27877417
ATAD2	8457	0,040777959	0,27669486
PPCS	8470	0,0398071	0,27642238
CETN3	8524	0,037906572	0,274067
ANAPC10	8632	0,033297472	0,26895064
MCM4	8636	0,033148795	0,26907632
RPL32	8809	0,025747294	0,26061985
EEF1A1	8861	0,024220936	0,25825095
SAMD10	8960	0,020427961	0,25348085
MYO1G	8962	0,020369995	0,2536006
PMF1	8971	0,020124994	0,25336537
SCAMP4	9051	0,017412599	0,24952798
RPL3	9063	0,016852332	0,24911417
EIF2S2	9088	0,015929315	0,24803725
ARFRP1	9092	0,015731314	0,24801742
PRKAB1	9231	0,009724466	0,24114124
KLF16	9287	0,007440278	0,23843053
PTBP1	9335	0,005435493	0,23610638
SLC3A2	9342	0,005294009	0,23584811
RPL23	9376	0,004081874	0,23421848
PLOD3	9407	0,002853165	0,23272984
IFRD1	9605	-0,004953056	0,22283928
IHPK1	9627	-0,005879312	0,22182965
CCNI	9732	-0,010406705	0,21667333
RPL23A	9748	-0,01105639	0,21600945
SNRPE	9770	-0,012135662	0,21505208
TRIM8	9801	-0,013024894	0,21364841
RPS3	9856	-0,015123445	0,21105228
MUTYH	9989	-0,021083463	0,20457351
IPMK	10071	-0,024409968	0,20069373
GNPAT	10097	-0,025324009	0,19964486
RPL7	10103	-0,025621641	0,19960682
SMARCAD1	10134	-0,026890334	0,19831897
NFKBIL1	10207	-0,02958221	0,19493614
RPS19	10277	-0,032288346	0,19172716
PNN	10283	-0,032537937	0,19174688
MRPS28	10336	-0,03420496	0,18941098
RPL35A	10413	-0,037139438	0,18588962
SLC6A9	10460	-0,038626287	0,18389314
RPLP1	10462	-0,038644683	0,18416555
DNAJC8	10527	-0,041568018	0,18128616
HTF9C	10650	-0,046812426	0,17552647

HSD17B12	10667	-0,047445655	0,17511615
RPS27A	10681	-0,047972001	0,17486148
RPS14	10683	-0,048101511	0,17521288
TBRG4	10790	-0,052023787	0,17030336
RPS24	10794	-0,052301183	0,17058901
ADHFE1	10819	-0,053098667	0,16982259
RPS27L	10877	-0,055163551	0,16740969
ZMYND12	10922	-0,056951381	0,16566713
THRAP3	11034	-0,061282061	0,16058287
ADH5	11089	-0,063065417	0,15838723
CNIH4	11150	-0,06550689	0,15590948
SLC7A6	11157	-0,065845989	0,15615703
MRPL49	11201	-0,067421079	0,15455234
LRDD	11334	-0,072693236	0,14850469
RPL37	11397	-0,075031757	0,14600568
PLEKHJ1	11457	-0,077046327	0,14367473
MRPS30	11481	-0,078453653	0,14317054
NDOR1	11667	-0,08607123	0,13456258
S100A13	11767	-0,089879617	0,1303222
CHCHD1	11904	-0,094822153	0,12425773
CSNK2A1	11913	-0,095166765	0,124649376
NFE2L1	11936	-0,095822848	0,124340676
BRD8	12085	-0,101734079	0,117728956
HMGA1	12108	-0,102963857	0,117479905
RPS25	12206	-0,107124232	0,11348442
DPH3	12283	-0,110397711	0,11057501
RDH10	12337	-0,112644926	0,108843945
ZFP36L2	12400	-0,115629964	0,106684074
CCT8	12449	-0,117531523	0,1052459
CFL1	12456	-0,118004918	0,10592916
GNB2L1	12550	-0,121853791	0,10225838
NEK8	12587	-0,123648837	0,101476304
ATF4	12825	-0,133220971	0,09064058
AMH	12878	-0,135384113	0,089149885
BRE	12896	-0,136517152	0,0894332
RPL28	13119	-0,145945668	0,07946002
GNAS	13164	-0,148344338	0,07848091
WBSCR16	13231	-0,151616588	0,07641998
GMFB	13253	-0,15250735	0,07663521
NMD3	13283	-0,153852031	0,07645835
DDX5	13388	-0,157760262	0,07253294
SOD1	13407	-0,158637092	0,07295062
EIF3B	13410	-0,158809483	0,0741764
JMJD6	13445	-0,160584971	0,0738037
LMAN2	13489	-0,162455916	0,07299288
COX8A	13549	-0,164788991	0,071394905
MAP3K12	13631	-0,168135732	0,06871573
ATP5E	13739	-0,17192407	0,06475739
ERH	13997	-0,183119982	0,05333018
ZBTB37	14030	-0,185194477	0,053263888
ATP6V1F	14119	-0,189218432	0,050407916
RLBP1	14177	-0,191928521	0,049137477
LNPEP	14181	-0,192323506	0,0505928
TIMM9	14292	-0,197834328	0,04669965
CANX	14344	-0,200624019	0,045804344
RNMT	14424	-0,2045535	0,043530215
GARS	14502	-0,207676828	0,041383013
RPS12	14542	-0,20947367	0,041166622
EIF4A2	14617	-0,21311529	0,039216094
RPS11	14622	-0,213369742	0,04079681
AKT1S1	14625	-0,213446543	0,042478997
RAB34	14675	-0,215915367	0,04181226
DDX17	15001	-0,23286286	0,027372293
PNPLA2	15058	-0,235849157	0,02651916
KCNK5	15167	-0,242075622	0,023096412
ATAD3A	15178	-0,243013069	0,024622254
ZNHIT1	15292	-0,249678984	0,02101094
MUS81	15644	-0,271707088	0,005584645
SLC9A8	15754	-0,279260486	0,002422103
OSGEP	15908	-0,289405793	-0,002873992
ABCB8	15989	-0,295456409	-0,00443918
CRAT	16002	-0,296785712	-0,002564981
MSH4	16094	-0,302235752	-0,004628114
XRN2	16138	-0,304947495	-0,004248628
PPP1R15B	16450	-0,328144372	-0,01718684

TRAPPC4	16584	-0,340747237	-0,021045737
MAF1	16723	-0,353057027	-0,025053883
SNX5	16836	-0,364256412	-0,027657663
PLA2G6	16851	-0,365752071	-0,025308186
ARV1	16961	-0,377702445	-0,027648397
CNOT4	16971	-0,378594458	-0,024939563
JUNB	17036	-0,385386974	-0,024946863
GDF9	17234	-0,406652242	-0,031481843
ARMET	17555	-0,44765237	-0,043875493
SLC5A5	17587	-0,451382428	-0,041667778
HDLBP	17624	-0,455966115	-0,03967385
STX12	17697	-0,467740238	-0,039396543
CCND1	17699	-0,468081534	-0,035536855
TNFAIP1	17794	-0,484433144	-0,036229253
LAPTM4A	17810	-0,486755043	-0,0329194
RGS14	17835	-0,490500629	-0,030031998
SLC33A1	17953	-0,513479292	-0,031641327
MIS12	17995	-0,522030771	-0,029347613
COG4	18103	-0,544494867	-0,030193696
DERL2	18107	-0,544754386	-0,025794355
RBKS	18206	-0,567961812	-0,025990663
TM9SF4	18343	-0,605745792	-0,02778715
ALAD	18547	-0,659479737	-0,032512642
MMP14	18606	-0,678971827	-0,029764993
ZGPAT	18677	-0,702687085	-0,027424233
KBTBD3	18816	-0,765001774	-0,027991211
DARS	18839	-0,771892548	-0,02265238
SESN2	18995	-0,842873275	-0,023425933
VPS16	19298	-1,004797578	-0,030257998
BIRC2	19431	-1,110942483	-0,027632676
GTF2IRD2	19476	-1,15118885	-0,020234555
KIN	19546	-1,221551299	-0,013509061
TBC1D17	19676	-1,39048028	-0,008397381
FN1	20170	-4,095345974	9,58E-04

ES1 GENESET SIGNIFICANTLY ENRICHED IN PC-3/Mc OVER PC-3/S (GSEA FDR q-value < 0.001)

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
CD24	4	2,102022886	0,022120135
GAL	7	1,954879522	0,042778604
CXADR	14	1,818594337	0,061788056
NTS	21	1,721550465	0,079766974
RBM35A	23	1,680611491	0,09756336
KIF5C	31	1,608549237	0,11429186
TACSTD1	44	1,482986093	0,12943476
ST6GAL1	51	1,437490106	0,14439717
ZIC2	79	1,248459458	0,15629296
PTPRZ1	84	1,214943647	0,16899298
CBS	93	1,167224288	0,1809845
COBL	97	1,14108479	0,19295062
FZD7	98	1,134749055	0,2050008
TUBB2B	108	1,09352541	0,21615925
GPM6B	113	1,067097306	0,22728926
DCLK1	137	0,95648241	0,23628622
GLDC	153	0,916493952	0,24526209
LGR4	155	0,911314785	0,25488913
DIAPH2	177	0,850649714	0,2628631
FKBP5	181	0,831243038	0,27153897
CAMKV	203	0,779393494	0,27875626
EGLN3	211	0,765792966	0,28653532
ELOVL6	249	0,714244246	0,29225373
NMU	268	0,690258265	0,2986758
PRMT3	300	0,660546124	0,30412662
FRAT2	301	0,660076797	0,31113616
ORC1L	315	0,651006043	0,3173936
MAP7	333	0,628676474	0,32321218
PASK	404	0,568892121	0,3257225
CCNA2	442	0,547608852	0,32967135
DNMT3B	445	0,545396268	0,33536217
PRIM1	457	0,539239407	0,3405336
GJA1	459	0,538640141	0,34620312
CBX5	491	0,519036412	0,3501512
CRABP2	545	0,483687609	0,35261422

PODXL	582	0,465573877	0,3557424
TTLL12	611	0,455498546	0,35916707
ADAM23	627	0,447564125	0,36316326
SLC6A8	645	0,442284763	0,3670025
RUVBL1	697	0,428679615	0,36898223
NARG2	710	0,424587488	0,37288573
PSIP1	743	0,417675912	0,375707
UNG	744	0,416963369	0,38013485
LRP8	749	0,414501518	0,38433477
GABRB3	754	0,413878113	0,38852808
LYPLA1	758	0,413366169	0,3927664
AURKB	772	0,409555167	0,39645982
FZD5	794	0,403126001	0,39968145
HELLS	795	0,403102905	0,4039621
MCM2	797	0,402981639	0,40819103
TMPO	817	0,398855537	0,41146818
RAD51AP1	839	0,392751336	0,41457963
GINS2	840	0,392732918	0,41875017
DNA2L	857	0,388801426	0,42207187
FGF13	858	0,388622552	0,42619875
SMS	867	0,386550754	0,4299001
CACHD1	869	0,386254221	0,4339514
CDC20	920	0,372546166	0,4353855
EPB41L4B	957	0,365513653	0,4374511
HRASLS3	1027	0,350749135	0,43769532
SLC29A1	1043	0,348219335	0,44063652
SLC39A10	1062	0,346393108	0,44340703
TEAD4	1080	0,341573179	0,44617677
RFC4	1085	0,340935498	0,44959548
HAS3	1087	0,340637147	0,45316234
PUS7	1096	0,339830875	0,45636755
HOMER1	1104	0,338749349	0,45961174
NCAPH	1107	0,338622719	0,46310678
FAM60A	1131	0,333295107	0,46548596
LCK	1138	0,332374483	0,46871287
GMNN	1143	0,331620187	0,47203267
CDC25A	1152	0,329770237	0,47513106
NUDT15	1156	0,328933328	0,47847274
CHEK2	1163	0,328042954	0,48165366
POLR3G	1176	0,32707575	0,4845217
MDN1	1247	0,316117257	0,4843477
NPM3	1303	0,31046018	0,48487028
NTHL1	1308	0,309607685	0,48795632
FAM64A	1353	0,304453582	0,48896998
PDK1	1356	0,303897172	0,49209625
ERCC6L	1368	0,302793801	0,49475685
SOX2	1381	0,300844669	0,49734628
PIK3CB	1385	0,300022572	0,500381
NARG1	1438	0,293677092	0,50087667
MCM10	1477	0,289231569	0,5020313
BUB1	1489	0,287826657	0,50453293
HESX1	1555	0,280734688	0,50423545
NASP	1585	0,276908517	0,5057132
KRT8	1589	0,276700318	0,5085002
CHEK1	1600	0,274838328	0,5109144
GAD1	1612	0,273314267	0,5132619
MFGE8	1661	0,268621325	0,51369333
KIF4A	1673	0,267261952	0,51597655
M6PR	1683	0,266540647	0,51835304
AASS	1716	0,263223439	0,5195342
NOLC1	1746	0,259743452	0,5208297
ARL5B	1810	0,25338006	0,5203425
C1QBP	1812	0,253194571	0,52298087
MCM5	1855	0,248577297	0,523502
DCC1	1892	0,245337337	0,5242914
RBBP8	1898	0,244536862	0,526636
RFC3	1947	0,240699187	0,5267709
MSH2	1990	0,236729622	0,52716625
LSR	1994	0,236280069	0,529524
HMGB3	2042	0,232179075	0,52961886
CDT1	2065	0,230242506	0,5309541
CASP3	2066	0,230161965	0,5333983
TUBB2C	2094	0,228113919	0,53445876
MCM7	2098	0,227793261	0,5367264
CCAR1	2100	0,227508664	0,53909194

IGF2BP3	2107	0,226850748	0,5411983
EXOSC5	2116	0,226138666	0,5431962
PRPS1	2156	0,2234191	0,5436015
RCC2	2181	0,221325636	0,5447412
TMEM48	2189	0,221068591	0,5467357
FAM108B1	2228	0,217452615	0,54712814
SERBP1	2328	0,2100254	0,54436475
LARP7	2360	0,207380459	0,5450033
GPR19	2366	0,206998155	0,54694927
CHAF1A	2456	0,201062828	0,5445951
PTPN2	2468	0,200318441	0,5461675
CAPRIN2	2499	0,197842732	0,5467552
MCM3	2503	0,197693259	0,5487032
WDR12	2519	0,196408853	0,5500323
NUP37	2533	0,195494115	0,5514526
FUBP1	2550	0,194698736	0,55271304
KIF2C	2552	0,194589898	0,55472904
TRIM24	2595	0,191088974	0,5546397
BMPR1A	2601	0,190968975	0,55641544
SNRPA	2639	0,188489765	0,55655074
OIP5	2654	0,187507257	0,55783576
BOP1	2669	0,186790437	0,55911314
CCNB1	2697	0,185280234	0,5597187
ANGEL2	2725	0,183989003	0,56031066
SLC16A1	2732	0,183537424	0,56195706
ABCE1	2769	0,180967525	0,5620629
EXOSC9	2937	0,171743512	0,55546296
NUP107	2963	0,170564026	0,5560132
CDCA5	2965	0,170539379	0,55777377
FEN1	2973	0,169773683	0,55922353
MTHFD1	3058	0,1649483	0,5567381
PRKD3	3078	0,163899183	0,5575202
POLE2	3115	0,16175808	0,55742204
HSPD1	3123	0,16147317	0,5587837
ZNF589	3153	0,159673333	0,55901647
GABRA5	3225	0,156813949	0,5571004
SSB	3240	0,156123459	0,5580521
PUS1	3270	0,154910609	0,5582344
SEPHS1	3287	0,154185161	0,5590646
RNF138	3306	0,153344139	0,55978507
NCAPG2	3344	0,151831165	0,5595311
MRS2L	3350	0,151636511	0,5608892
PARP1	3368	0,15087536	0,5616338
MTA3	3427	0,148032635	0,56028026
FABP5	3448	0,147123888	0,56083375
NUDT21	3473	0,146009609	0,5611737
GYLTL1B	3498	0,144846663	0,56150126
MSH6	3532	0,143308699	0,5613585
ECT2	3550	0,142587006	0,5620152
TMEM177	3580	0,141302496	0,5620529
SFRS18	3648	0,138122469	0,5601401
PFAS	3679	0,136737749	0,5600789
PDCD2	3710	0,135743931	0,56000715
SFRS1	3755	0,134098351	0,55921173
DEK	3789	0,132830963	0,55895776
DLG3	3830	0,131311372	0,5583345
HNRPAB	3925	0,127383009	0,55494577
ISG20L1	3935	0,126911402	0,5558395
HMMR	3949	0,126518562	0,55652726
CDC2	3950	0,126504138	0,5578706
ILF3	3994	0,124957837	0,55702865
FOXO1	3996	0,124939837	0,55830497
MAT2A	3999	0,12491253	0,55953056
NOC3L	4033	0,123902179	0,55918175
PRIM2	4060	0,122839957	0,5591747
PRKX	4074	0,12242353	0,55981904
FXR1	4078	0,122312665	0,5609666
DLG7	4096	0,121840745	0,5614029
GNL3	4107	0,121449545	0,5621882
BXDC2	4137	0,120636985	0,5620065
MED14	4226	0,117603034	0,5588165
MGST1	4285	0,115648516	0,557119
TIA1	4500	0,108175538	0,5474733
MCM6	4567	0,105722249	0,54526687
PAICS	4645	0,103228889	0,5424791

PPM1B	4685	0,101809561	0,541593
PHF17	4828	0,096818551	0,5354585
SPC25	4849	0,095840938	0,53546745
RRM2	4857	0,09563794	0,53612995
PCDH1	4882	0,095076703	0,53592896
BIRC5	4892	0,0948034	0,53648174
EPRS	4923	0,093831941	0,53596497
GART	4938	0,093175709	0,5362482
AMD1	4987	0,091783367	0,5348017
BUB1B	5011	0,091017425	0,53460807
MARS	5133	0,087523267	0,5294341
PIM2	5156	0,087037265	0,52924865
CDC6	5267	0,084232911	0,5245946
LOC91431	5280	0,083891965	0,5248802
G3BP1	5292	0,083574109	0,5252128
PRPF40A	5370	0,081508942	0,5221944
JARID2	5389	0,081068508	0,52214736
DSG2	5397	0,080909878	0,52265346
ACTA1	5423	0,080307461	0,5222452
GPR64	5504	0,078085952	0,51903915
ACTC1	5516	0,077759072	0,51931
SLC7A3	5530	0,077276975	0,5194749
NANOG	5706	0,072432876	0,51141685
KPNA2	5725	0,072069928	0,5112742
PAK1	5914	0,067272358	0,50250566
MTHFD2	5977	0,065974884	0,50007886
MTF2	6156	0,061659861	0,4917551
PTTG1	6226	0,060016137	0,488912
NOL5A	6286	0,058676653	0,48655903
DDX18	6357	0,057011489	0,48363355
TFAM	6380	0,056432288	0,48312312
TDGF1	6396	0,056170367	0,482963
CLDN6	6495	0,054151531	0,47859478
CECR1	6550	0,053195816	0,47643584
RRP15	6571	0,052746408	0,47598714
ITPR3	6587	0,052510731	0,47578815
SLC38A1	6629	0,051551756	0,47426748
HSPA8	6736	0,049580336	0,46944723
CEBPZ	6746	0,04938497	0,46951768
PIPOX	6797	0,048396561	0,46750954
NOL11	6806	0,048095163	0,46761674
PYCR2	6864	0,047013056	0,46524084
PROM1	6889	0,046413954	0,4645231
AUTS2	6923	0,045661859	0,46334344
HPS3	7001	0,044213779	0,459929
ATP1A2	7012	0,044049248	0,45989233
ZNF195	7024	0,043779075	0,4598024
UTF1	7084	0,042573918	0,45727843
MRE11A	7196	0,040736042	0,45211205
UGP2	7242	0,039622292	0,45026293
CSE1L	7332	0,037928235	0,44617644
SALL1	7349	0,037514839	0,44576773
PAK1IP1	7397	0,036691081	0,44378662
RBM13	7475	0,035220739	0,44027665
KCNS3	7562	0,03367405	0,4362963
PRDM14	7572	0,033364434	0,43619663
NPM1	7609	0,032397155	0,43472478
GARNL4	7699	0,03074928	0,43056202
INDO	7734	0,030177139	0,42916748
TXNDC1	7742	0,030086981	0,4291339
USP9X	7786	0,029181195	0,4272748
CER1	7877	0,027518637	0,4230273
ETV4	8130	0,023175934	0,4105622
SFRS7	8148	0,022856334	0,4099474
CYP2S1	8157	0,022752151	0,40978548
TNPO3	8221	0,021417603	0,4068351
LECT1	8293	0,019983333	0,403466
SEMA6A	8324	0,019386273	0,40215862
NLN	8372	0,01833467	0,39998257
LIN28	8380	0,018197784	0,39982274
DDX21	8424	0,01723076	0,39783671
HSPC111	8485	0,016046522	0,39498064
DPPA4	8540	0,015215721	0,39241838
RAD54B	8554	0,01495974	0,39192152
MCM4	8636	0,013532989	0,38797948

KLKB1	8671	0,013053603	0,38640308
NUDT1	8823	0,010348686	0,37889633
COCH	9006	0,007725944	0,36979806
ROBO1	9008	0,007699593	0,3698294
LRRN1	9057	0,007025491	0,3674828
BUB3	9075	0,006694447	0,3666964
CTSC	9162	0,00523909	0,36241406
SCNN1A	9211	0,004330062	0,36003885
RARRES2	9265	0,0033693	0,35740125
PPP2R1B	9282	0,003127207	0,3566274
CNTNAP2	9469	2,18E-04	0,34724763
NLGN4X	9474	9,94E-05	0,3470469
MAL2	9516	-5,53E-04	0,34498468
PPP2R2B	9778	-0,005028412	0,33187288
IGF2BP2	10061	-0,009767778	0,31775215
USP44	10087	-0,010253864	0,3166
G3BP2	10112	-0,010564368	0,3155016
DHFR	10166	-0,011479476	0,3129501
FGF2	10182	-0,011693258	0,31231767
HSPA4	10210	-0,01209784	0,3110842
PNN	10283	-0,013283605	0,3075935
NFYB	10339	-0,0140132	0,30496803
CYP26A1	10371	-0,014421769	0,3035575
GINS4	10763	-0,020814972	0,28405595
RPS24	10794	-0,021352064	0,28276947
MYST2	10820	-0,021705613	0,28173894
LEFTY1	11005	-0,024525357	0,27271816
FGFR1	11178	-0,027198525	0,26433107
SLC13A3	11358	-0,029980127	0,25562042
WDHD1	11487	-0,032104399	0,24950486
SCGB3A2	11512	-0,032477155	0,24863915
ALPL	11632	-0,034571949	0,24300376
DNAJB6	11642	-0,034733336	0,24291863
ZIC3	11691	-0,035589382	0,24087538
ZNF267	12092	-0,041683566	0,22114149
LRIG1	12095	-0,041751992	0,22148398
HMGA1	12108	-0,042036306	0,22132507
RRAS2	12425	-0,047598913	0,20589106
ZSCAN10	12426	-0,047603462	0,20639658
SFRP1	12468	-0,0483016	0,2048414
VSNL1	12520	-0,049222153	0,2027916
ITGB1BP3	12552	-0,04975979	0,20175633
FAM46B	12561	-0,049982574	0,20188357
OLFM1	12668	-0,051828649	0,19708717
OAZ2	12783	-0,053790879	0,19190808
FOXD3	12837	-0,054583836	0,18981433
ADD2	13051	-0,058391184	0,17969039
EPHA1	13120	-0,05964566	0,17689377
WBSCR16	13231	-0,061901957	0,17200257
GNPTAB	13364	-0,06400308	0,16602397
GDF3	13705	-0,069648132	0,14961351
MAD2L2	13731	-0,07005351	0,1490964
ZNF770	13950	-0,073699318	0,13888282
L1TD1	14018	-0,075311981	0,136303
SFRP2	14048	-0,075999685	0,13564725
RABGAP1L	14451	-0,083962679	0,11626145
CRABP1	14490	-0,084566794	0,11524272
PITPNC1	14558	-0,085796595	0,11277424
CRMP1	14704	-0,088707887	0,106402256
BCL2L12	14709	-0,088819325	0,107143685
PPAP2A	14890	-0,09279196	0,09904962
DIDO1	15062	-0,096380442	0,09144764
SALL4	15396	-0,104591712	0,07576135
LOC157627	15432	-0,105617285	0,07511748
TMEFF1	15697	-0,112438142	0,062994964
AP1M2	15698	-0,112550288	0,064190164
LGALS8	15750	-0,113830417	0,06282645
PLCB3	16349	-0,130927861	0,034052867
POU5F1	16441	-0,133761391	0,030883148
MICB	16726	-0,144318029	0,018090349
NODAL	16849	-0,149301752	0,013521976
KLHL7	16856	-0,149711818	0,014809154
PMAIP1	16875	-0,150532782	0,015499755
SIRT1	16935	-0,153135613	0,014149899
ERBB2	16968	-0,154549852	0,014176979

DBNDD1	17025	-0,157171965	0,01302131
GPRC5B	17188	-0,164112881	0,006592565
TNNT1	17282	-0,168191656	0,003687587
SORL1	17401	-0,174152598	-4,15E-04
ZNF398	17685	-0,190017715	-0,012672184
ACTN3	17817	-0,199232295	-0,017164305
ESF1	17921	-0,207107395	-0,020160442
SRA1	17986	-0,212881237	-0,021128051
AZIN1	18009	-0,214424208	-0,019960737
SYNGR3	18333	-0,246239856	-0,033638414
CHST7	18466	-0,260471165	-0,037530668
TERF1	18521	-0,266996235	-0,0374192
SNRPN	18730	-0,297187209	-0,0447551
ETV1	18826	-0,314831823	-0,04620375
FANCL	18998	-0,345270157	-0,051162716
DNMT3A	19089	-0,363155991	-0,051845994
KIFC2	19139	-0,373129249	-0,05035527
BRWD1	19278	-0,404903293	-0,05301641
MAN2A1	19378	-0,440138221	-0,05333617
SALL2	19440	-0,458167076	-0,051547702
TRIM22	19654	-0,559201539	-0,056353413
NFE2L3	19745	-0,60878098	-0,05442834
RAB3B	19813	-0,67216754	-0,050669998
SPRY4	19834	-0,691820979	-0,044332206
GPR176	19837	-0,695502758	-0,037047375
GPC4	19980	-0,886371315	-0,034797452
TCF7L1	20000	-0,946539164	-0,025704304
IFITM1	20005	-0,95821774	-0,015730517
KAL1	20046	-1,079515457	-0,006284531
DMKN	20103	-1,26631856	0,004338103

ES2 GENESET SIGNIFICANTLY ENRICHED IN PC-3/Mc OVER PC-3/S (GSEA FDR q-value = 0.007)

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
ORC1L	315	0,651006043	0,08589875
HELLS	795	0,403102905	0,124996744
RRP9	818	0,398060858	0,18598603
NCAPH	1107	0,338622719	0,22450575
CDC25A	1152	0,329770237	0,27375284
CHEK2	1163	0,328042954	0,3244177
ERCC6L	1368	0,302793801	0,36151794
MYBL2	1391	0,299408078	0,40712148
MCM10	1477	0,289231569	0,44801167
HESX1	1555	0,280734688	0,4879737
DCC1	1892	0,245337337	0,50956285
ORC2L	2003	0,235480726	0,5408296
GPR19	2366	0,206998155	0,5551492
PRKD3	3078	0,163899183	0,5454289
SLC5A6	3686	0,136505648	0,536597
ISG20L1	3935	0,126911402	0,5440834
GJA7	4978	0,092163563	0,5067501
NANOG	5706	0,072432876	0,48197082
GOLGA7	5919	0,067197248	0,48193073
DTYMK	6341	0,057382677	0,46998882
TDGF1	6396	0,056170367	0,47606945
CLDN6	6495	0,054151531	0,4796518
ABHD9	7375	0,037029121	0,4418083
PRDM14	7572	0,033364434	0,4372857
ETV4	8130	0,023175934	0,41326025
LIN28	8380	0,018197784	0,40374225
CYP26A1	10371	-0,014421769	0,30724195
GPR23	10925	-0,023305664	0,28343523
RBM4	11030	-0,02500274	0,2821738
ZIC3	11691	-0,035589382	0,2549732
EPHA1	13120	-0,05964566	0,193414
GDF3	13705	-0,069648132	0,17529647
L1TD1	14018	-0,075311981	0,17155968
WSCD1	14736	-0,089452907	0,14993103
PWP2	15911	-0,118319489	0,11012672
POU5F1	16441	-0,133761391	0,10473745
DBNDD1	17025	-0,157171965	0,10031963
DNMT3A	19089	-0,363155991	0,054584935

Supplementary Table 5A. Genes for secreted proteins significantly overexpressed in PC-3/S cells relative to PC-3/Mc cells

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
INHBA	1	2,357851028	0,07263439
INHBB	5	1,968301415	0,13316041
SRGN	7	1,914293051	0,19212145
LIF	164	1,021191955	0,21580787
TBX3	166	1,020048022	0,24720244
PYCARD	173	0,9861812	0,27730325
DOPEY2	242	0,822589397	0,2992637
OPTN	609	0,519135177	0,29698238
SCG5	626	0,510695398	0,311926
ANG	678	0,486340076	0,32437035
NPHS2	865	0,422123492	0,32809085
LTBP2	968	0,393460661	0,3351242
STX1A	1035	0,377747953	0,34347168
ARFGAP3	1079	0,366979331	0,3526362
SLC22A4	1521	0,284558356	0,33937687
BACE2	1639	0,269941449	0,34185317
SCRN1	1680	0,264663815	0,34801355
ERGIC3	1752	0,256671876	0,35237885
STEAP2	1794	0,252602637	0,35811746
IL11	1847	0,247493744	0,36314905
PDIA4	1874	0,244005591	0,36937198
NFAT5	1910	0,239468902	0,37500545
FAM3B	1977	0,233389989	0,37890288
IFT88	2087	0,222183257	0,38030663
SNAP23	2209	0,212237835	0,38080433
SEC22A	2362	0,199864417	0,3793719
BET1	2453	0,193530664	0,3808416
NPHP1	2519	0,188820675	0,38341504
GBF1	2562	0,186605304	0,38706923
GOSR2	2622	0,183311418	0,3897726
SCAMP3	2631	0,183018789	0,39501476
SCAMP1	2783	0,174680293	0,39285597
ABCA1	2829	0,171726078	0,3959016
YKT6	3100	0,159596696	0,38733286
RER1	3121	0,158834532	0,39123002
COPA	3208	0,155214652	0,39171842
ARFGEF1	3528	0,142158151	0,3801642
COG3	3592	0,139572233	0,3813194
NLRP12	3599	0,139448419	0,38531837
CARD8	3799	0,132308975	0,37945545
SYT1	3842	0,130849943	0,38139087
EXOC5	3893	0,129236609	0,3828769
SEC22B	4054	0,124471828	0,37872073
BAIAP3	4172	0,12143039	0,37661898
AQP3	4464	0,113247685	0,36557236
GRM4	4499	0,112130605	0,3673304
SLC22A2	4606	0,109565236	0,36541244
ERGIC1	4682	0,107772082	0,36498785
TACR2	4689	0,107565798	0,368004
NPPB	4833	0,103684083	0,3640563
CLCNKA	4847	0,103217579	0,36658868
AQP5	4931	0,10127648	0,3655642
AP3B2	4949	0,100738041	0,36782032
GHRH	4984	0,099937499	0,3692025
NOD2	5018	0,098971985	0,37060484
KCNK5	5022	0,098846264	0,37350208
LMAN1	5029	0,098601714	0,37624186
APOA1	5068	0,097676061	0,3773545
VTI1B	5328	0,091971956	0,36725068
CANX	5845	0,081915401	0,34399775
VAMP3	5911	0,080452003	0,34323058
AQP7	5948	0,07980331	0,34389216
NMUR1	5989	0,078976937	0,34432843
NAPA	6105	0,076730981	0,34094867

CCL5	6179	0,075057521	0,33961552
CKLF	6275	0,073222004	0,33712673
SLC34A1	6353	0,071810082	0,33549368
PPY	6388	0,071079008	0,33598623
CLCNKB	6396	0,070987307	0,33782482
ERGIC2	6413	0,070780978	0,33920744
GCK	6503	0,069373235	0,33689976
NLRP2	6536	0,068956435	0,33742678
NMUR2	6666	0,066842757	0,3330428
ARFGEF2	6816	0,064204656	0,32757834
GUCA2B	7139	0,05835909	0,313291
SNAP29	7158	0,058112502	0,31418318
AQP9	7215	0,057097159	0,31314567
FOXP3	7230	0,056881767	0,31419975
KNG1	7448	0,053119823	0,30499643
INHA	7877	0,045599896	0,2850203
NRBP1	7894	0,045270223	0,28561652
DPH3	7906	0,045071509	0,28645638
SNAP25	8017	0,043042962	0,2822879
CCL8	8070	0,042202312	0,28099108
AQP6	8270	0,038939822	0,2722499
LMAN2L	8416	0,036761977	0,2661393
COPB2	8419	0,036735196	0,2671718
GOSR1	8487	0,035713796	0,26492557
CPLX2	8671	0,032559149	0,25678703
GHRL	8810	0,030327426	0,2508278
OSM	8884	0,029072566	0,2480771
CLDN16	9143	0,025188964	0,23596454
KCNJ1	9185	0,024509232	0,23467182
CADPS	9218	0,023987804	0,23381263
SYN3	9400	0,021217173	0,22542436
APOA2	9430	0,020778738	0,22461614
NLRC4	9625	0,017688159	0,21546964
COPB1	9854	0,013938477	0,20450899
FAM3D	9951	0,012507532	0,20009863
COG2	10339	0,006074292	0,18095231
CADM1	10412	0,005022243	0,17751019
SCTR	10515	0,003262799	0,17251511
CARTPT	10614	0,001388257	0,16766205
CCL3	10711	-3,51E-05	0,16286722
HRH3	10740	-5,79E-04	0,16148627
ACHE	10741	-5,85E-04	0,16150431
CIDEA	10763	-8,31E-04	0,16048083
STX7	10893	-0,002896358	0,1541256
SCNN1A	10978	-0,004330062	0,15006264
MON2	11049	-0,005602907	0,14673834
UNC13B	11057	-0,00571533	0,14656481
SNCAIP	11299	-0,009427727	0,13481568
TMED10	11397	-0,010786871	0,13030231
ARFIP1	11427	-0,01137169	0,1292041
ATP6V1B1	11540	-0,013356911	0,1240206
SCAMP2	11596	-0,014392073	0,12171659
GUCA1B	11621	-0,014709745	0,12097106
NPR1	11724	-0,016395129	0,116380796
SLC22A16	11772	-0,017335705	0,11456719
SEC22C	11914	-0,020396294	0,10815193
CPLX1	12016	-0,022388045	0,103796355
INS	12025	-0,022508629	0,104090564
AQP1	12040	-0,022832541	0,104095004
CRTAM	12122	-0,024507217	0,10080391
SLC22A18	12263	-0,026717644	0,094633475
ATP6V0A4	12441	-0,029886613	0,086712286
NLRP3	12575	-0,032291643	0,081063375
AQP2	12765	-0,036144733	0,072735615
ZW10	12796	-0,036751777	0,07236982
RAB14	12966	-0,03995413	0,06515864
COPZ1	13035	-0,041418668	0,06303833
CLCN5	13437	-0,04927187	0,044524234
SLC26A3	13775	-0,055771261	0,029407777
TPD52	13778	-0,055906098	0,031031253
PYDC1	14016	-0,061286584	0,021080568
RIMS1	14363	-0,069271155	0,005930651
COPG2	14376	-0,069552675	0,007475227
GHSR	14480	-0,072360538	0,004560225
AKAP3	14500	-0,072761834	0,005854026

NAPG	14628	-0,076534331	0,001868706
COPE	14779	-0,080689624	-0,003137544
PKDREJ	14784	-0,080801003	-8,47E-04
NLGN1	14847	-0,082153678	-0,001411415
CEL	15228	-0,092680082	-0,017538274
AVPR2	15232	-0,092758976	-0,014828711
STX16	15374	-0,097170666	-0,01887729
NPHS1	15377	-0,097350314	-0,015976233
DOPEY1	15786	-0,111509368	-0,032921463
MYO6	16182	-0,124698736	-0,04881066
RPH3AL	16230	-0,126151487	-0,04726985
VIP	16577	-0,139767408	-0,060246617
KIF1C	16808	-0,150407404	-0,06710032
ADORA2B	16818	-0,150780678	-0,062901884
GLMN	17403	-0,180268794	-0,08652003
SERGEF	17499	-0,185749397	-0,08553999
SYTL4	17726	-0,200785458	-0,09064087
COG7	17729	-0,200961038	-0,08454586
AQP4	17830	-0,20752874	-0,083144225
SCG2	18015	-0,221562892	-0,085506395
KRT18	18274	-0,24323912	-0,090897225
STX18	18324	-0,247619629	-0,08571189
RABEPK	18347	-0,24978362	-0,079110995
RAB2A	18382	-0,25368765	-0,07298924
DNAJC1	18408	-0,2561647	-0,06634151
ARL4D	18495	-0,265646726	-0,0624489
RAB3A	18512	-0,266840875	-0,055022445
GRHPR	18535	-0,269508928	-0,047813486
RAB26	19096	-0,340114117	-0,06530517
SCIN	19480	-0,424605995	-0,071349785
SCNN1B	19569	-0,450639218	-0,061854403
KCNMB4	19779	-0,565092087	-0,054875705
SCNN1G	19950	-0,723178685	-0,0410754
LIN7A	20167	-1,718165278	0,001098829

Supplemental Table 5B. Genes for cytokines significantly overexpressed in PC-3/S cells relative to PC-3/Mc cells

PROBE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES
INHBA	1	2,357851028	0,13277043
INHBB	5	1,968301415	0,24349774
SRGN	7	1,914293051	0,35128212
IL6	79	1,289369106	0,4203844
PYCARD	173	0,9861812	0,47131428
TGFB2	250	0,807879746	0,51304525
SIGIRR	414	0,63124311	0,54050165
CHRNA7	869	0,420848787	0,5416416
SMAD3	951	0,398430467	0,56005937
LTB	1370	0,313032389	0,5569154
TLR3	1699	0,262843281	0,5554178
ABCA1	2829	0,171726078	0,5089724
BCL3	3024	0,162838563	0,5085022
NLRP12	3599	0,139448419	0,48782578
TLR9	3603	0,139294878	0,49552327
CARD8	3799	0,132308975	0,49328357
SMAD4	4649	0,108603574	0,4572003
ATP6AP2	4832	0,103687696	0,4539945
IL12A	5009	0,099394798	0,45084512
NOD2	5018	0,098971985	0,45602268
AFAP1L2	5051	0,09809415	0,4599578
APOA1	5068	0,097676061	0,4646647
HIF1A	5348	0,091430873	0,4559469
PRG3	5431	0,089848772	0,45693222
TLR4	5649	0,08547432	0,45096073
NLRP2	6536	0,068956435	0,41080496
SOD1	6782	0,064768754	0,4022753
FOXP3	7230	0,056881767	0,38326058
IRF4	7725	0,048225168	0,36142203
IL9	7836	0,046344046	0,35856488
INHA	7877	0,045599896	0,3591453
MAST2	8669	0,032591868	0,32166323
GHRL	8810	0,030327426	0,31641266
CD40LG	8824	0,030043637	0,31745887
CD28	9081	0,026056338	0,30620173

TNFRSF8	9182	0,024549276	0,30261394
TLR7	9387	0,021430224	0,29368094
APOA2	9430	0,020778738	0,29276374
NLRC4	9625	0,017688159	0,28411704
TLR8	9814	0,014475322	0,2755876
SPN	9886	0,013473444	0,27281737
CADM1	10412	0,005022243	0,24700426
AZU1	10468	0,004041054	0,24449801
IFNG	10634	0,001049464	0,23635553
CIDEA	10763	-8,31E-04	0,2300399
CALCA	10768	-8,93E-04	0,22989136
IL18	10855	-0,00222572	0,22574195
CARD11	11394	-0,010743264	0,19960491
INS	12025	-0,022508629	0,16955762
CRTAM	12122	-0,024507217	0,16616628
NOD1	12231	-0,026067976	0,16226639
NLRP3	12575	-0,032291643	0,147036
IL12B	12889	-0,038671777	0,13365622
TLR1	13034	-0,041417867	0,12883155
SFTPD	13133	-0,043114353	0,12638897
IL27	13148	-0,04338561	0,12813704
PYDC1	14016	-0,061286584	0,08849363
GHSR	14480	-0,072360538	0,06955556
EREG	14591	-0,075645156	0,06834899
CD276	14681	-0,077999249	0,068318866
MALT1	14768	-0,080343321	0,068569906
EBI3	15029	-0,086941965	0,06054369
BCL10	15530	-0,102737941	0,041477658
TRAF2	16000	-0,118472718	0,024838889
TRAF6	16064	-0,121045329	0,02852597
IL4	16481	-0,135885581	0,015502545
CEBPG	16701	-0,145308942	0,012802172
GLMN	17403	-0,180268794	-0,011887522
IL17F	17433	-0,182121292	-0,003069938
MAP3K7	17533	-0,187404737	0,002565795
TLR6	17634	-0,194327772	0,008541804
CD24	20185	-2,102022886	1,99E-04

Supplementary Table 6. Expression of cell surface markers in parental PC-3 cells and derived clones PC-3/Mc and PC-3/S, determined by flow cytometry.

Surface Markers	CD24	CD44	CD40	CD71 (Tf receptor)
PC-3 (parental)	**4.0%**	**82.0%**	**24.3%**	**4.7%**
PC-3/Mc	**39.0%**	**90.0%**	**3.6%**	**64.5%**
PC-3/S	**2.4%**	**86.0%**	**44.0%**	**15.3%**

Supplemental Table 7. Enhanced bone and lymph node colonization of PC-3/Mc cells co-injected with PC-3/S cells

Metastasis Sites	Bone	Thoracic Lymph-nodes	Adrenal gland	Abdominal Lymph-nodes	No mets
PC-3/Mc	4/8	4/4	0/4	0/4	0/4
PC-3/S	0/8	0/4	1/4*	0/4	3/4
PC-3/Mc + PC-3/S	16/20	8/10	6/10	1/10	0/10

Note: PC-3/Mc colonized adrenal glands 25 days after inoculation, an organ that PC-3/Mc cells alone were not observed to colonize upon prolonged monitoring.
Asterisk: One out of 4 mice inoculated with PC-3/S cells via intracardiac injection had one adrenal gland (but no other locations) colonized with tumor cells 75 days after inoculation.

Supplemental Table 8

Genes in the ESC-like module most significantly enriched in PC-3/Mc cells, as determined by GSEA

BCAT1
ELOVL6
CDCA7
CCNA
LMNB1
PLK1
CRABP2
SLC2A1
BLM
BAX
TTK
CCNF
NDC80
CDCA3
TCOF1
AURKB

MCM2
CDC7
CDC20
RCC1
MYC
NCAPH
CHEK2
KIF11
PLK4
NTHL1
ERCC6L
SOX2
MYBL2
BUB1
CDCA8
CHEK1
CCNB2
MAD2L1
KIF4A
RACGAP1
MRPL12
KIF20A
AURKA
EXO1
TRIP13
MRTO4
TOP2A
NEK2
CHAF1A
MCM3
MRPL11
DTL
CDCA5
PUS1
CKS1B
GTSE1
PDCD2
BANF1
SPAG5
COX4NB
E2F3
RRM2
BIRC5
BUB1B
CDC6
SDHD
TRIP6
HN1
PIPOX
TGIF2
POLD1
MCM4
RPL27A
TGIF1

Supplemental Table 9. Immunohistochemical analysis of SOX2 expression in prostate cancer samples.

CASE CODE	% SOX2 (NUCL)	SOX2 (NUC) PREDOMINANT INTENSITY	GLEASON 1º	GLEASON 2º	GLEASON SCORE	TNM
1	0	0	3	4	7	T2A
2	0	0	2	3	5	T2AN0
3	0	0	2	3	5	T2AN0
4	0	0	3	3	6	T2AN0
5	0	0	4	2	6	T2AN0
6	0	0	3	4	7	T2AN0
7	0	0	3	4	7	T2AN0M0
8	70	1	3	4	7	T2AN0M0
9	0	0	3	3	6	T2AN0
10	0	0	4	3	7	T2AN0
11	70	1	3	4	7	T2AN0M0
12	0	0	3	4	7	T2AN0M0
13	80	1	2	3	5	T2C
14	0	0	3	2	5	T2C
15	0	0	2	3	5	T2C
16	0	0	3	3	6	T2C
17	0	0	3	4	7	T2C
18	0	0	3	4	7	T2C
19	20	1	3	4	7	T2C
20	0	0	3	4	7	T2C
21	0	0	3	4	7	T2C
22	0	0	3	4	7	T2C
23	0	0	4	3	7	T2C
24	0	0	3	4	7	T2C
25	0	0	3	4	7	T2C
26	0	0	3	4	7	T2C
27	0	0	4	5	9	T2C
28	5	1	4	3	7	T2C
29	70	1	3	2	5	T2CN0
30	0	0	2	3	5	T2CN0
31	0	0	3	3	6	T2CN0
32	20	1	3	3	6	T2CN0
33	0	0	3	3	6	T2CN0
34	0	0	3	3	6	T2CN0
35	0	0	3	4	7	T2CN0
36	0	0	3	4	7	T2CN0
37	0	0	4	3	7	T2CN0
38	80	1	3	4	7	T2CN0
39	50	1	3	4	7	T2CN0
40	0	0	3	4	7	T2CN0
41	5	1	4	3	7	T2CN0
42	<1	1	3	4	7	T2CN0
43	90	1	3	4	7	T2CN0
44	0	0	4	4	8	T2CN0
45	0	0	3	5	8	T2CN0
46	0	0	3	3	6	T2CN0M0
47	0	0	3	3	6	T2CN0M0
48	0	0	3	4	7	T2CN0M0
49	<1	1	3	4	7	T2CN0M0
50	0	0	3	4	7	T2CN0M0
51	5	1	3	4	7	T2CN0M0
52	0	0	4	3	7	T2CN0M0
53	0	0	3	4	7	T2CN0M0
54	70	1	3	4	7	T2CN0M0
55	0	0	3	4	7	T2CN0M0
56	0	0	4	3	7	T2CN0M0

58	0	0	3	2	5	T2CN0
59	0	0	4	2	6	T2CN0
60	0	0	2	3	5	T2CN0M0
61	5	1	3	2	5	T2CN0M0
62	0	0	4	3	7	T2CN0M0
63	50	1	3	4	7	T2CN0M0
64	30	1	3	4	7	T2CN0M0
65	0	0	3	4	7	T3A
66	20	1	3	4	7	T3A
67	0	0	3	4	7	T3A
68	0	0	3	4	7	T3AN0
69	10	1	4	3	7	T3AN0
70	<1	1	3	4	7	T3AN0
71	15	1	3	4	7	T3AN0
72	0	0	4	3	7	T3AN0
73	10	1	3	5	8	T3AN0
74	70	1	4	5	9	T3AN0
75	50	1	4	5	9	T3AN0
76	0	0	3	4	7	T3AN0M0
77	0	0	4	5	9	T3AN0M0
78	50	1	4	3	7	T3AN1
79	0	0	5	4	9	T3AN1
80	0	0	3	4	7	T3AN1M0
81	<1	1	3	4	7	T3AN0
82	5	1	4	3	7	T3AN0
83	80	1	4	3	7	T3B
84	30	1	4	3	7	T3BN0
85	<1	1	4	5	9	T3BN0
86	0	0	4	5	9	T3BN0
87	5	1	4	5	9	T3BN0
88	0	0	5	4	9	T3BN0
89	0	0	3	4	7	T3BN0M0
90	90	2	4	5	9	T3BN0M0
91	60	1	4	3	7	T3BN1
92	5	1	3	5	8	T3BN1
93	0	0	5	4	9	T3BN0
94	0	0	4	3	7	T3C
95	<1	1	5	4	9	
96	0	0	5	5	10	
97	0	0				
98	0	0				

Article 2
Nature Medicine
Published 07/08/11

Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization

Manav Korpal, Brian J Ell, Francesca M Buffa, Toni Ibrahim, Mario A Blanco, **Toni Celià-Terrassa**, Laura Mercatali, Zia Khan, Hani Goodarzi, Yuling Hua, Yong Wei, Guohong Hu, Benjamin A Garcia, Jiannis Ragoussis, Dino Amadori, Adrian L Harris & Yibin Kang

ARTICLES

nature medicine

Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization

Manav Korpal[1], Brian J Ell[1], Francesca M Buffa[2], Toni Ibrahim[3], Mario A Blanco[1], Toni Celià-Terrassa[1,4], Laura Mercatali[3], Zia Khan[5,6], Hani Goodarzi[1,6], Yuling Hua[1], Yong Wei[1], Guohong Hu[1], Benjamin A Garcia[1], Jiannis Ragoussis[7], Dino Amadori[3], Adrian L Harris[2] & Yibin Kang[1,8]

Although the role of miR-200s in regulating E-cadherin expression and epithelial-to-mesenchymal transition is well established, their influence on metastatic colonization remains controversial. Here we have used clinical and experimental models of breast cancer metastasis to discover a pro-metastatic role of miR-200s that goes beyond their regulation of E-cadherin and epithelial phenotype. Overexpression of miR-200s is associated with increased risk of metastasis in breast cancer and promotes metastatic colonization in mouse models, phenotypes that cannot be recapitulated by E-cadherin expression alone. Genomic and proteomic analyses revealed global shifts in gene expression upon miR-200 overexpression toward that of highly metastatic cells. miR-200s promote metastatic colonization partly through direct targeting of Sec23a, which mediates secretion of metastasis-suppressive proteins, including Igfbp4 and Tinagl1, as validated by functional and clinical correlation studies. Overall, these findings suggest a pleiotropic role of miR-200s in promoting metastatic colonization by influencing E-cadherin–dependent epithelial traits and Sec23a-mediated tumor cell secretome.

Early events of metastatic dissemination are thought to be initiated by epithelial-to-mesenchymal transition (EMT) in carcinoma cells, promoting tumor cell migration and invasion[1–3]. Although the involvement of EMT in cancer progression is widely recognized, the potential role of the reverse process, mesenchymal-to-epithelial transition (MET), is less clear. Histological analysis has revealed morphological similarities of primary tumors and metastatic lesions[4], and it has been reported that E-cadherin levels are elevated in lymph node metastases relative to matched primary tumor samples, both suggesting that EMT in primary tumors may be followed by MET at distant sites[5,6]. Despite these correlative clinical findings, rigorous functional studies linking MET with metastatic colonization ability are scarce.

miRNAs have been recognized as key regulators of normal and pathological processes in metazoan organisms[7–9]. Several miRNAs have also been shown recently to serve as promoters[10–15] or suppressors[16,17] of metastasis, particularly in the early step of tumor invasion. However, relatively little is known about the role of miRNAs in the late step, metastatic colonization of distant organs. The miR-200 family of miRNAs is involved in neurogenesis[18], regulation of embryonic and adult stem cells and cancer stem cells[19–22], chemosensitivity and apoptosis[23,24]. Notably, miR-200s have recently been shown to inhibit EMT and promote MET by direct targeting of E-cadherin transcriptional repressors *Zeb1* and *Zeb2* (refs. 25–28). The observation that the miR-200 family enforces the epithelial phenotype and inhibits EMT and invasion *in vitro* suggests that these miRNAs are likely to suppress metastasis. However, functional studies have yielded conflicting results in different models of metastasis[29–31], casting doubt on the potential therapeutic utility of miR-200s. Furthermore, it remains unclear whether the metastasis-related functions of the miR-200s are mediated entirely or only partially through the Zeb–E-cadherin axis.

Here we show that miR-200s promote metastatic colonization of breast cancer not only by influencing cell-intrinsic epithelial traits through targeting of the Zeb–E-cadherin axis, but also by altering tumor cell–derived secretome through targeting of the Sec23 homolog A (Sec23a)-mediated secretion of metastasis-suppressive proteins, including insulin-like growth factor binding protein 4 (Igfbp4) and tubulointerstitial nephritis antigen-like 1 (Tinagl1). These findings provide new insights into the molecular functions of miR-200s as well as the role of MET and the tumor secretome in metastasis.

RESULTS

Correlation of miR-200 expression with metastatic colonization

To investigate the clinical relevance of miR-200 expression (cluster 1: miR-200b, miR-200a and miR-429; cluster 2: miR-200c and miR-141; **Supplementary Fig. 1a**), we performed a retrospective analysis on a series of breast tumor samples (n = 210, Oxford collection)[32]. A composite miR-200 family expression score calculated in each sample as either the median (**Fig. 1a**; $P = 0.034$) or the mean ($P = 0.030$)

[1]Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA. [2]Cancer Research UK, Molecular Oncology Laboratories, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, UK. [3]Osteoncology Center, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori, Meldola, Italy. [4]Department of Cell Biology, Institut de Biologia Molecular de Barcelona, Consejo Superior de Investegaciones Cientificas, Barcelona, Spain. [5]Department of Computer Science, Princeton University, Princeton, New Jersey, USA. [6]The Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey, USA. [7]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [8]Genomic Instability and Tumor Progression Program, Cancer Institute of New Jersey, New Brunswick, New Jersey, USA. Correspondence should be addressed to Y.K. (ykang@princeton.edu).

Received 8 October 2010; accepted 18 May 2011; published online 7 August 2011; doi:10.1038/nm.2401

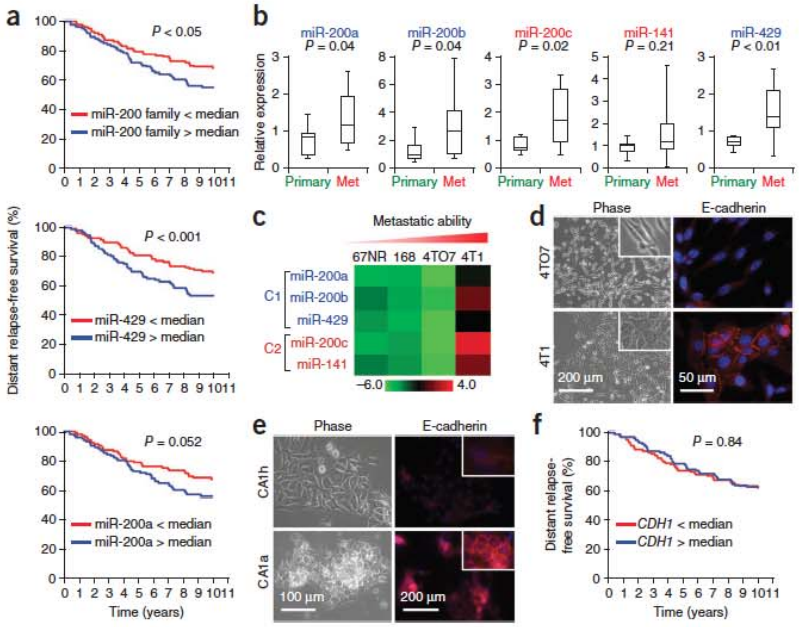


Figure 1 miR-200s are associated with poor prognosis in breast cancer. (**a**) Kaplan-Meier curves showing the DRFS of 210 subjects with high or low expression of the entire miR-200 family (top), miR-429 (middle) and miR-200a (bottom) in breast tumors. *P* values were computed by a likelihood-ratio test. (**b**) Box plots showing miR-200 expression levels in ten human primary and metastasis (met) samples as assessed by qRT-PCR analysis. Data (mean ± s.e.m.) are normalized to U6, and *P* values were computed by Student's *t* test. (**c**) Heat map showing miRNA expression levels in 4T1 series. 168, 168FARN. (**d**) Phase-contrast images (left) and immunofluorescence images of 4TO7 and 4T1 cells stained for E-cadherin (right). (**e**) Imaging as in **e** of MCFCA1h and MCFCA1a cells. Insets highlight the membrane localization of E-cadherin. (**f**) Kaplan-Meier curves showing the DRFS of 210 subjects with high or low *CDH1* expression. *P* values were computed by a likelihood-ratio test.

expression was significantly associated with poor distant relapse-free survival (DRFS)[33]. In particular, miR-429 and miR-200a showed the most significant association with DRFS (**Fig. 1a**; $P = 0.001$ for miR-429 and $P = 0.050$ for miR-200a). The miR-200 family expression was associated with estrogen receptor (ER)-positive status ($P = 0.019$, median expression) and correlated with poor DRFS only in the ER-positive tumors ($P = 0.028$, $n = 122$) and not in ER-negative tumors ($P = 0.48$, $n = 77$). Furthermore, profiling of miR-200 levels in ten human primary and lung-pleural metastasis samples, including six matched pairs (which we denote as the Meldola collection), revealed higher expression in metastases, reinforcing the potential role of miR-200s in metastatic colonization (**Fig. 1b**). These findings are consistent with a clinical correlation observed in individuals with serous ovarian carcinoma[34] and an earlier observation in xenograft studies[31].

To further investigate the importance of miR-200s in metastasis, we profiled miR-200 expression levels in three cancer cell line series that model the progression of breast and bladder cancers. The 4T1 series, including 67NR, 168FARN, 4TO7 and 4T1, are near-isogenic mouse mammary tumor cell lines. Of these lines, only 4T1 cells are capable of spontaneously metastasizing and colonizing distant organs after orthotopic implantation[35,36]. We found that miR-200s showed greatest expression in the highly metastatic 4T1 cells (**Fig. 1c** and **Supplementary Fig. 1b**), which was consistent with acquisition of epithelial traits (**Fig. 1d** and **Supplementary Fig. 2a,b**) in 4T1 compared to the weakly metastatic 4TO7 cells[37]. Similarly, elevated expression of miR-200s was also correlated with epithelial traits and greater metastatic ability for the MCF10A human breast cancer[38] (**Fig. 1e** and **Supplementary Fig. 1b**) and the TSU bladder carcinoma[39] progression series (**Supplementary Fig. 2c–e**, **Supplementary Table 1** and **Supplementary Results**). These observations collectively point to the possibility that MET induced by the miR-200 family may be crucial for successful completion of metastasis, particularly in the colonization step.

As E-cadherin is a key mediator of MET, we investigated the clinical relevance of E-cadherin expression in the Oxford cohort. As expected, miR-200 expression was positively correlated with E-cadherin expression ($P < 0.001$, Spearman correlation) and inversely correlated with Vimentin expression ($P < 0.001$, Spearman correlation). However, in contrast to the clinical associations observed for miR-200s, E-cadherin expression alone did not have any prognostic power (**Fig. 1f**), suggesting that the influence of miR-200s on breast cancer metastasis goes beyond the regulation of E-cadherin and the epithelial phenotype and probably involves genetic targets yet to be described.

miR-200 overexpression enhances lung colonization

To directly test the functional role of miR-200s in metastasis, we stably overexpressed cluster 1 (C1 line), cluster 2 (C2 line) and clusters 1 and 2 simultaneously (C1+C2 line) in the mesenchymal-like, weakly metastatic 4TO7 cells (**Supplementary Fig. 3a**). We also generated E-cadherin (CDH1)–overexpressing 4TO7 cells (CDH1 line) to test the importance of E-cadherin as a major downstream effector of miR-200s in metastasis. Notably, although ectopic expression of cluster 1 raised the expression of miR-200b, miR-200a and miR-429, cluster 2 elevated expression of all five miRNAs from both clusters (**Supplementary Fig. 3a**), possibly owing to a double-negative feedback mechanism involving Zeb factors[25,40]. As expected, the C1, C2 and C1+C2 lines expressed *CDH1* at higher levels and *Zeb1* at lower levels (**Fig. 2a,b** and **Supplementary Fig. 3b,c**) and adopted an epithelial-like phenotype (**Fig. 2b**). Notably, EMT inducers such as Snail and Twist, and mesenchymal markers such as N-cadherin and vimentin, remained unaffected upon miR-200 expression (**Fig. 2a,b**). The CDH1 line also showed elevated expression of E-cadherin but maintained a mesenchymal morphology (**Fig. 2a,b**). All the engineered lines had similar growth kinetics *in vitro* (data not shown) and *in vivo* (**Supplementary Fig. 3d**).

Parental 4TO7 cells, when inoculated orthotopically in the mammary fat pad, spontaneously disseminate to lungs but have a very low efficiency of colonization[35,36]. C2 ($P = 0.035$) and C1+C2 ($P = 0.01$) lines, both of which overexpress all five members of the miR-200 family, formed 10- to 30-fold more lung-derived tumor colonies relative to vector control cells, suggesting that ectopic miR-200 expression can enhance lung-colonization efficiency (**Fig. 2c**, **Supplementary Fig. 4a,b** and **Supplementary Methods**), consistent with an earlier report[31]. Of note, the C1 line, which expresses elevated levels of E-cadherin similarly to the C2 and C1+C2 lines, was incapable of efficiently colonizing lungs (**Fig. 2c**, $P = 0.38$). This implies that E-cadherin is probably not the only functional downstream effector

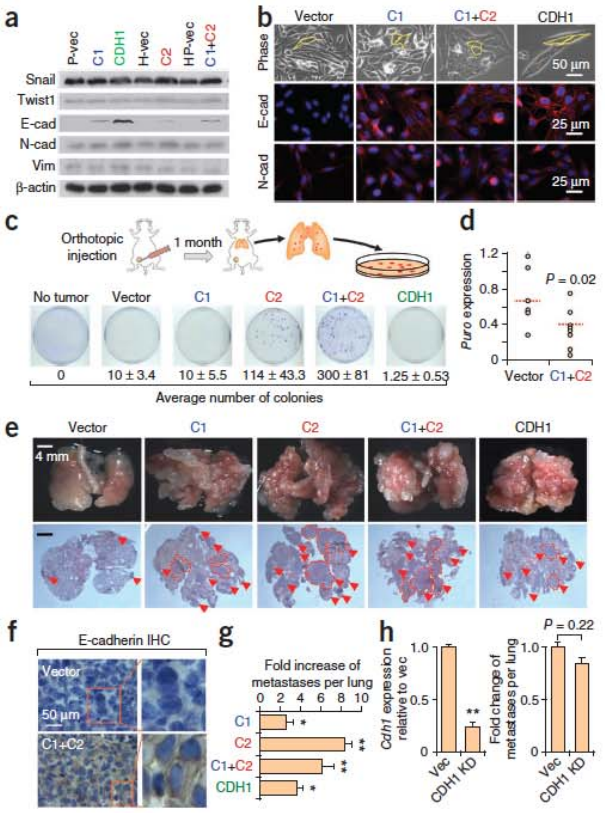


Figure 2 Ectopic miR-200 expression enhances spontaneous metastasis and colonization of distant organs. (**a**) Western blot showing expression of indicated proteins in various genetically modified 4TO7 cell lines. (**b**) Phase-contrast and immunofluorescence images of cell lines stained for E-cadherin (E-cad) and N-cadherin (N-cad). Yellow outline emphasizes cell morphology. (**c**) Plated colonies showing lung colonization by various cell lines used to generate orthotopic mammary gland tumors. Average numbers of colonies are listed below representative plate images. Data represent mean ± s.e.m. from a single representative experiment of three independent experiments. ($n = 9$ or 10). (**d**) Relative expression of puromycin-resistance gene, an indicator of circulating tumor cells, by qRT-PCR analysis of genomic DNA from whole-blood samples. Red dotted lines represent median values. $P = 0.02$ (Student's t test). (**e**) Representative gross lung and H&E-stained lung sections from mice intravenously injected with various 4TO7 cell lines. Red arrowheads and dashed lines mark metastatic nodules. Scale bar, 4 mm. (**f**) Immunohistochemical (IHC) staining for E-cadherin of lung nodules established from indicated cells. (**g**) Fold increase in number of pulmonary metastasis nodules for each group. Data represent mean fold increase ± s.e.m. from a single representative experiment of three independent experiments. ($n = 9$ or 10). (**h**) Left, RT-PCR showing expression of *Cdh1* in C1+C2 cells with or without stable *Cdh1* knockdown. Right, fold change in number of pulmonary lesions after intravenous inoculation of tumor cells. $^*P < 0.05$, $^{**}P < 0.01$ (Student's t test).

of miR-200s, and gene targeting mediated by all five members may be necessary to efficiently stimulate metastasis. In support of this, the CDH1 line also did not phenocopy the C2 and C1+C2 lines ($P = 0.08$), confirming that miR-200–mediated E-cadherin regulation is not sufficient to promote spontaneous metastatic colonization, but rather other genetic pathways must also be involved (**Fig. 2c** and **Supplementary Fig. 4a,b**).

As metastasis is a multistep process, we aimed to elucidate the influence of miR-200s on early and late steps of metastasis. To determine whether invasion or intravasation is regulated by miR-200s, we analyzed whole-blood samples for circulating tumor cells. In contrast to the results obtained from the lung-colony assays, miR-200 expression reduced tumor cell entry into circulation from primary tumors (**Fig. 2d** and **Supplementary Fig. 5a**), possibly through inhibition of EMT, migration and invasion, as observed *in vitro* (**Supplementary Fig. 5c**). Similarly, there was also reduced dissemination of CDH1 cells from primary tumors (**Supplementary Fig. 5b**). Together, these data suggested that although miR-200 expression can hinder entry of tumor cells into circulation, those that do intravasate may be more capable of colonizing distant organs.

To test this hypothesis, we inoculated cells directly into venous circulation in mice and measured the incidence of pulmonary metastasis. Tail-vein inoculation resulted in greater metastasis burden for all cell lines tested, with C2 ($P < 10^{-7}$) and C1+C2 ($P < 0.01$) lines showing the greatest metastasis potential (**Fig. 2e–g**). miR-200–overexpressing lines were only partly dependent on Cdh1 for colonization, as stable knockdown of *Cdh1* in C1+C2 lines modestly but insignificantly reduced colonization potential (**Fig. 2h**). In summary, our results from both clinical and experimental analyses collectively show that miR-200s can promote distant colonization of breast cancer cells. However, the fact that E-cadherin overexpression alone cannot fully recapitulate the metastasis potential of miR-200–overexpressing lines suggests that other genes or signaling pathways are also likely to be simultaneously targeted to enhance colonization efficiency.

miR-200 family induces global changes in gene expression

To identify such functional gene and pathway targets of miR-200s, we performed microarray analysis. Although we observed global changes in gene expression in C2 and C1+C2 lines, substantially less gene expression change was evident in C1 and CDH1 lines (**Fig. 3a**). To analyze the extent of the global changes in gene expression observed, we performed an unbiased hierarchical clustering based on a 1,218-gene-set signature that distinctively defines the parental 4TO7 and the highly metastatic 4T1 cell lines. Notably, we found that C2 and C1+C2 lines clustered with the highly metastatic 4T1 cells, whereas the C1 and CDH1 lines remained clustered with the parental 4TO7 cells (**Fig. 3a**), mirroring their spontaneous metastasis potentials (**Fig. 2c**). Rigorous gene-set enrichment analyses were also performed to (i) confirm the genome-wide shifts in gene expression, (ii) reveal that PicTar-derived miR-200 targets were regulated at a global level, (iii) highlight the global repression of the EMT-like genetic program upon miR-200 expression in 4TO7 cells, (iv) show that *Cdh1* overexpression alone causes only a modest expression shift, suggesting that E-cadherin is unlikely to be the only functional mediator of miR-200s and (v) reveal that these genetic changes may influence many cellular functions, including epithelial traits and protein transport and secretion (**Fig. 3b,c**, **Supplementary Fig. 6**, **Supplementary Table 2** and **Supplementary Results**). Furthermore, there was a significant negative association between miR-200 family expression and the miR-200–downregulated gene signature in the Oxford cohort of tumor samples ($P < 0.001$), suggesting that the global changes in gene expression observed *in vitro* are likely to be pathologically relevant in breast tumors.

Genomic and proteomic identification of miR-200 targets

Recent studies suggest that in mammalian cells, miRNAs regulate most of their direct targets at both mRNA and protein levels[41]. Therefore, we combined microarray and mass spectrometry (MS) analysis to identify

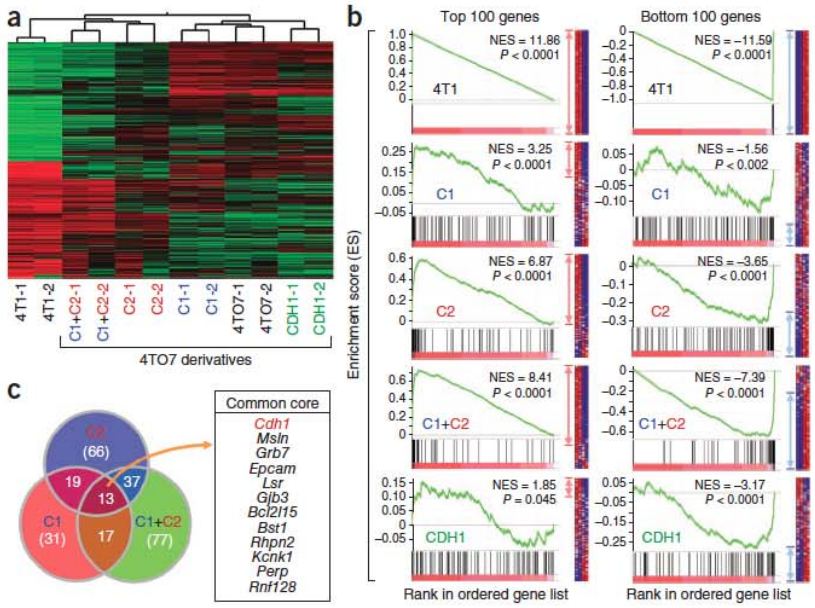


Figure 3 Ectopic miR-200 expression promotes global changes in gene expression. (**a**) Unsupervised clustering highlighting genome-wide changes in gene expression upon miR-200 expression in 4TO7 cells. Experiment was performed twice in duplicates. (**b**) Gene-set enrichment analysis showing influence of miR-200 overexpression on the overall gene expression profile of 4TO7 cells. Gene sets used are the top 100 (left) and bottom 100 (right) differentially expressed genes in the test lines (C1, C2, C1+C2 and CDH1) compared with control lines. The gene list used included all mouse genes, ranked by their differential expression between 4T1 and 4TO7 variants. Enrichment of top and bottom 100 genes from 4T1 compared with 4TO7 (ranked list) is shown as an example of maximum possible enrichment. NES, normalized enrichment score. Red and blue double-sided arrows denote the relative number of core genes for each analysis. (**c**) Venn diagram showing substantial overlap of core genes from top 100 gene sets for C1 (red circle), C2 (blue circle) and C1+C2 (green circle) lines from **b**. Core genes shared among all three lines are listed. *Cdh1* is listed in red to emphasize the positive influence of miR-200s on E-cadherin expression.

nine candidate genes downregulated at both the RNA and protein level in C1+C2 cells compared with controls (**Fig. 4a,b**, **Supplementary Table 3**, **Supplementary Results** and **Supplementary Methods**). We confirmed reduced expression of the nine candidate genes in C1+C2 cells by quantitative RT-PCR (qRT-PCR) analysis (**Fig. 4c**) and further validated the findings in MDA-MB-231 human breast cancer and TSU-PR1 bladder cancer cell lines, respectively (**Fig. 4d**). Furthermore, miR-200b and miR-200c levels were negatively correlated with mean expression of the nine candidate genes in the NCI-60 panel of cell lines (**Fig. 4e**), implying conserved targeting.

To assess which of these candidate genes are directly targeted by miR-200s, we cloned the 3′ untranslated regions (3′ UTRs) of eight of the nine candidates for standard luciferase assays. We confirmed direct targeting of three candidates, including cofilin 2 (*Cfl2*), low-density lipoprotein receptor-related protein 1 (*Lrp1*) and *Sec23a* (**Fig. 4f**). The 3′ UTRs of *Cfl2* and *Lrp1* each contain one functional miR-200 target sequence, whereas *Sec23a* contains two evolutionarily conserved miR-200 target sites that function cooperatively to suppress Sec23a expression (**Fig. 4f** and **Supplementary Fig. 7**).

Sec23a is a miR-200 target that suppresses metastasis

To test which of the three direct targets most probably mediates miR-200 function in metastasis, we stably knocked down each of the genes in 4TO7 cells (**Supplementary Fig. 8a–c**). Stable knockdown did not influence morphology (**Supplementary Fig. 8a**), cell proliferation or E-cadherin expression (data not shown). Of the three genes, *Sec23a* knockdown reduced transwell migration (**Fig. 5a**) and increased pulmonary colonization (**Fig. 5b,c**), phenocopying miR-200 overexpression (**Supplementary Fig. 5c** and **Fig. 2e,f**). In support of the functional data, endogenous *Sec23a* transcripts were less abundant in highly metastatic 4T1 and MCFCA1a lines than in the weakly metastatic 4TO7 and MCFCA1h lines and were reduced in the 4TO7 line upon miR-200 overexpression (**Supplementary Fig. 8d,e**). Furthermore, *SEC23A* levels were significantly lower in clinical metastases relative to primary tumors (**Fig. 5d,e**), consistent with *SEC23A*'s role as a suppressor of metastatic colonization.

Neither knockdown in 4TO7 cells nor overexpression of Sec23a in 4TO7-C1C2 cells significantly ($P > 0.05$) influenced spontaneous metastasis (**Supplementary Fig. 8f–h**). As Sec23a inhibits migration but increases colonization, it is possible that these two effects may balance each other out. Indeed, evaluation of *SEC23A* expression as a primary tumor prognostic marker in a large public clinical database of breast cancer[42] did not reveal significant associations ($P > 0.05$). Overall, both clinical data analysis and experimental animal models suggest that Sec23a is involved in suppressing metastasis, specifically at the step of colonization. However, it should be noted that although *Sec23a* knockdown can enhance metastatic colonization, overexpression of Sec23a alone is not sufficient to significantly ($P > 0.05$) suppress metastasis (**Supplementary Fig. 8i,j**).

miR-200 overexpression suppresses Sec23a-mediated secretion

Sec23a is an essential component of COPII vesicles and is involved in anterograde transport of proteins from the endoplasmic reticulum to the Golgi apparatus. Although several recent studies have shown that Sec23a is indispensible for the secretion of extracellular matrix components such as collagens and cartilage oligomeric matrix protein (Comp), and in craniofacial and chondrocyte development[43–45], little is known of other classes of proteins influenced by Sec23a-mediated secretion and their potential role in metastasis. To clarify this, we performed MS analysis on conditioned medium from two different *Sec23a*-knockdown 4TO7 lines. There was a global reduction in secretion of proteins (**Supplementary Fig. 9a**), including those involved in wounding ($P = 6.4 \times 10^{-6}$), cell adhesion ($P = 1.6 \times 10^{-5}$), extracellular structure organization ($P = 3.8 \times 10^{-3}$) and inflammatory and immune responses ($P = 1.0 \times 10^{-4}$ and $P = 4.0 \times 10^{-4}$, respectively) (**Supplementary Table 4**) as defined by the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources[46]. Reduced secretion led to intracellular accumulation of proteins such as Comp[44] (**Supplementary Fig. 9b**), which in turn resulted in modest endoplasmic reticulum distension as assessed by electron microscopic (EM) analysis (**Supplementary Fig. 9c**). There were strong correlations in the levels of Sec23-dependent

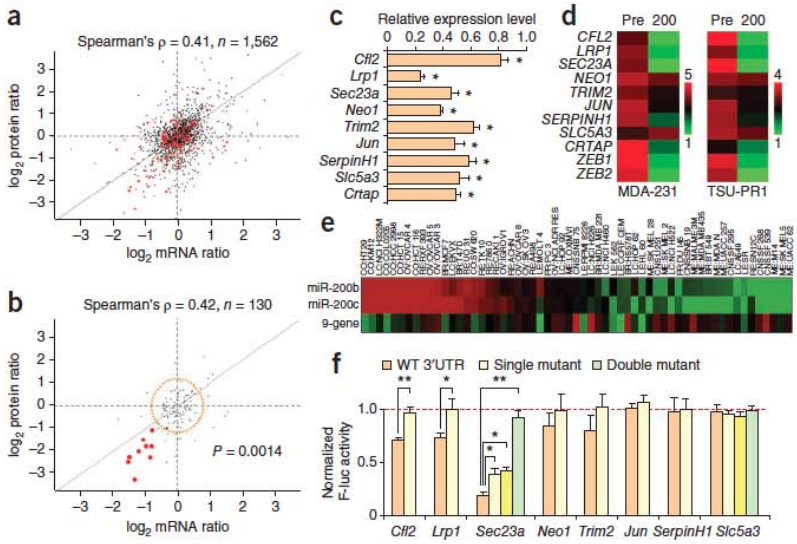


Figure 4 Identification of putative miR-200 targets using MS. (**a**) Protein abundance compared with mRNA abundance, relative to control cells, for 1,562 genes in C1+C2 cells. Red dots represent genes with miR-200 target sites. (**b**) Protein and mRNA abundances as in **a**, for only genes containing miR-200 target sites ($n = 130$). Red dots denote significantly reduced expression at both mRNA and protein levels ($n = 9$). Orange dotted circle represents little or no difference from gene expression in control cells. (**c**) qRT-PCR validation of reduced expression in C1+C2 cells, relative to control cells, for the nine candidate genes highlighted as red dots in **b**. Data represent mean ± s.e.m. $*P < 0.05$ (Student's t test). (**d**) Heat map showing expression (expressed as fold difference) of the nine candidate genes in MDA-MB-231 (left) and TSU-PR1 (right) cells upon transient transfection of miR-200s (200) relative to pre-miR controls. Pre, control pre-miRNA. *ZEB1* and *ZEB2* were included as positive controls. (**e**) Heat map showing average expression (fold difference) of the nine–candidate gene signature (9-gene) and miR-200b and miR-200c in NCI-60 panel of cell lines. (**f**) Luciferase assays in HeLa cells testing direct targeting of eight of nine candidate genes by miR-200s. Data represent percentage difference in normalized luciferase activity upon cotransfection of miR-200s, relative to transfection with the negative control pre-miRNA (mean ± s.e.m.). $*P < 0.05$, $**P < 0.01$ (Student's t test).

secreted proteins between the two knockdown lines Sec23a-KD2 and Sec23a-KD3 (**Fig. 6a**; $R = 0.9$, $P < 0.0001$) and between Sec23a-KD2 and the C1+C2 line (**Fig. 6a**; R=0.43; $R = 0.43$, $P < 0.001$). Similarly to the Sec23a-KD2 cell line, the C1+C2 line also showed modest endoplasmic reticulum distension by EM analysis, suggesting functional disruption of protein transport (**Supplementary Fig. 9c**). Together, these data suggest that expression of Sec23a and the secretory pathway are inhibited by miR-200s.

Igfbp4 and Tinagl1 are Sec23a-dependent metastasis suppressors

To evaluate the clinical relevance of the secretome in cancer progression in humans with breast cancer, we used a large public microarray database[42] to analyze the expression of 35 of 38 genes whose products were reduced in secretion in both *Sec23a* knockdown lines (secretome gene signature; **Fig. 6a**). Low expression was significantly correlated with low relapse-fee survival (RFS; **Fig. 6b**; $P = 0.0236$), reinforcing the idea that these genes are metastasis suppressors. In contrast, average expression of eight of nine candidate miR-200 target genes identified through combined genomic and proteomic analysis (**Fig. 4a,b**) correlated with high RFS, but the correlation did not reach significance ($P = 0.13$, data not shown). This highlights the prominent influence of the Sec23a-mediated secretome in metastasis-free survival.

To uncover the central mediators of the secretome that have metastasis-related functions, we functionally tested three of six genes—*Axl* (Axl receptor tyrosine kinase), *Tinagl1* and *Igfbp4*—from the secretome signature that were significantly ($P < 0.05$) associated with RFS (**Supplementary Fig. 9d**). Stable knockdown of *Tinagl1* and *Igfbp4* in 4TO7 cells (**Supplementary Fig. 9e**) increased colonization after intravenous injection (**Fig. 6c,d**), phenocopying miR-200 overexpression and *Sec23a* knockdown. Notably, both *TINAGL1* and *IGFBP4* were also significantly associated with better distant metastasis-free survival in an independent EMC286 clinical data set[47] (**Fig. 6e**), again

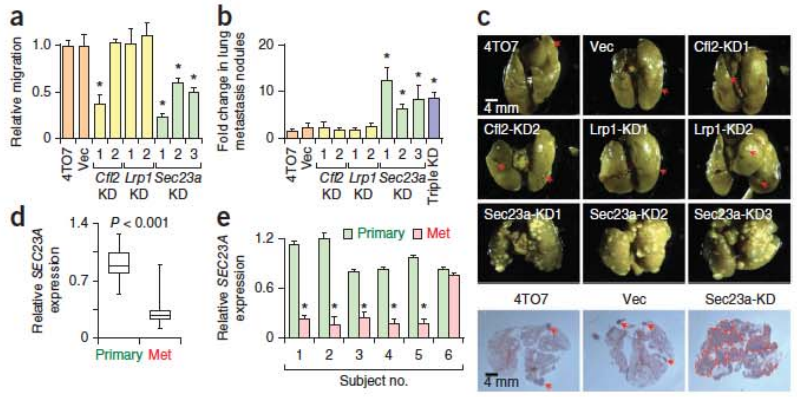


Figure 5 *Sec23a* knockdown phenocopies miR-200s in inhibiting migration and promoting metastatic colonization. (**a**) Transwell migration assays. Shown are ratios of migration of knockdown (KD) lines over migration of parental 4TO7 cells (mean ± s.e.m. from triplicate experiments. (**b**) Fold change in number of pulmonary nodules relative to 4TO7 parental line (mean ± s.e.m.). 'Triple KD' denotes knockdown of all three genes. (**c**) Representative gross lung images and H&E-stained lung sections (bottom) from mice intravenously injected with indicated cell lines (KD1 and KD2 denote two different knockdown lines for each gene). Red arrows mark metastatic nodules, except in *Sec23a* knockdown samples, which contained large numbers of nodules that were outlined with red dashed lines. (**d,e**) Relative *SEC23A* expression in ten human primary tumors compared with expression in ten lung metastases (box plots show 25th, 50th and 75th percentiles (horizontal bars) and 1.5 interquartile ranges (error bars)) (**d**), and in matched primary and lung metastasis samples collected from six individuals (**e**). *GAPDH* was used to normalize expression. Error bars show s.e.m. $*P < 0.05$ (Student's t test).

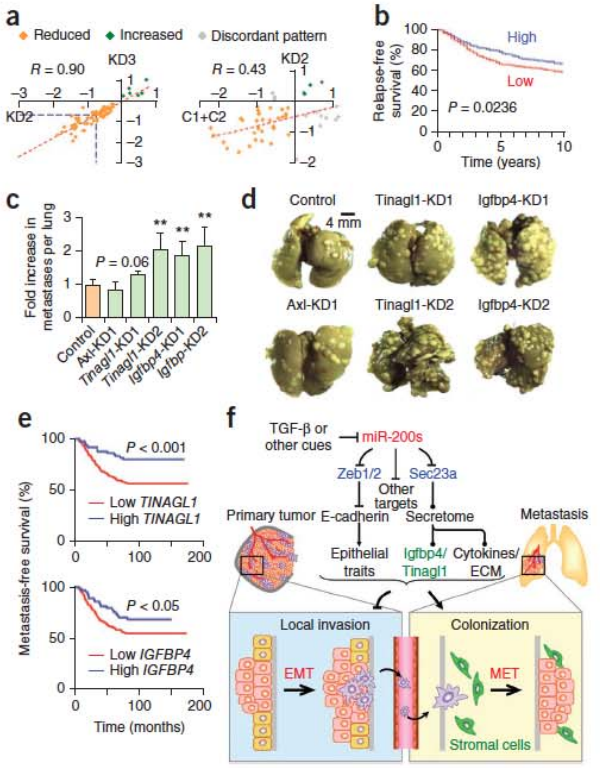


Figure 6 Sec23a knockdown disrupts secretion of proteins that are correlated with suppression of clinical metastasis. (**a**) Correlation of secretome profiles between two different *Sec23a* knockdown lines (Sec23a-KD2 and Sec23a-KD3) and between Sec23a-KD2 and C1+C2 lines. Proteins in common between different lines were used to generate the plots. Orange, proteins less abundant in both lines; green, more abundant in both lines; gray, discordant expression patterns. (**b**) Kaplan-Meier curves showing RFS of subjects with high or low median expression of 35 genes whose secreted products were reduced in *Sec23a*-knockdown lines. (**c**) Fold increase in number of pulmonary metastases in 4TO7-derived lines with stable knockdown of *Axl*, *Tinagl1* or *Igfbp4*, relative to vector control (KD1 and KD2 signify different knockdown lines). (**d**) Representative gross lung images from animals injected via lateral tail vein with knockdown lines from **c**, along with vector control. $^{**}P < 0.01$ (Student's *t* test). (**e**) Kaplan-Meier plots of distant metastasis-free survival of patients in the EMC286 data set stratified by expression of *TINAGL1* (top) or *IGFBP4* (bottom). *P* values were computed by log-rank test. (**f**) Schematic model of miR-200 function during metastasis. miR-200s simultaneously target several genes including *Zeb1* and *Zeb2* (*Zeb1/2*) and *Sec23a* to inhibit local invasion but promote metastatic colonization. Targeting of *Zeb1/2* influences cell-intrinsic epithelial traits, whereas targeting of *Sec23a* modulates tumor-derived secretion of factors such as Igfbp4 and Tinagl1, which influence metastatic colonization by altering tumor-stromal interactions.

supporting their role in reducing metastasis. Furthermore, *IGFBP4* expression showed a significant ($P < 0.05$) lung-tropic association (**Supplementary Fig. 9f**) with metastases in the MSK82 clinical data set[48]. Together, our data imply that the secretome is a major downstream mediator of miR-200s in metastasis and reveal that Tinagl1 and Igfbp4 are secreted suppressors of lung metastasis with potential therapeutic applications.

DISCUSSION

In this study, we show that miR-200s promote metastatic colonization through mechanisms that go beyond cell-intrinsic regulation of epithelial traits through the Zeb–E-cadherin axis (**Fig. 6f**). miR-200s also influence tumor cell secretome by direct targeting of the Sec23a-mediated transport pathway, which affects cell-extrinsic tumor-stromal interactions. We further identified Tinagl1 and Igfbp4 as key Sec23a-mediated secretory proteins that substantially reduce metastatic colonization. Our findings support dynamic roles for EMT and MET during different stages of metastasis: whereas EMT and low levels of miR-200s promote invasion and intravasation, MET and high miR-200 expression is required for efficient colonization of secondary organs.

This is, to our knowledge, the first large-scale breast cancer clinical study showing a clear clinical association of miR-200 family expression with poor DRFS, particularly in ER-positive breast cancers. Furthermore, miR-200 expression was greater in lung metastases than in primary tumors. In support of the clinical associations, a survey across several different isogenic series of cancer cell lines revealed a strong correlation between miR-200 expression and metastatic ability, which was further supported by functional analysis *in vivo*.

Several recent studies in animal metastasis models have reported conflicting roles of miR-200s in metastatic progression[29–31]. These contradictory findings may in part be due to the different models applied in these studies: miR-200s may promote metastasis of breast cancer cells[31] and hinder metastasis of lung adenocarcinoma[29] and pancreatic neuroendocrine cells[30]. However, it is more likely that the net effect of miR-200s in metastasis progression of various cancers depends on several variables across different models, such as the difference in the rate-limiting step of the metastatic cascade and the importance of MET in colonization in different models. In the model system we applied, the 4TO7 tumor cells are inherently highly migratory and invasive *in vitro* and *in vivo*, and although ectopic miR-200 expression reduces dissemination, it has a net effect of enhancing the rate-limiting colonization step. In contrast, this outcome is less likely in model systems where earlier steps of metastasis, such as early dissemination through acquisition of EMT-like properties, is the rate-limiting process, as may be the case for the model systems showing metastasis-suppressor functions of miR-200s (refs. 29,30) (see **Supplementary Discussion** for further discussion on the biphasic role of miR-200s).

Here both clinical and functional data revealed that E-cadherin alone is insufficient to recapitulate all miR-200 phenotypes, suggesting that miR-200s probably influence other genes or pathways during metastasis. We used integrated genomic and proteomic analysis to comprehensively examine the impact of miR-200s on gene expression in cancer cells. miR-200 overexpression led to widespread changes in gene expression toward that of the most metastatic cells. In addition to genes and pathways controlling epithelial characteristics, other cellular processes such as protein transport and secretion were also implicated in miR-200–dependent regulation of metastatic ability. Notably, we showed that the expression of Sec23a is directly suppressed by miR-200s through two evolutionarily conserved target sequences in its 3′ UTR (see **Supplementary Discussion** for a summary of evidence supporting Sec23a as a functional target of miR-200s).

Thus, in addition to the well-established function of miR-200s in regulating intrinsic cellular properties through Zeb1, Zeb2 and E-cadherin, miR-200s may also have the potential to influence the microenvironment via directed targeting of Sec23a-dependent secretome. Through regulation of this pathway, miR-200s may extend their

reach to manipulate a neighboring population of tumor and stromal cells, influencing these cells' collective behavior during metastasis. The ability of miRNAs to influence cell-intrinsic and cell-extrinsic properties of tumor cells has also recently been uncovered for the miR-17/20 cluster [49–51]. The miR-200 family and the miR-17/20 cluster may represent a growing number of miRNAs known to influence both intracellular regulatory machineries and intercellular communication of tumor and stromal cells (see **Supplementary Discussion** for further comments on the role of secreted components in metastasis).

In summary, we combined clinical and experimental studies to establish a biphasic role for miR-200s in metastasis. We showed that miR-200s promote metastatic colonization by enhancing cell-intrinsic epithelial traits via the Zeb–E-cadherin axis and by inhibiting Sec23a-dependent regulation of tumor secretome. The fact that miR-200 targeting of Sec23a seems to have dichotomous roles in metastasis, hindering early steps of migration and invasion while promoting the late step of metastatic colonization, may explain the contradictory roles of miR-200s in various models of metastasis. Although miR-200s and the Sec23a pathway may represent new therapeutic opportunities for metastatic cancer, their dichotomous functions warrant careful assessment of potential therapeutic benefits and adverse side effects when treatments are applied at different stages of the disease.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemedicine/.

Accession codes. Gene expression microarray data used to analyze global changes in gene expression induced by miR-200 family or Cdh1 alone have been deposited at the NCBI Gene Expression Omnibus with the accession code GSE19631.

Note: Supplementary information is available on the Nature Medicine website.

ACKNOWLEDGMENTS

We thank S. Kyne at the mass spectrometry core facility, D. Storton and J. Buckles at the microarray core facility, and E. Williams at the electron microscopy facility of Princeton University for expert technical advice and support, N. Sethi for insightful discussions and blinded verification of tumor counts, S.J. Parkinson for technical advice, and K. Socha for colony counting. We also thank E. Williams (Monash Institute of Medical Research) and F. Miller (Barbara Ann Karmanos Cancer Institute) for the TSU-PR1 and 4T1 cell line series, respectively. Y.K. is a Champalimaud Investigator and a Department of Defense Era of Hope Scholar Award recipient. This research was supported by grants from the US National Institutes of Health (1R01-CA141062) and the Brewster Foundation to Y.K. and from Cancer Research UK, Oxford NHS Biomedical Research Centre and Friends of Kennington Cancer Fund to A.L.H. M.K. and Y.H. are recipients of US Department of Defense predoctoral fellowships.

AUTHOR CONTRIBUTIONS

M.K. and Y.K. designed experiments. M.K., B.J.E., T.C.-T. and Y.H. performed the experiments. F.M.B., T.I., L.M., J.R., D.A. and A.L.H. provided clinical samples and associated analyses. M.A.B., Z.K., H.G., Y.W., G.H. and B.A.G. contributed genomic and proteomic analyses. M.K. and Y.K. wrote the manuscript. All authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemedicine/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Thiery, J.P., Acloque, H., Huang, R.Y. & Nieto, M.A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
2. Thompson, E.W. & Williams, E.D. EMT and MET in carcinoma—clinical observations, regulatory pathways and new models. *Clin. Exp. Metastasis* **25**, 591–592 (2008).
3. Yang, J. & Weinberg, R.A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev. Cell* **14**, 818–829 (2008).
4. Chaffer, C.L., Thompson, E.W. & Williams, E.D. Mesenchymal to epithelial transition in development and disease. *Cells Tissues Organs* **185**, 7–19 (2007).
5. Jeschke, U. *et al.* Expression of E-cadherin in human ductal breast cancer carcinoma in situ, invasive carcinomas, their lymph node metastases, their distant metastases, carcinomas with recurrence and in recurrence. *Anticancer Res.* **27**, 1969–1974 (2007).
6. Park, D., Karesen, R., Axcrona, U., Noren, T. & Sauer, T. Expression pattern of adhesion molecules (E-cadherin, α-, β-, γ-catenin and claudin-7), their influence on survival in primary breast carcinoma, and their corresponding axillary lymph node metastasis. *APMIS* **115**, 52–65 (2007).
7. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
8. Esquela-Kerscher, A. & Slack, F.J. Oncomirs—microRNAs with a role in cancer. *Nat. Rev. Cancer* **6**, 259–269 (2006).
9. Johnson, R. *et al.* A microRNA-based gene dysregulation pathway in Huntington's disease. *Neurobiol. Dis.* **29**, 438–445 (2008).
10. Huang, Q. *et al.* The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. *Nat. Cell Biol.* **10**, 202–210 (2008).
11. Ma, L., Teruya-Feldstein, J. & Weinberg, R.A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449**, 682–688 (2007).
12. Ma, L. *et al.* miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat. Cell Biol.* **12**, 247–256 (2010).
13. Asangani, I.A. *et al.* MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **27**, 2128–2136 (2008).
14. Zhu, S., Si, M.L., Wu, H. & Mo, Y.Y. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *J. Biol. Chem.* **282**, 14328–14336 (2007).
15. Zhu, S. *et al.* MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. *Cell Res.* **18**, 350–359 (2008).
16. Tavazoie, S.F. *et al.* Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* **451**, 147–152 (2008).
17. Valastyan, S. *et al.* A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. *Cell* **137**, 1032–1046 (2009).
18. Karres, J.S., Hilgers, V., Carrera, I., Treisman, J. & Cohen, S.M. The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in *Drosophila*. *Cell* **131**, 136–145 (2007).
19. Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
20. Wellner, U. *et al.* The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* **11**, 1487–1495 (2009).
21. Shimono, Y. *et al.* Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* **138**, 592–603 (2009).
22. Iliopoulos, D. *et al.* Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Mol. Cell* **39**, 761–772 (2010).
23. Schickel, R., Park, S.M., Murmann, A.E. & Peter, M.E. miR-200c regulates induction of apoptosis through CD95 by targeting FAP-1. *Mol. Cell* **38**, 908–915 (2010).
24. Cochrane, D.R., Howe, E.N., Spoelstra, N.S. & Richer, J.K. Loss of miR-200c: a marker of aggressiveness and chemoresistance in female reproductive cancers. *J. Oncol.* **2010**, 821717 (2010).
25. Burk, U. *et al.* A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582–589 (2008).
26. Gregory, P.A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10**, 593–601 (2008).
27. Korpal, M., Lee, E.S., 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. *J. Biol. Chem.* **283**, 14910–14914 (2008).
28. Park, S.M., Gaur, A.B., Lengyel, E. & Peter, M.E. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* **22**, 894–907 (2008).
29. Gibbons, D.L. *et al.* Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev.* **23**, 2140–2151 (2009).
30. Olson, P. *et al.* MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. *Genes Dev.* **23**, 2152–2165 (2009).
31. Dykxhoorn, D.M. *et al.* miR-200 enhances mouse breast cancer cell colonization to form distant metastases. *PLoS ONE* **4**, e7181 (2009).
32. Camps, C. *et al.* hsa-miR-210 Is induced by hypoxia and is an independent prognostic factor in breast cancer. *Clin. Cancer Res.* **14**, 1340–1348 (2008).
33. Hudis, C.A. *et al.* Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J. Clin. Oncol.* **25**, 2127–2132 (2007).
34. Nam, E.J. *et al.* MicroRNA expression profiles in serous ovarian carcinoma. *Clin. Cancer Res.* **14**, 2690–2695 (2008).
35. Aslakson, C.J. & Miller, F.R. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res.* **52**, 1399–1405 (1992).

36. Yang, J. *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927–939 (2004).
37. Lou, Y. *et al.* Epithelial-mesenchymal transition (EMT) is not sufficient for spontaneous murine breast cancer metastasis. *Dev. Dyn.* **237**, 2755–2768 (2008).
38. Santner, S.J. *et al.* Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. *Breast Cancer Res. Treat.* **65**, 101–110 (2001).
39. Chaffer, C.L. *et al.* Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2. *Cancer Res.* **66**, 11271–11278 (2006).
40. Bracken, C.P. *et al.* A double-negative feedback loop between ZEB1–SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res.* **68**, 7846–7854 (2008).
41. Guo, H., Ingolia, N.T., Weissman, J.S. & Bartel, D.P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–840 (2010).
42. Győrffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123**, 725–731 (2010).
43. Lang, M.R., Lapierre, L.A., Frotscher, M., Goldenring, J.R. & Knapik, E.W. Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. *Nat. Genet.* **38**, 1198–1203 (2006).
44. Saito, A. *et al.* Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. *Nat. Cell Biol.* **11**, 1197–1204 (2009).
45. Townley, A.K. *et al.* Efficient coupling of Sec23-Sec24 to Sec13-Sec31 drives COPII-dependent collagen secretion and is essential for normal craniofacial development. *J. Cell Sci.* **121**, 3025–3034 (2008).
46. Huang, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat. Protoc.* **4**, 44–57 (2009).
47. Wang, Y. *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365**, 671–679 (2005).
48. Minn, A.J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
49. O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V. & Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843 (2005).
50. Yu, Z. *et al.* A cyclin D1/microRNA 17/20 regulatory feedback loop in control of breast cancer cell proliferation. *J. Cell Biol.* **182**, 509–517 (2008).
51. Yu, Z. *et al.* microRNA 17/20 inhibits cellular invasion and tumor metastasis in breast cancer by heterotypic signaling. *Proc. Natl. Acad. Sci. USA* **107**, 8231–8236 (2010).

ONLINE METHODS

Tumor xenografts. All procedures involving mice, such as housing and care, and all experimental protocols were approved by the Institutional Animal Care and Use Committee of Princeton University. Cells were collected from subconfluent cell culture plates, washed with PBS twice and resuspended at the appropriate concentration in PBS. We injected 2×10^5 cells in 0.1 ml PBS on day 0 into the lateral tail vein of 4-week-old female BALB/c mice (US National Cancer Institute) using 26-G needles to generate pulmonary metastasis. For orthotopic primary tumor formation, female BALB/c mice at 6 weeks old were anesthetized, and a small incision was made to reveal the mammary gland. We injected 1×10^6 cells resuspended in 10 µl PBS directly into the mammary fat pad. The primary tumor growth was monitored weekly by measurement of the tumor size. For colonization experiments, after 1 month of colonization, lungs were excised, dissociated and plated in selective medium for colony formation.

Statistical analysis. Results were reported as mean ± s.e.m. (s.e.m.). Two-sided independent Student's *t* test without equal variance assumption or the Wilcoxon signed-rank test was performed to analyze gene and miRNA expression levels, differences in number of tumor colonies and nodules, end points of *in vitro* luciferase assays and histology data. For clinical associations, Spearman rank correlation coefficients were used for studying the association between continuous variables. Tests of hypotheses on the location parameter (median) were performed using Mann-Whitney and Kruskall-Wallis methods. The log-rank test was used to test for differences in survival in univariate analysis. Expression values were introduced into survival analysis as continuous variables using a nonparametric approach in which samples were ranked by expression values and ranks were normalized between 0 and 1. Statistical analyses were performed using R. DRFS and RFS were calculated as described by the STEEP criteria[33]. A standard χ^2 goodness-of-fit test was used to determine significance of overlap in number of enriched gene ontology categories between samples.

Additional methods. Sequences of primers used in the study were listed in **Supplementary Table 5**. Detailed methodology is described in the **Supplementary Methods**.

Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization

Manav Korpal, Brian J. Ell, Francesca M. Buffa, Toni Ibrahim, Mario A. Blanco, Toni Celià-Terrassa, Laura Mercatali, Zia Khan, Hani Goodarzi, Yuling Hua, Yong Wei, Guohong Hu, Benjamin A. Garcia, Jiannis Ragoussis, Dino Amadori, Adrian L. Harris and Yibin Kang

SUPPLEMENTARY RESULTS

Metastatic cell lines express high levels of miR-200s and present epithelial traits

The MCF10A progression series consisting of normal human mammary epithelial (MCF10A) and tumor cell lines with weak (MCF10AT and MCFCA1h) and high (MCFCA1a) metastatic potential[1]. We found that miR-200 expression was highest in the highly metastatic MCFCA1a lines (**Supplementary Fig. 1b**). The strongly metastatic MCFCA1a line also appeared more epithelial relative to the weakly metastatic MCFCA1h line, expressed elevated levels of E-cadherin and reduced levels of *ZEB1/2* (**Fig. 1e** and **Supplementary Fig. 2c**, *upper panel*). In support, gene set enrichment analysis (GSEA) of transcriptomic profiles revealed an enrichment for gene sets associated with adhesion/polarity in the highly metastatic 4T1 and MCFCA1a cells relative to the 4TO7 and MCFCA1h cells, respectively (**Supplementary Table 1**).

Similar observations were made for a series of bladder carcinoma cell lines that were selected *in vivo* for increasing metastasis potential to bone[2]. The strongly metastatic TSU-Pr1–B1 and TSU-Pr1–B2 cells showed elevated miR-200 levels, possessed epithelial traits, and showed ~20–30 fold increase in *CDH1* and reduced expression of *ZEB2* relative to the parental TSU-Pr1 line (**Supplementary Fig. 2c–e**).

miR-200 family induces global changes in gene expression

To rigorously test how extensively C1, C2, or C1+C2 overexpression in 4TO7 cells caused a genome-wide shift to a 4T1-like gene expression state, we performed GSEA analysis. We used gene sets consisting of the 100 most up- or down-regulated genes following C1, C2, or C1+C2 overexpression (as compared to corresponding vector

Nature Medicine: doi:10.1038/nm.2401

controls) in 4TO7 cells. These gene sets were tested for enrichment in the list of all mouse genes ranked by how overexpressed they are in 4T1 cells relative to 4TO7 cells. Strikingly, the C1, C2, and C1+C2 up-regulated gene sets were each highly enriched in the 4T1 phenotype (genes overexpressed in 4T1 relative to 4TO7 cells). The degree of enrichment was greater in C2 than in C1 overexpression (normalized enrichment scores, or NES = 6.87 and 3.25, respectively) and greatest in C1+C2 overexpression (NES = 8.41), with $P < 0.0001$ in all cases (**Fig. 3b**, *left panel*). Conversely, the C1, C2, and C1+C2 down-regulated gene sets were strongly enriched in the 4TO7 phenotype (genes underexpressed in 4T1 relative to 4TO7). Here again the NESs increased in magnitude from C1 to C2 to C1+C2 (–1.56 to –3.65 to –7.39), with $P < 0.002$ for C1 and $P < 0.0001$ for C2 and C1+C2 (**Fig. 3b**, *right panel*). For comparison, we performed similar GSEA analyses using sets of the 100 most up- and down-regulated genes in the CDH1 line. Cdh1 up-regulated genes were moderately enriched in the 4T1 phenotype (NES = 1.85, $P = 0.016$) (**Fig. 3b**, *left panel*) and Cdh1 down-regulated genes were strongly enriched in the 4TO7 phenotype (NES = –3.17, $P < 0.0001$) (**Fig. 3b**, *right panel*). Overall, these results strongly suggest that C1, C2, or C1+C2 overexpression in 4TO7 cells causes a remarkably wide-spread shift in gene expression to a 4T1-like transcriptomic state. Moreover, the degree of enrichment increased from C1 to C2 to C1+C2, mirroring previous findings that phenotypic changes were stronger in C2 than in C1 overexpression and strongest in C1+C2 overexpression (**Fig. 2c**). GSEA of PicTar-derived miR-200 targets revealed global targeting of predicted miR-200 targets in the C1+C2 line (**Supplementary Fig. 6a**), whereas GSEA for four distinct EMT gene signatures revealed a global repression of the EMT-like genetic program[3-6], confirming the acquisition of 4T1-like epithelial properties (**Supplementary Fig. 6b**). Finally, as *Cdh1* overexpression alone causes only a modest genome-wide expression shift, E-cadherin is unlikely to be the only functional mediator of miR-200s, confirming our *in vivo* and clinical data.

To assess the degree of overlap between gene expression changes after C1, C2, and C1+C2 overxpression, we investigated the "enrichment core" subsets of the top-100 gene sets. Enrichment cores consist of the gene set members most directly contributing to the GSEA scores (see **Supplementary Methods**) when tested for

Nature Medicine: doi:10.1038/nm.2401

enrichment in the 4T1 phenotype. The C1, C2, and C1+C2 enrichment cores consisted of 31, 66 and 77 genes, respectively. Of these, 19 were in common between C1 and C2, 17 were in common between C1 and C1+C2, 37 were in common between C2 and C1+C2, and strikingly, 13 were in common between all conditions (**Fig. 3c**, Venn diagram). This final common core gene set included *CDH1* and other genes of various functions, including epithelial-related functionality (**Fig. 3c**).

For an unbiased investigation of which functional classes of genes were globally up-regulated following C1+C2 overexpression, we tested the enrichment of 645 gene ontology (GO) category gene sets. In addition to abundance of GO gene sets related to epithelial characteristics, others included protein transport, channel activity and endoplasmic reticulum functionality (**Supplementary Table 2**). These GO category GSEA analyses suggested that C1+C2 overexpression causes genome-wide expression changes in 4TO7 cells that may influence many cellular functions, particularly those pertinent to protein transport and secretion, in addition to promoting epithelial traits during metastasis.

Genomic and proteomic identification of miR-200 targets

Of the 3,769 proteins quantified by mass spectrometry, 1,562 protein abundance values were cross referenced using their gene symbols with gene expression microarray data. We observed a Spearman's rank correlation of 0.41 between protein and mRNA abundance for these genes (**Fig. 4a**, all dots, $P < 0.001$), and a correlation of 0.42 for genes containing miR-200b/c/429 and/or miR-200a/141 target sites in their 3'-UTR (**Fig. 4a**, red dots; **Fig. 4b**, all dots, $P < 0.001$, $n = 130$). Furthermore, we found a significant enrichment of genes containing miR-200 target sites among the ~46 genes showing a 40% or more decrease in both mRNA and protein abundance ($P < 0.004$, hyper geometric) (**Fig. 4a**, lower left quadrant, red dots; **Supplementary Table 3**). Although most of the genes containing target sites for miR-200s were not significantly changed in expression in C1+C2 lines (**Fig. 4b**, within orange dotted circle), nine showed reduced abundance at both the mRNA and protein level, serving as our initial list of candidate metastasis suppressor genes targeted by the miR-200s (**Fig. 4b**, lower left quadrant, red dots).

Nature Medicine: doi:10.1038/nm.2401

SUPPLEMENTARY DISCUSSION

The biphasic role of miR-200s in metastasis

Results from our studies and others suggest a biphasic role of miR-200s in metastasis that will likely provide the phenotypic and functional plasticity necessary for tumor cells to maximize the potential for escaping the primary tumor and colonizing distant organs (**Fig. 6f**). During early steps of metastasis, tumor cells may undergo EMT in response to external cues at the invasive front, such as high levels of TGF-β[7], stimulating expression of EMT-related transcription factors such as Zeb1/2[8], and hence suppressing the expression of miR-200s. This chain of events promotes the transition of tumor cells to a more motile and invasive mesenchymal phenotype. It is possible that once the disseminated tumor cells reach a secondary organ, changes in the microenvironment may allow tumor cells to regain miR-200 expression and epithelial properties, with enhanced colonization abilities.

Sec23a as a functional target of miR-200s

We provide several lines of evidence supporting Sec23a as a functional downstream miR-200 target in metastasis. Firstly, Sec23a knockdown using three independent shRNAs hinders migration (**Fig. 5a**) and enhances pulmonary colonization (**Fig. 5b,c**), phenocopying miR-200 overexpression (**Fig. 2e,g**). Secondly, both knockdown of Sec23a and overexpression of miR-200s significantly hinder protein secretion via the canonical secretory pathway (**Fig. 6a**), leading to intracellular protein accumulation and modest ER distension (**Supplemental Fig. S9a–c**). Genes encoding proteins significantly reduced in secretion were associated with improved relapse-free survival (**Fig. 6b**, $P = 0.0236$), suggesting their role as metastasis suppressors. Lastly, lung metastases of clinical breast cancer have significantly higher expression of miR-200s (**Fig. 1b**) and lower expression of *SEC23A* (**Fig. 5d,e**) relative to primary tumors, supporting their respective roles in promoting and suppressing metastatic colonization. Interestingly however, *SEC23A* expression alone failed to associate with distant metastasis-free survival (data not shown), implying that gene expression prognosis analysis using primary tumors may not be the only criteria for investigating the clinical importance of candidate metastasis genes, particularly those possessing biphasic roles.

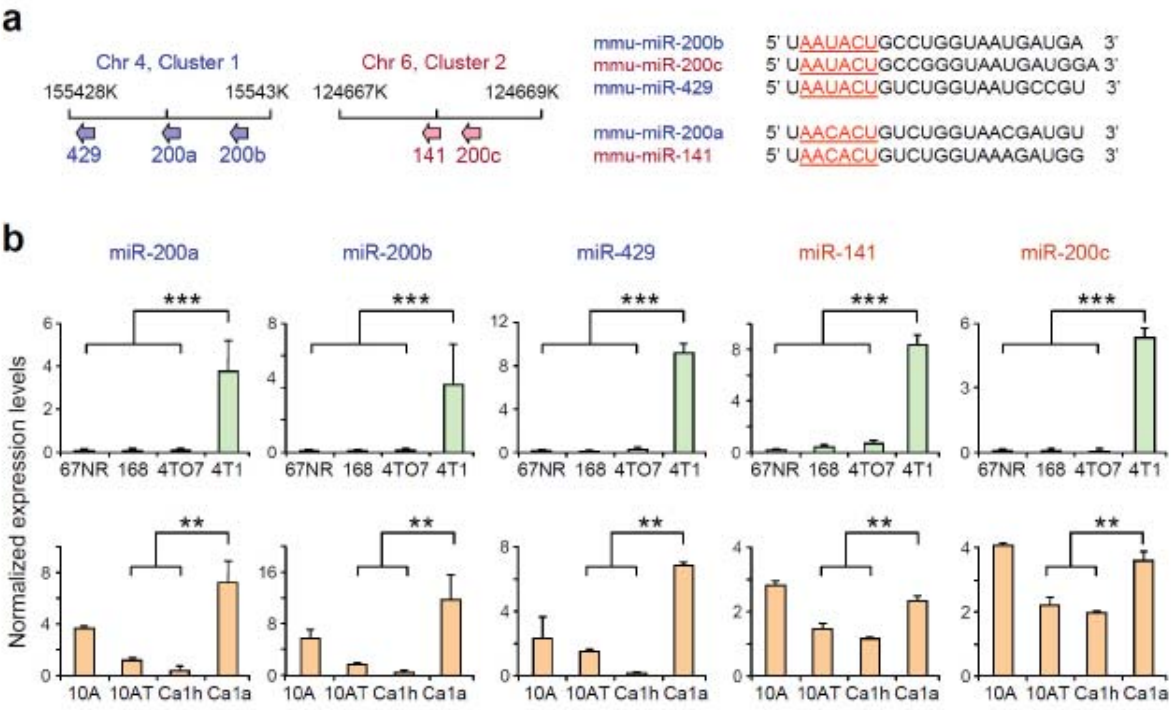
Instead, comparing relative expression levels in primary and metastases (preferably matched) samples may prove to be more informative, as was the case for *SEC23A*.

Sec23a KD reduced secretion of multiple classes of proteins

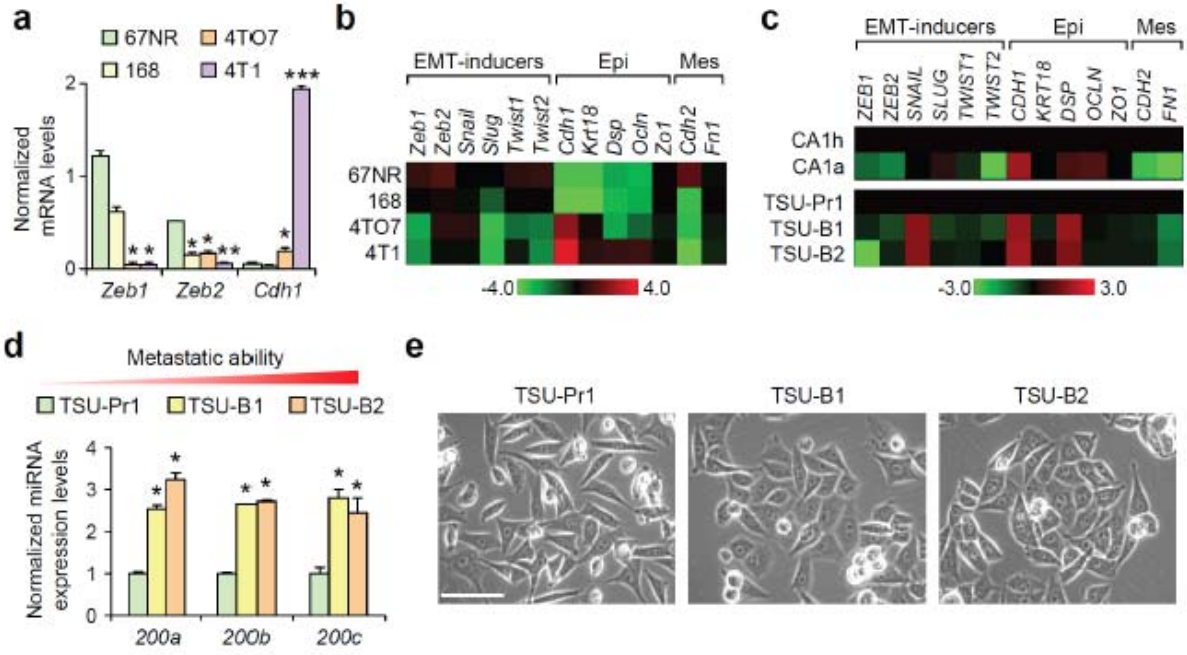
Recent functional genomic analysis of breast cancer metastasis to bone[9], lung[10] and brain[11] revealed that many of the organ-specific metastasis genes encode secreted cytokines, growth factors, proteases and membrane receptors. In our present study, proteins involved in cell adhesion, angiogenesis and immune/inflammatory responses are reduced in secretion in Sec23a KD and miR-200 expression lines. Interestingly, clinical correlation studies indicate potential metastasis suppressive function of these proteins (**Fig. 6b**). In support, we functionally confirmed Tinagl1 and Igfbp4 as two metastasis suppressor proteins that are significantly reduced in secretion upon disruption of the Sec23a-mediated secretory pathway (**Fig. 6c,d**). Although other proteins were not functionally tested, many of them possess various properties that may endow them with the potential to influence tumor progression. For example, the importance of immune/inflammatory cells in regulating metastasis has been well-recognized[12-15]. In light of this, we tested the possibility that Sec23a may modulate immune response as one potential mechanism of influencing colonization. However, we failed to observe any changes in total circulating $CD45^+$ leukocytes following orthotopic inoculation of the control and Sec23a knockdown cells (data not shown). Future studies are needed to fully characterize changes in subpopulations of leukocytes locally (intratumorally) and systemically (peripheral circulating cells and in immune organs) as a potential mechanism of Sec23a-mediated regulation of metastasis.

Nature Medicine: doi:10.1038/nm.2401

SUPPLEMENTARY FIGURES

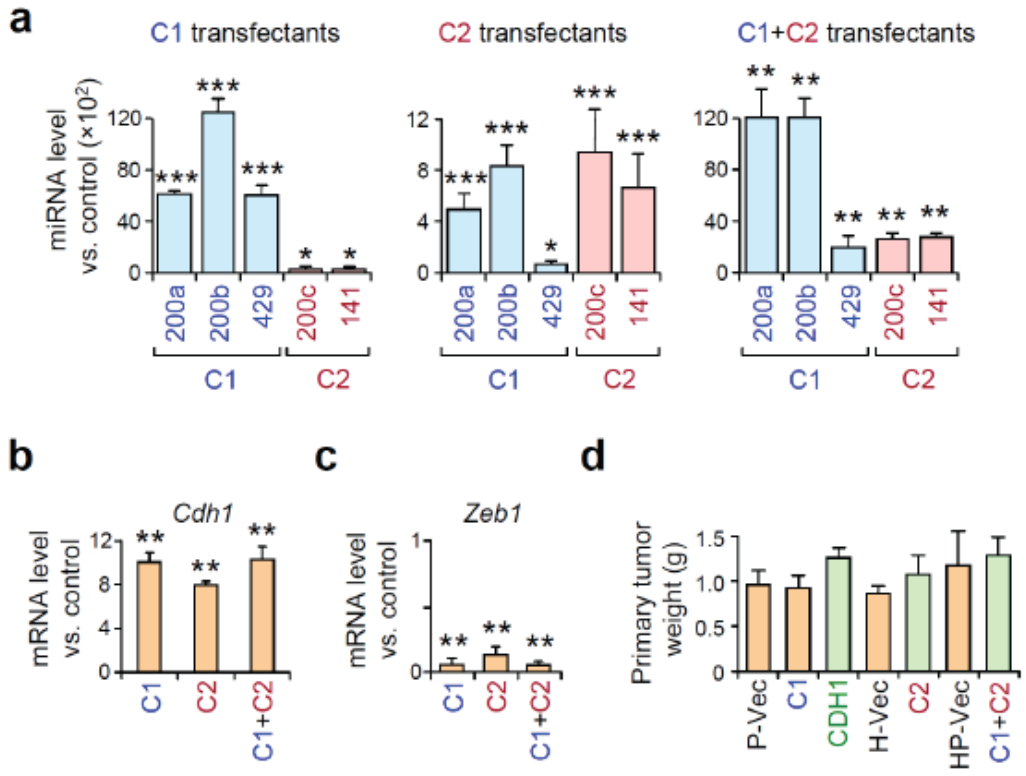


Supplementary Figure 1. miR-200s are abundantly expressed in highly metastatic lines. (**a**) Schematic of genomic organization of cluster 1 (blue font) and cluster 2 (red font) miR-200 members (left panel) and sequences of mature miR-200 members (right panel). Red, underlined font represents the seed sequence and the bolded nucleotide represents the difference in seed sequence between miR-200b/c/429 and miR-200a/141. (**b**) Expression levels of miR-200 members quantified in mouse 4T1 (upper panel) and human MCF10A (lower panel) progression series by TaqMan real-time PCR analysis. Data were normalized to U6 expression and are presented as means ± s.e.m. ** $P < 0.01$, *** $P < 0.001$ (Student's t-test).



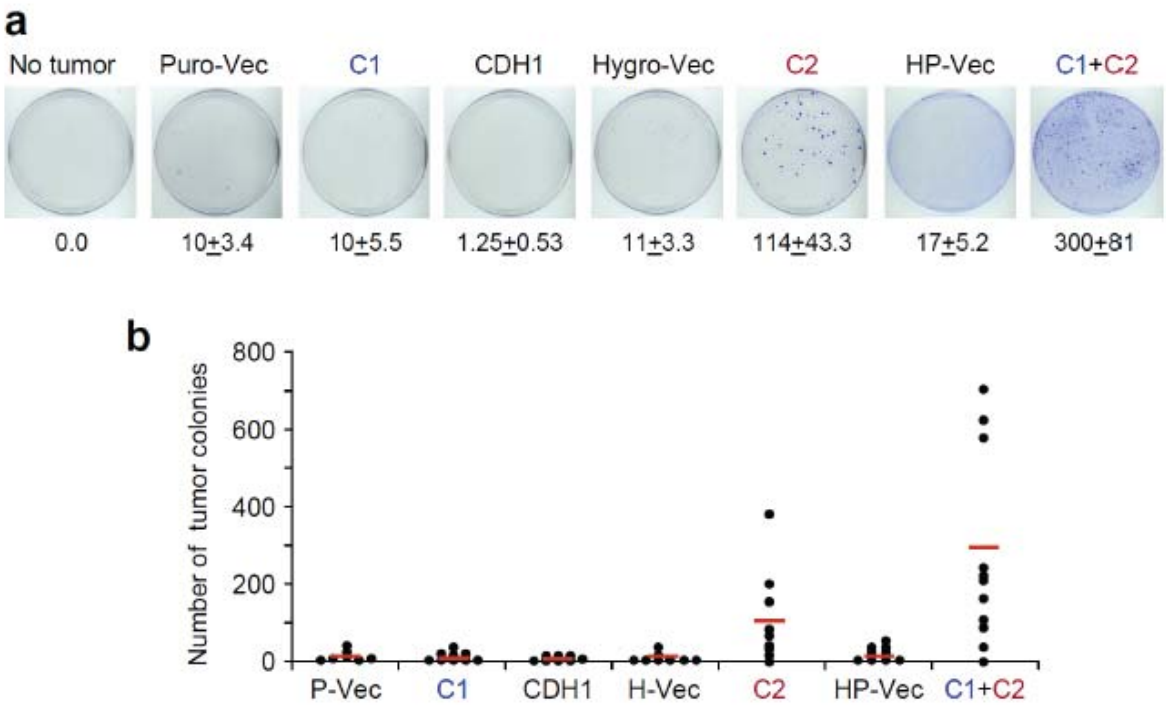
Supplementary Figure 2. Highly metastatic lines express high endogenous miR-200 levels and possess epithelial traits. (**a**) Bar graph showing normalized expression of *Zeb1/2* and *Cdh1* in the 4T1 progression series as assessed by qPCR analysis. Data is normalized to U6 expression and is presented as means ± s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. (**b,c**) Heat maps showing expression levels of various EMT-inducers, epithelial markers and mesenchymal markers in (**b**) 4T1 progression series and in (**c**) CA1h/CA1a cells (upper panel) and TSU progression series (lower panel) as measured by qRT-PCR analysis. Data is normalized to *Gapdh* or *GAPDH* expression and is presented on a $\log_{10}$ scale. Epi: epithelial markers. Mes: mesenchymal markers. (**d**) miR-200 levels in TSU progression series as measured by real-time PCR analysis. Expression was normalized to U6 expression. Experiments were performed in triplicate and data are means ± s.e.m. * $P < 0.05$ (Student's t-test). (**e**) Phase contrast images of TSU-Pr1, -B1 and –B2 cells. Scale bar 100 μm.

Nature Medicine: doi:10.1038/nm.2401



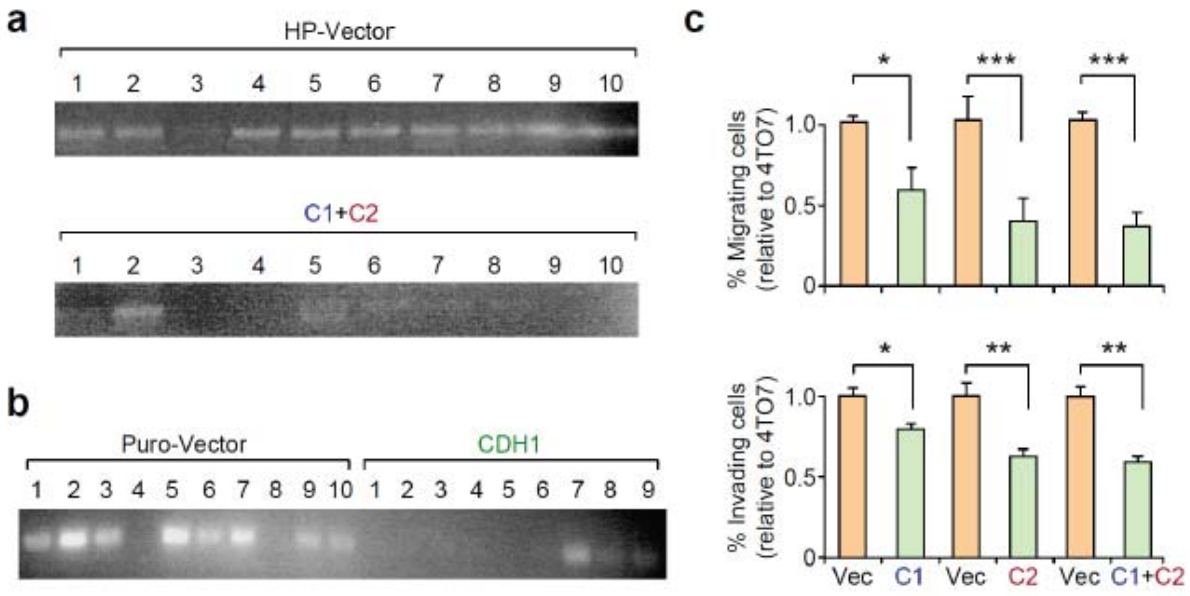
Supplementary Figure 3. Stable overexpression of miR-200 family in 4TO7 cells reduces basal *Zeb1* levels and enhances E-cadherin expression. (**a**) Expression levels of each miR-200 member in 4TO7 cells stably overexpressing cluster 1 (C1 line, blue bars), cluster 2 (C2 line, pink) or both clusters (C1+C2 line) as assessed by qRT-PCR analysis. All data is normalized to U6 expression and data is presented as the mean ratio of expression in each subline to its respective vector control ± s.e.m. (**b,c**) Changes in expression of *Cdh1* and *Zeb1* in C1, C2 and C1+C2 sublines as assessed by qRT-PCR analysis (normalized to *Gapdh*). Data is presented as the mean ratio of expression in each subline relative to its respective vector control ± s.e.m. (**d**) Bar graph showing primary tumor weight at last time point. Data are means ± s.e.m. Data are means ± s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (Student's t-test).

Nature Medicine: doi:10.1038/nm.2401

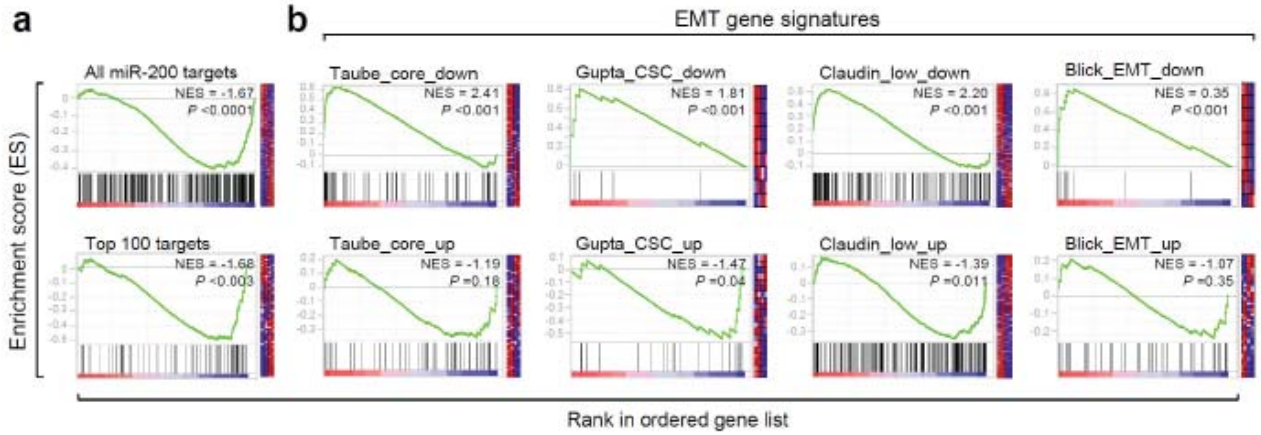


Supplementary Figure 4. Ectopic miR-200 expression enhances spontaneous metastasis through E-cadherin-independent mechanisms. (**a**) Representative plate images of tumor colonies from the lungs of mice injected with miR-200 and CDH1 lines, or with corresponding vector controls, are shown. Data are means ± s.e.m. (n = 9–10). (**b**) Data from (**a**) is presented as dot plot with each dot representing a unique sample. Red lines represent mean values. Experiment was performed three times and data from a representative experiment is presented.

Nature Medicine: doi:10.1038/nm.2401

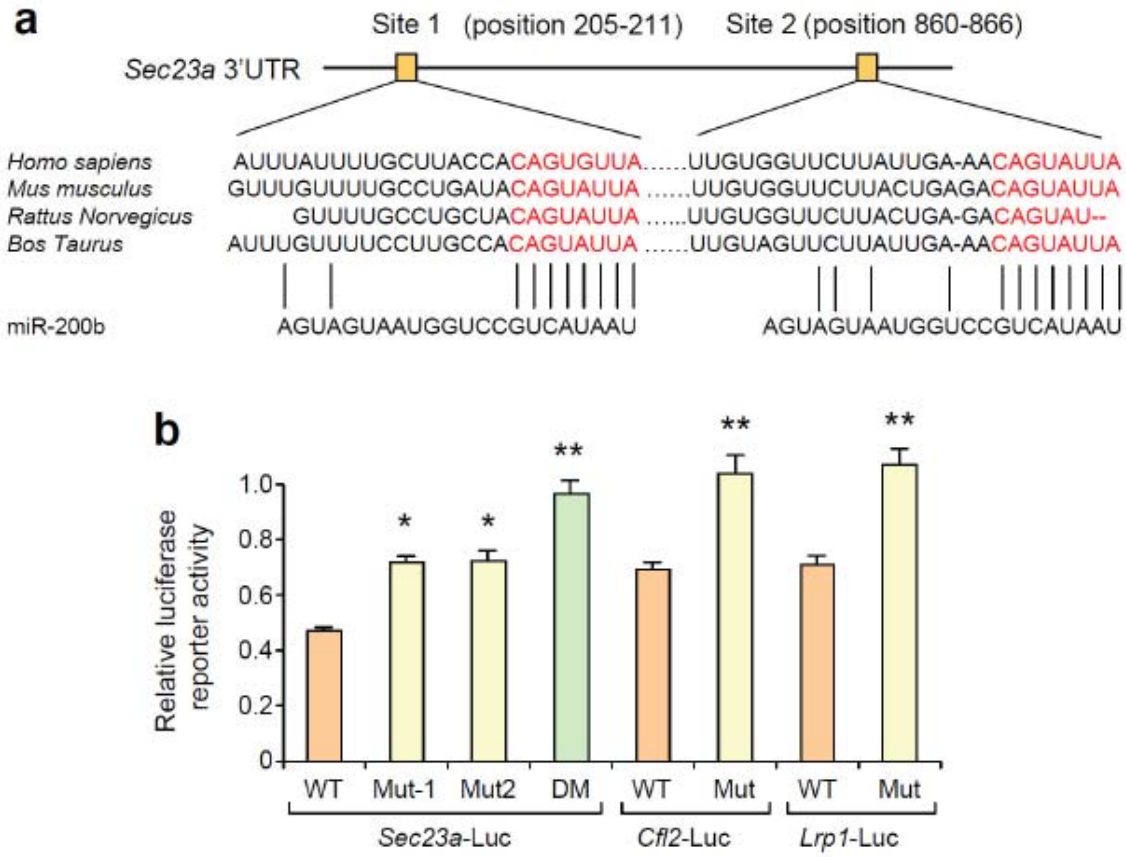


Supplementary Figure 5. Enhanced epithelial character reduces migration/invasion ability of tumor cells *in vivo* and *in vitro*. (**a**,**b**) Genomic DNA was extracted from 50 μl of whole blood and genomic PCR analysis was performed using primers specific for puromycin drug resistance gene. Genomic PCR analysis was performed for C1+C2 (**a**) and *Cdh1* overexpressing cells (**b**) relative to vector controls. (**c**) Bar graphs showing *in vitro* migration and invasion potential of miR-200 lines. Data is presented as the mean ratio of migration of experimental cells relative to vector control cells ± s.e.m from a representative experiment. Transwell migration and invasion assays were performed at least twice. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (Student's t-test).



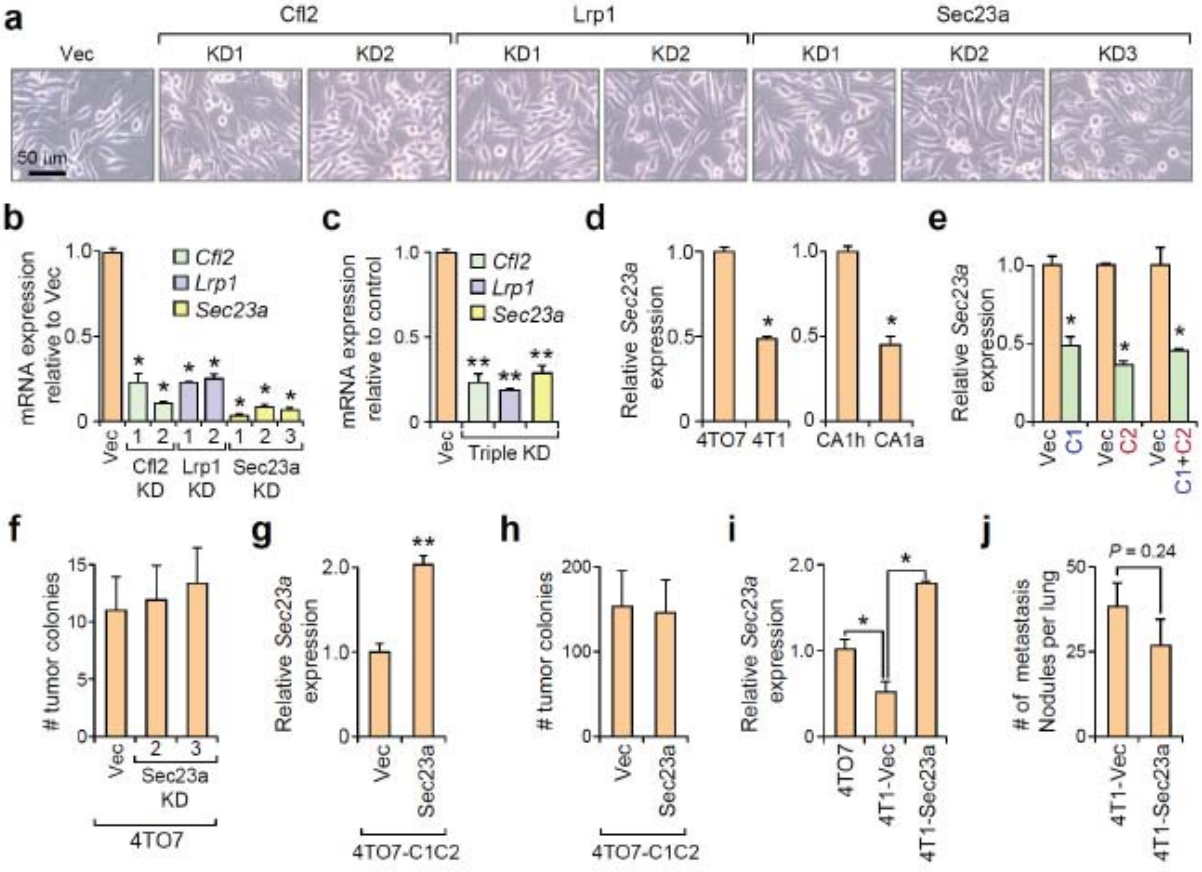
Supplementary Figure 6. Bioinformatic analysis confirms global targeting of predicted miR-200 targets and repression of the EMT program in C1C2 line. (**a**) GSEA using PicTar predicted targets of miR-200s as gene sets. All targets (top) or top 100 targets (bottom) were tested for enrichment in the C1C2 line versus vector control ranked list. (**b**) GSEA of four distinct EMT gene signatures tested for enrichment in C1C2 line versus vector control ranked list. Downregulated (top) and upregulated (bottom) EMT gene sets were used in the analysis.

Nature Medicine: doi:10.1038/nm.2401



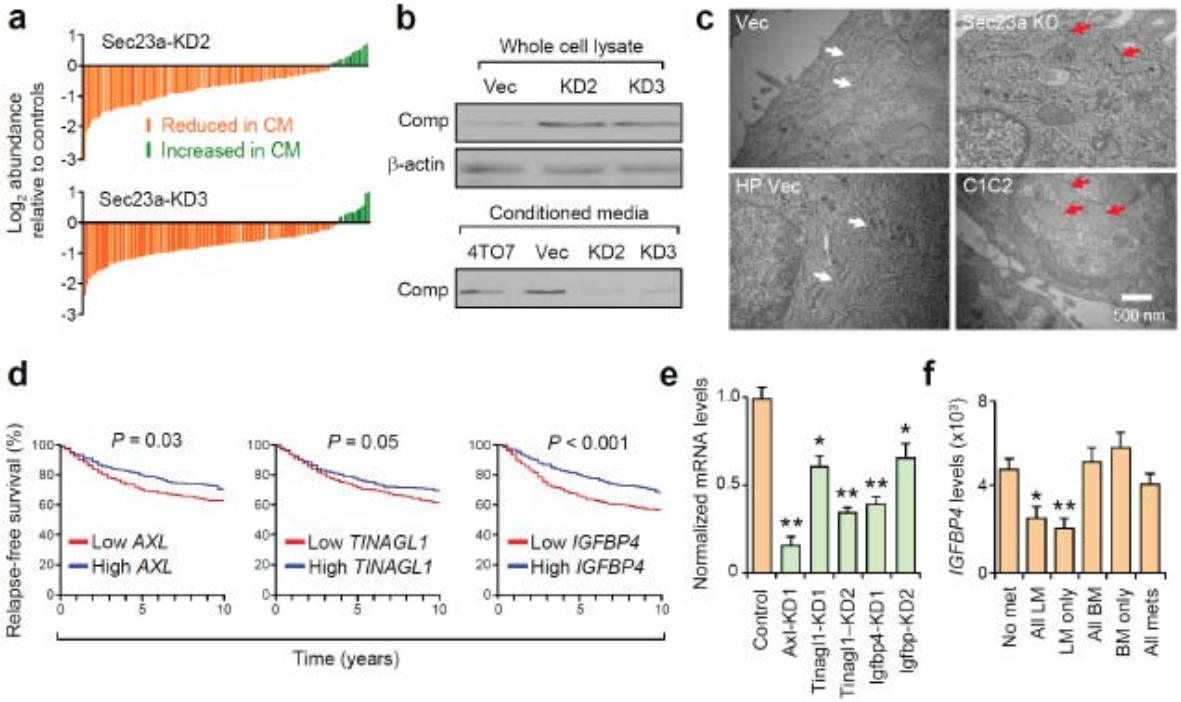
Supplementary Figure 7. *Sec23a*, *Cfl2* and *Lrp1* are direct targets of miR-200s. (**a**) Schematic showing evolutionarily conserved target 'seed' sequences for miR-200b in the Sec23a 3'UTR of various species. Red bolded font represents target 'seed' sequence. Beige boxes represent locations of two miR-200 target sites in Sec23a 3'UTR. (**b**) Luciferase assays confirming direct targeting of *Sec23a*, *Cfl2*, and *Lrp1* in the more physiologically relevant 4TO7 cells. Wild type 3'UTRs (WT, beige bars), 3'UTRs with single mutated sites (MUT 3'UTR, yellow bars), or 3'UTR with both sites mutated (green bars, for *Sec23a* only). Data represents mean change in normalized luciferase activity after co-transfection of miR-200s relative to the negative control pre-miR ± s.e.m. Three independent experiments were performed. * $P < 0.05$, ** $P < 0.01$ (Student's t-test).

Nature Medicine: doi:10.1038/nm.2401



Supplementary Figure 8. Sec23a is necessary but not sufficient to influence metastasis. (**a**) Phase contrast images of Cfl2, Lrp1, Sec23a KD and triple KD cells. (**b**) qRT-PCR showing endogenous expression of direct miR-200 targets relative to vector control (normalized to *Gapdh*) in single knockdown cell lines. (**c**) qRT-PCR confirming stable knockdown of all three direct target genes of miR-200s in the triple KD line. (**d**,**e**) *Sec23a* mRNA levels as measured by qRT-PCR in highly metastatic 4T1 and CA1a cells (**d**) and 4TO7 derived cell lines overexpressing miR-200s (**e**), relative to weakly metastatic 4TO7 and CA1h cells (**d**) and 4TO7-derived vector control lines (**e**), respectively. (**f**) Bar graph showing number of tumor colonies attained from dissociated lungs from mice orthotopically injected with vector control or two independent Sec23a KD lines. (**g**) Bar graph showing ectopic expression of Sec23a in C1C2 line relative to C1C2-vector control line. (**h**) Bar graph showing number of tumor colonies attained following dissociation of lungs from mice orthotopically injected with C1C2 vector control

and Sec23a overexpression lines. (**i**) Bar graph showing ectopic expression of *Sec23a* in 4T1 cells comparable to 4TO7 cells. (**j**) Bar graph showing the number of pulmonary lesions induced when vector control and Sec23a overexpressing 4T1 cells were inoculated into the lateral tail vein. All data is presented as means ± s.e.m. * $P < 0.05$, ** $P < 0.01$ (Student's t-test).



Supplementary Figure 9. Tinagl1 and Igfbp4 represent two secreted components that possess metastasis suppressor activities *in vivo*. (**a**) Bar graphs showing relative protein abundance in the conditioned media (CM) of two independent Sec23a KD lines relative to vector controls. Orange bars represent abundance of proteins reduced in CM whereas green bars represent proteins increased in abundance. (**b**) Western blots showing intracellular accumulation (upper) and reduced secretion (lower) of Comp in two independent Sec23a KD lines relative to parental (4TO7) and vector control (Vec) lines. (**c**) Electron microscopy analysis of the ER structure in Sec23a KD (upper right panel, red arrows) and C1+C2 (lower right panel, red arrows) lines relative to their vector controls (upper left and lower left panels respectively, white arrows). (**d**) Kaplan-

Meier plots of relapse-free survival with high and low median expression of level of *AXL*, *TINAGL1* and *IGFBP4* in the Gyorffy et al database[16]. *P* values computed by log-rank test. (**e**) Expression levels of *Axl*, *Tinagl1* and *Igfbp4* assessed by qPCR in stable knockdown lines (green bars) relative to vector control line (beige). Data is presented as means ± s.e.m. (**f**) Expression of *IGFBP4* in MSK82 dataset grouped according to incidence and site of metastasis. NM – no metastasis; All LM – lung metastasis (including those with additional metastases at other sites); LM only – lung metastasis only (excluding those with additional metastases at other sites); All BM – bone metastasis; BM only – bone metastasis only; All mets – all cases with metastasis (any site). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (Student's t-test).

Nature Medicine: doi:10.1038/nm.2401

Supplementary Tables

Supplementary Table 1. Enrichment of gene sets associated with adhesion and polarity in highly metastatic lines. Top functional gene sets enriched in highly metastatic 4T1 and MCF10A-CA1a relative to lowly metastatic 4TO7 and CA1h cells respectively are listed. Rows highlighted in yellow represent gene sets that are common between the 4T1 and MCF10A series.

Functional enrichment in 4T1 relative to 4TO7	Nominal *P* value
Epidermis development	0.0001
Ectoderm development	0.002
Regulation of cell migration	0.007
Tight junction	0.011
Apicolateral plasma membrane	0.025
Apical junction complex	0.03
Basolateral plasma membrane	0.035
Epidermis development	0.0001

Functional enrichment in CA1a relative to CA1h	Nominal *P* value
Calcium-independent adhesion	0.0062
Apicolateral plasma membrane	0.023
Apical junction complex	0.026
Tight junction	0.031

Nature Medicine: doi:10.1038/nm.2401

Supplementary Table 2. Functional (GO category) gene sets enriched in the ranked gene list according to their differential expression in C1+C2 versus control. Size of each gene set and *P*-value is listed. Beige shade represents functions associated with epithelial phenotype, aqua shade represents protein transport functions and yellow shade represents endoplasmic related functions.

GO category	GO subcategory	Size	*P*-value
Epithelial	Epidermis development	34	0
	Ectoderm development	38	0.002
	Tight junction	17	0.004
	Apicolateral plasma membrane	19	0.006
	Apical junction complex	19	0.012
	Cell junction	41	0.024
	Intercellular junction	31	0.024
Plasma membrane	Intrinsic to plasma membrane	418	0
	Plasma membrane part	496	0.002
	Integral to plasma membrane	411	0.002
Protein transport	Protein targeting	52	0.002
	Intracellular protein transport	71	0.004
	Protein processing	25	0.012
	Establishment of protein localization	83	0.024
	Protein transport	77	0.027
Channel activity	Cation channel activity	48	0
	Substrate-specific channel activity	59	0.006
	Potassium channel activity	18	0.008
	Ion channel activity	57	0.01
	Gated channel activity	49	0.02
	Voltage-gated cation channel activity	28	0.027
	Voltage-gated channel activity	30	0.039
Receptor	G-protein coupled receptor protein signaling	109	0.012
	Receptor complex	29	0.014
	Rhodospin-like receptor activity	35	0.016
	G-protein coupled receptor activity	58	0.018
	Peptide receptor activity	17	0.019
Endoplasmic reticulum	Endoplasmic reticulum	142	0.026
	Endoplasmic reticulum membrane	38	0.035
	Nuclear envelope endoplasmic reticulum	40	0.038

Nature Medicine: doi:10.1038/nm.2401

Supplementary Table 3. 46 genes reduced at protein and RNA levels by miR-200s in C1+C2 line relative to the control. miR-200 targets are highlighted in red.

Gene	Log_2 protein ratio	Log_2 RNA ratio	Target site(s)
Tcp11l2	-4.32	-1.07	-----
Fth1	-4.07	-0.94	-----
Kalrn	-3.92	-3.22	-----
Ak1	-3.44	-1.66	-----
Trim2	-3.25	-1.31	1 miR-200a/141
BC005685	-2.74	-1.02	-----
2010011I20RIK	-2.48	-1.06	-----
Slc5a3	-2.44	-1.50	1 miR-200a/141 site 1 miR-200b/c/429 site
S100a4	-2.37	-1.28	-----
Sec23a	-2.27	-1.46	2 miR-200b/c/429 sites
Cfl2	-2.24	-0.91	1 miR-200b/c/429 site
Slc29a1	-2.15	-0.95	-----
Crlf1	-2.09	-2.18	-----
Cotl1	-2.04	-2.12	-----
Crtap	-2.01	-1.19	1 miR-200b/c/429 site
Nfat5	-1.88	-0.88	-----
Osbpl8	-1.86	-0.74	-----
Itgb2	-1.84	-0.78	-----
Jun	-1.81	-0.85	1 miR-200b/c/429 site
Neo1	-1.80	-0.96	1 miR-200b/c/429 site
Cpne2	-1.79	-1.22	-----
Sned1	-1.72	-1.31	-----
Arhgap18	-1.66	-1.67	-----
Mrps6	-1.52	-0.81	-----
Lrp1	-1.52	-1.03	1 miR-200b/c/429 site
Plec1	-1.51	-0.97	-----
Gpnmb	-1.49	-1.29	-----
Sord	-1.46	-1.45	-----
Blvrb	-1.34	-1.00	-----
Acsl1	-1.30	-0.82	-----
Rlf	-1.29	-0.86	-----
Dbn1	-1.27	-0.77	-----
Dap3	-1.23	-1.19	-----
Naglu	-1.22	-0.79	-----
Tanc1	-1.21	-1.24	-----
Selenbp2	-1.19	-2.53	-----
Myo10	-1.14	-0.79	-----
SerpinH1	-1.10	-0.77	1 miR-200a/141 site
Cand2	-1.09	-0.78	-----
Pola2	-1.02	-1.06	-----
Itgb7	-1.01	-1.35	-----
Xdh	-0.99	-1.42	-----
Dgcr8	-0.98	-0.76	-----
AW146020	-0.85	-0.88	-----
Akr1b8	-0.82	-0.84	-----
Camk2B	-0.75	-1.32	-----

Nature Medicine: doi:10.1038/nm.2401

Supplementary Table 4. List of 38 proteins reduced in abundance in conditioned media in two independent Sec23a knockdown lines (H) relative to vector controls (L).

Gene	Log_2 protein (Ratio H:L)-Sec23a KD#2	Log_2 protein (Ratio H:L)-Sec23a KD#3	DAVID functional category
C4b	-3.04	-2.85	W,IM
Sema7a	-1.87	-1.72	
Sned1	-1.75	-1.56	CA
Col18a1	-1.68	-1.31	CA,ESO
Lbp	-1.64	-1.27	W,IM
Tinagl1	-1.53	-1.19	IM
Lgals3bp	-1.49	-1.67	CA
Sdc4	-1.46	-1.26	
Serpine2	-1.45	-1.32	
Csf3	-1.43	-2.03	IM
Mmp3	-1.38	-1.43	
Ltbp3	-1.37	-1.18	
Crlf1	-1.34	-1.83	
Sema3b	-1.34	-0.90	
F5	-1.33	-1.32	W,CA
Lama4	-1.29	-1.00	CA
Col4a2	-1.28	-1.18	
Hspg2	-1.28	-1.06	CA,ESO
Vasn	-1.28	-0.89	
Lcn2	-1.27	-1.00	
Cxcl1	-1.26	-0.92	W,IM
Agrn	-1.26	-1.26	ESO
Igfbp4	-1.25	-1.11	W,IM
Col4a1	-1.13	-1.12	
Ace	-1.13	-1.49	
Esm1	-1.10	-1.50	
Cp	-1.08	-1.16	
Oaf	-1.01	-1.16	
Cd14	-1.00	-0.87	W,IM
Axl	-1.00	-0.89	
Spp1	-0.96	-1.09	CA
Gm2a	-0.89	-1.19	
Efemp2	-0.88	-1.09	W
Lamc1	-0.86	-0.81	CA,ESO
Col6a1	-0.85	-0.87	CA
Masp1	-0.83	-0.89	W,IM
Npc2	-0.82	-1.04	
Bmp1	-0.80	-0.96	

Note: W: response to wounding; CA: cell adhesion; IM: immune response/inflammatory, ESO: extracellular structure organization

Nature Medicine: doi:10.1038/nm.2401

Supplementary Table 5. Sequences of primers used in the study. Red and non-capitalized fonts indicate mutated nucleotides.

Overexpression	Forward 5'-3'	Reverse 5'-3'
miR200 Cluster 1	GTTGACCTCTCCACTACCTA	ACCAGTGTTGATAGCACAGG
miR200 Cluster2	TTCTTCTGCACCACTCTAGG	CCTTGGGTTCTTAGAGGAAG
Sec23a	TCACGGAGCTTCTAGACAGA	TGAGTGTCTG ACACATCTGG
CDH1	AGTGTTTGCTCGGCGTCTGC	TAGTCCCCTAGTCGTCCTCACCAC

Sec23a ShRNAs	Sense 5'-3'	Antisense 5'-3'
KD #1	GCCTACAGCTTTGGTTGGACTT	AAGTCCAACCAAAGCTGTAGGT
KD #2	GGTTCTTACTGAGACAGTA	TACTGTCTCAGTAAGAACC
KD #3	GGAGATGGTTCTGTTTGAT	ATCAAACAGAACCATCTCC

qRT-PCR	Forward 5'-3'	Reverse 5'-3'
Mouse primers		
Cfl2	ATTCTGGGCTCCTGAAAGTG	GAAACTACAACACTGCCCCC
Jun	CGCCTGATCATCCAGTCC	GACACTGGGAAGCGTGTTCT
Lrp1	CGTGCAAAGATTTTGACGAGT	GCCGATCTACTGGCTCATTC
Neo1	ACAGAAAGCTCTGGCAGGTT	TCCATGGATTCGTGAGCATA
Sec23a	ATGTAGTTCTGCGTGGTCCC	ATGAACCTGCACCATTCTCC
Trim2	CCATGATGGAAATGTGATGG	CTTTTGTTGACTGCGTCCAG
SerpinH1	TGACCTGCAGAAACATCTGG	TGTCTTGGTCAAAGGGGTTG
Slc5A3	TCTGTCATCCTGCTCATTGG	AGGTGAGGCCAACATGTACC
Crtap	TCGTACGAGAGCCTGTTTGTC	TTGAAGTCCTTGATCTCCCG
Cdh1	TGGGCAGAGTGAGATTTGAA	CGGACGAGGAAACTGGTCT
Zeb1	GGAAGAGAGCAAAGACATGTGA	AAACCGTTTCTTGCAGTTCG
Zeb2	GTACCTTCAGCGAAGCGACA	TCAGCAGTTGGGCAAAAGCA
Zo1	CCCACAAGGAGCCATTCCTGAAGG	AGGGTCACAGTGTGGCAAGCG
Fn1	TGTGTGGGGAACGGTCGTGGA	TGGCACTGGTCAATGGGGTCACA
Dsp	TCTACTCCCACTCCCGGCGC	TCAGTTCCGGCTGCGCGATC
Krt18	CCAGGGCGTCAGAGACTGGG	CTCCACAGACTGGCGCATGGC
Ocln	CGGCCGCCAAGGTTCGCTTAT	GCATCGGCCGGACATGCATCT
Human primers		
CDH1	GCCCCATCAGGCCTCCGTTT	ACCTTGCCTTCTTTGTCTTTGTTGGA
ZEB1	AGTGGTCATGATGAAAATGGAACACCA	AGGTGTAACTGCACAGGGAGCA
ZEB2	GACAGATCAGCACCAAATGC	GCTGATGTGCGAACTGTAGG
ZO1	ACAAGCGCAGCCACAACCAAT	CTGCTTTCTGTTGAGAGGCTGGCT
FN1	ATCATCCCAGAGGTGCCCCAAC	CCACCTCAGGCCGATGCTTGAA
DSP	TGTCCCTCGAGTCCGCAGGG	TCTCGCGCAGGTCGGCTTTG
KRT18	ACCTGAGGGCTCAGATCTTCGCA	CACAGACTGGCGCATGGCCA
OCLN	CCTCTCGGGCCGCAACATCG	AACCAATCTGCTGCGTCCTAGACC
SEC23A	GGACTGCTGGAGTGTACTTTTCCCA	GCAGCTCGATTAGCCAATGCTTCA

3'UTR Reporter	Forward 5'-3'	Reverse 5'-3'
Cfl2	CCGCTCGACGCTGGGAGAGA	AGCAGCCTTACCAAAAAGCCACA
Jun	AGCTGGCATCCACGGCCAAC	TCCAGCGGGCTGACCCTCTC
Lrp1	CCATTCCCTGGCCAGCACGG	TGCCCGCTTGCCCTTCCTTG
Neo1	ACGACCTGCCTTTTGCTTTGTGGT	AGCTGCTAGGCCCCAGCTCA
Sec23a	ACACGGAGCATGGTGGCAGC	CGGGTGCAGTCAGTCCCCAC
Trim2	AGGCCCTGAGTTGAGCCCATCA	ACAGCCCATGGCAGCCTCTT
SerpinH1	GCGCAGCCCCAAGCTGTTCT	ACAAAGGCTCATATTTCCCTTCCCCC
Slc5a3	CCCACTCAGCATTGCCTTCCTCC	CGTCTGCAGGCTAATGTAGCAGCTC

Nature Medicine: doi:10.1038/nm.2401

Mutagenesis	Forward 5'-3'	Reverse 5'-3'
Cfl2 MUT	GGGAGACTTGTTTgAcTtTaATTTATAGTTGCACTT GATTACCG	CGGTAATCAAGTGCAACTATAAATTAAAGTCA AACAAGTCTCCC
Jun MUT	CCTAACATTCGATCTCATTgAcTtTaAAAGGGGGG TGGGAGGGG	CCCCTCCCACCCCCCTTTTAAAGTCAATGAG ATCGAATGTTAGG
Lrp1 MUT	GCAAGCGAGCAAGCAgAcTtTaATCTCTTTGCATT TCCTTCC	GGAAGGAAATGCAAAGAGATTAAAGTCTGCT TGCTCGCTTGC
Neo1 MUT	GTTTGTTGCTTTCTGTGCAGATTCtGaAaTGGGGT GGGATTGGGG	CCCCAATCCCACCCCATTTCAGAATCTGCAC AGAAAGCAACAAAC
Sec23a MUT1	CAAGTTTGTTTTGCCTGATAgAcTtTaATAATTG GTTTTGTAACTTTG	CAAAGTTACAAAACCAATTATTAAAGTCTATC AGGCAAAACAAACTTG
Sec23a MUT2	GAATTTGTGGTTCTTACTGAGAgAcTtTaAATA TAATGTTTAATTATGTC	GACATAATTAAACATTATATTTAAAGTCTCTC AGTAAGAACCACAAATTC
Trim2 MUT	GATTTCCCCTTAGAGAGCAAcTcTaACCAAAG TTCTGTTGAGC	GCTCAACAGAACTTTGGTTAGAGTTGCTCTC TAAGGGGAAATC
SerpinH1 MUT	CTGCCTCAACAGTCAATCtGaGaTCATATTTATGG CCAGGC	GCCTGGCCATAAATATGATCTCAGATTGACT GTTGAGGCAG
Slc5a3 MUT1	ATTTTAATATGTAAcTtTaAATGCATTTAAAAAGAT GTCTGCG	CGCAGACATCTTTTTAAATGCATTTAAAGTTA CATATTAAAAT
Slc5a3 MUT2	CTGTAGACTGTGTTAAcTcTaACTGTTGTTTAAAAA TGGG	CCCATTTTTAAACAACAGTTAGAGTTAACACA GTCTACAG

Nature Medicine: doi:10.1038/nm.2401

SUPPLEMENTARY METHODS

Cell culture. 67NR, 168FARN, 4TO7, 4T1, MDA-MB-231 parental and derivative cells, TSU parental and derivatives cells, as well as H29 cells, a packaging cell line used for retrovirus production, were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) containing 10% fetal bovine serum, supplemented with 1 mM L-glutamine and penicillin/streptomycin (Gibco). MCF10A, MCF10AT, MCFCA1h and MCFCA1a cells were maintained in 1:1 DMEM:Hams F12 medium (Invitrogen) containing 5% horse serum, supplemented with 20 ng ml^{-1} EGF, 0.5 μg ml^{-1} hydrocortizone, 10 μgml^{-1} insulin, 1% Pen/Strep and 50 ng ml^{-1} Cholera toxin.

RNA extraction and quantitative real-time PCR. Total RNA was extracted using the miRVana miRNA isolation kit (Ambion) according to the manufacturer's instructions. This RNA preparation is enriched for both small and large RNA species. For miRNA analysis, mature miRNAs were reverse transcribed (Applied Biosystems), and real-time PCR was performed using TaqMan miRNA assays (Applied Biosystems). All miRNA data was normalized to U6 expression. For mRNA analysis, cDNA was oligo(dT) primed from 1–2 μg of total RNA using a SuperScript® III First-Strand Synthesis System (Invitrogen). cDNA from each sample was diluted 15–20 fold and real-time was performed in triplicates using Power SYBR® green PCR master mix (Applied Biosystems) on an ABI 7900HT series PCR machine. Expression levels were normalized to *GAPDH* expression. All analysis was performed using the SDS2.3 software.

miRNA microarrays for cell line samples. MiRNA microarrays were synthesized by spotting cDNA probes to 377 miRNAs (mirVana miRNA probe set v1; Ambion) onto epoxide-coated slides by the Microarray Core Facility at Lewis-Sigler Institute of Integrative Genomics at Princeton University. Total RNA was extracted using the miRVana miRNA isolation kit (Ambion). Small RNA species were enriched from 50 μg total RNA using the FlashPAGE fractionator system (Ambion) and the entire fraction of small RNAs were labeled for hybridization using the mirVana miRNA labeling kit (Ambion). Competitive hybridization experiments were performed in duplicate. The arrays were analyzed using the G2565BA scanner (Agilent Technologies), and median

fluorescent intensities were obtained after subtracting background. To identify differential miRNA expression between samples, the median fluorescent intensities were normalized using the median within the array.

Agilent cDNA microarrays for cell line samples. Total RNA was isolated using the miRVana miRNA isolation kit. 125 ng of total RNA was profiled by Agilent cDNA microarrays and subsequently analyzed using procedures previously described[17].

Human tumor samples

Tumor specimens used in the study (Oxford collection and Meldola collection) were obtained with informed consent from all subjects in accordance with the Local Ethics Committees of the University of Oxford and the of Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (I.R.S.T.).

miRNA microarrays for primary breast tumors of the Oxford Cohort. This study includes samples from a historical series of breast cancer patients (n = 210) treated in Oxford between 1989 and 1992 with > 10 years follow-up. Patients received surgery followed by adjuvant hormone therapy or no adjuvant treatment. The demographics, treatment and other clinical details have been previously described[18,19]. miRNA expression was measured using the Illumina miRNA arrays version 1.0. The protocol and reagents supplied by the manufacturer were used to prepare and hybridize the samples. Briefly, 200 ng of total RNA was poly-adenylated and converted to biotinylated cDNA, which was attached to a solid phased and hybridized with a pool of miRNA-specific oligonucleotides (MSO). Each single MSO was used to detect one miRNA on the panel. Universal PCR amplification was then performed, creating fluorescently labeled products identifiable by their unique MSO sequence. These products were hybridized on the Illumina miRNA array. Hybridization signals were detected and quantified using an Illumina scanner and BeadStudio Software. Average signal values were background subtracted by using a local background subtraction method (BeadStudio). Expression was normalized using quantile normalization; QC of the arrays and more details on the methods are provided in a previous study[18].

Nature Medicine: doi:10.1038/nm.2401

Illumina mRNA arrays for primary breast tumors of the Oxford Cohort. Illumina Human RefSeq-8 arrays (Illumina Inc.) were used. RNA was amplified using Ambion Illumina Amplification Kit. 850 ng of amplified RNA product was hybridized to the Illumina Sentrix Beadchip 8x1 GAP REFSEQ2 using single chamber hybridization cartridges. Washing and staining were carried out as described in the Illumina Whole Genome Expression Manual version 1. Beadchips were scanned using the Illumina BeadArray Reader. Expression data was extracted using the BeadStudio software, using background subtraction. Rescaling was used to eliminate negative values and normalization was done in Bioconductor (R) using quantile normalization.

Analysis of primary tumors and lung metastases of the Meldola cohort. Women with resected breast cancer were selected from patients followed from 1995 to 2010 in I.R.S.T., Meldola, Italy. Surgical tumor specimens were fixed in formalin and embedded in paraffin. Tissues collected were ten primary breast tumors and ten lung-pleural metastatic tissues derived from primary breast lesions. For six patients, matched primary-metastatic tissues were available. Total RNA was collected from 20 μm thick sections from formalin-fixed paraffin embedded (FFPE) tissue blocks using the FFPE RNA/DNA Purification kit (Norgen) according to the manufacturer's instructions and subjected to RT-PCR analysis.

Generation of stable miR-200 and CDH1 overexpressing and Cfl2, Lrp1, Sec23a, Axl, Tinagl1 and Igfbp4 knockdown lines. To overexpress all five miR-200 family miRNAs simultaneously, genomic fragments encoding cluster 1 (miR-200b/200a/429) and cluster 2 (miR-200c/141) were cloned into pMSCV-puro and pMSCV-hygro retroviral vectors respectively. To overexpress E-cadherin, the coding sequence was cloned into pMSCV-puro retroviral vector. Short hairpin RNAs (shRNA), cloned into lentiviral vector pLKO.1-puro, targeting *Cfl2* (shRNA#1, Cat# TRCN0000071542; shRNA#2, Cat# TRCN0000071541), *Lrp1* (shRNA#1, Cat# TRCN0000119622; shRNA#2, Cat# TRCN0000119626), *Sec23a* (shRNA#1, Cat# TRCN0000100894), *Axl* (Cat# TRCN0000023312), *Tinagl1* (shRNA#1, Cat# TRCN0000030659; shRNA#2, TRCN0000030660), and *Igfbp4* (shRNA#1, Cat# TRCN0000114796; shRNA#2, TRCN0000114798) were purchased from Sigma Aldrich. Sec23a directed shRNA sequences #2 and #3 were generated for *Sec23a* using an in-house algorithm, and

were cloned into the pSUPER-Retro retroviral vector system[20]. Retrovirus was produced by transfecting the vectors into the H29 packaging cell line[10] and harvested 48h after transfection, filtered, and used to infect cell cultures in the presence of 5 μg ml^{-1} polybrene.

Lentivirus was produced by transfecting 293FT cells with VSVG:deltaR8.9:shRNA constructs at a ratio of 1:2.5:1.25. Virus was harvested 2–3 days after transfection, filtered, and used to infect cell cultures in the presence of 5 μg ml^{-1} polybrene. In all cases, infected cells were selected with appropriate drugs in the media and at least 1000 independent clones were pooled to generate stable cell lines to avoid clonal variations. Stable cell lines infected with control vectors were generated to be used as negative controls for *in vitro* and *in vivo* experiments. All primer sequence information is in **Supplementary Table S5**.

3'UTR luciferase reporter assays. The 3'UTRs of *Cfl2*, *Jun*, *Lrp1*, *Neo1*, *Sec23a, Trim2, SerpinH1 and Slc5a3* were PCR amplified from mouse genomic DNA. Amplified 3'UTRs were cloned into appropriate sites downstream of the firefly luciferase coding region in the pMIR-REPORT™ microRNA expression reporter plasmid (Cat # AM5795, Ambion). Mutations in miR-200 target sites were generated using the QuikChange Multi site-directed mutagenesis kit (Stratagene). Primers used to amplify WT and mutant (MUT) 3'UTR are listed in Table S5. Reporter assays were performed as follows: 5×10^4 4TO7 cells were seeded in 24-well plates 24 h prior to transfection. The following day, 200 ng of reporter plasmid along with 200 ng of control plasmid, constitutively expressing *renilla*-luciferase, was co-transfected using Lipofectamine 2000 (Cat # 11668027, Invitrogen). Cells were collected 24 h post-transfection and assayed for luciferase activity using the Glomax 96 microplate luminometer (Promega). To assess the effect of candidate miRNAs on reporter activity, 10 pM of synthetic precursor microRNAs (pre-miRs) (Ambion) was co-transfected. All experiments were performed in triplicates and repeated at least twice.

Transwell migration and invasion assays. Control and genetically modified 4TO7 cells were trypsinized and 1×10^5 cells were resuspended in serum-free media and placed in inserts (Costar) containing 8-μm pores with (invasion assay) or without

Nature Medicine: doi:10.1038/nm.2401

matrigel (1mg/ml) (migration assay). These inserts were placed in wells with serum-containing media. 12 h post-seeding, serum-containing media was aspirated, and 500 µl of trypsin was placed into the wells to trypsinize the cells that had passed through the pores. Trypsin was neutralized with serum-containing media and centrifuged for 2 min at 1000 rpm. 900 µl of media was aspirated and the cell pellet was resuspended in the remaining 100 µl. 10 µl of this mixture was used to count the number of cells that had migrated using a hemacytometer.

Immunoblot analysis. Cells were lysed in 1% SDS Laemmli sample buffer containing 3% β-mercaptoethanol, sonicated 10 times at 1 sec intervals and subsequently boiled for 5 min. Approximately 20 µg of protein was loaded per lane and resolved by SDS-polyacrylamide electrophoresis. Protein was transferred onto nitrocellulose membranes, blocked and probing was carried out with antibodies to E-cadherin (1:1000 dilution, Cat # 610181, BD Transduction Laboratories), Comp (1:1000 dilution, Cat # ab74524, Abcam), Snail (1:1000 dilution, Cat # 3895S, Cell Signaling Technology), Twist1 (1:1000 dilution, Cat # T6451, Sigma Aldrich), N-cadherin (1:1000 dilution, Cat # 610920, BD Transduction Laboratories), Vimentin (1:1000 dilution, Cat # 550513, BD Pharmingen) or β-actin (1:4000 dilution, Cat # AB6276, Abcam). Membranes were incubated with horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody (1:4000 dilution, Cat # NA931V, GE Healthcare) or anti-rabbit secondary antibody (1:4000 dilution, Cat # NA934V, GE Healthcare) for 1 h and signals were developed using the ECL method (GE Healthcare).

Immunofluorescence. 4TO7 and 4T1 cells were seeded onto gelatin coated glass coverslips and placed in 24-well plates. 48 h subsequent to seeding, media was aspirated and cells were fixed with ice cold methanol for 10 min. Following PBS washes, cells were permeabilized with 0.2% Triton for 3 min and subsequently blocked in 10% goat serum for 1 h at room temperature. E-cadherin or N-cadherin was probed with mouse antibody to E-cadherin or N-cadherin respectively (1:200 dilution) for 1 h at room temperature followed by detection of the primary antibody with a rhodamine-conjugated goat anti-mouse secondary antibody (1:500 dilution) for 1 h at room temperature. Hoechst dye (1mg/ml) was subsequently used to stain nuclei (1:1000

Nature Medicine: doi:10.1038/nm.2401

dilution). Cells were observed on a Zeiss microscope and pictures were taken using an Axiocam Icc3 camera with optical deconvolution.

Histology and Immunohistochemical staining. Lungs were excised, fixed in 10% neutral-buffered formalin or bouins fixative for 24 h, embedded in paraffin, sectioned at 5 μm thickness and stained with hematoxylin and eosin (H&E). A few overt pulmonary lesions were dissected prior to fixation, homogenized in lysis buffer (Ambion), total RNA was extracted using the mirVana microRNA isolation kit and PCR analysis was performed using reverse transcribed RNA samples to ensure stable overexpression of miR-200s, enhanced expression of E-cadherin in miR-200 overexpressing lesions and knockdown of *Sec23a*. Immunohistochemical staining of E-cadherin was performed on paraffin embedded lung sections. Sections were dehydrated, treated with 0.3% hydrogen peroxide/methanol for 20 min at RT, boiled in 10 mM sodium citrate (pH 6.0) for 30 min and blocked with 10% normal goat serum for 30 min (catalog no. 16210-064, Gibco) and avidin/biotin blocks for 15 min each at RT prior to staining. Incubation with primary antibody to E-cadherin (Cat # 610181, BD-Biosciences) at a 1:200 dilution was carried out overnight at 37 °C followed by a 30 min incubation with biotinylated anti-mouse secondary antibody at a 1:1000 dilution (Vectastain ABC Kit Rabbit IgG, Cat # PK-4001, Vector Laboratories) and the DAB detection kit (00-2014, Zymed) according to the manufacturer's instructions.

Quantification of number of macro- and micrometastases. The number of overt macrometastases was counted manually on fixed lungs. Two-sided independent Student's t-test without equal variance assumption was performed to assess statistical differences between groups. To quantify micrometastases, mice were sacrificed ~3–4 weeks following orthotopic inoculation, lungs were excised, minced, digested in 300 U/ml collagenase type 1A (Sigma, Cat # C2674)/ 100 U ml^{-1} hyaluronase (Sigma, Cat # H3506) enzyme cocktail at 37 °C for 60–90 min, washed and strained using a 70 μm strainer and plated in 60 μM 6-thioguanine selection in two 15 cm tissue culture plates (serve as duplicates). Following 1–2 wks of selection, tumor colonies were stained with crystal violet for 30 min, rinsed with ultrapure water and dried overnight prior to counting.

Nature Medicine: doi:10.1038/nm.2401

Quantification of tumor cells in circulation. Whole blood was collected from the retroorbital sinus weekly from mice orthotopically injected with tumor cells in the mammary gland. Genomic DNA was isolated from 50 µl of whole blood using the DNeasy Blood & Tissue kit (Qiagen) and genomic PCR was performed using primers designed to amplify a region of the puromycin resistance gene (part of the pMSCV-puro vector backbone stably integrated in miR-200 overexpressing lines). Primer sequences are: 5'-ATCGGCAAGGTGTGGGTCGC-3' (forward), and 5'-GCGCCAGGAGGCCTTCCATC-3' (reversed).

Mass spectrometry. Mass spectrometry was performed on two biological replicates of control and C1+C2 lines. To maximize the number of proteins detected by MS, the stable isotope labeling of amino acids in cell culture (SILAC) approach was applied. Heavy labeled isotopic arginine (Sigma, Cat # 608033) and lysine (Sigma, Cat # 608041) was used to label the proteome of the C1+C2 line whereas naturally occurring arginine (Cat # A5006) and lysine (Sigma, Cat # L5501) was used to label the proteome of the vector control line. Cells were grown in SILAC media (Sigma, D-9943) supplemented with 10% dialyzed FBS (Invitrogen, 26400-044) and isotopic/non-isotopic arginine and lysine. Following ten population doublings, nuclear and cytoplasmic fractions were isolated, reduced and alkylated with DTT and iodoacetamide. Modified trypsin enzyme was subsequently added to perform in-solution digestion overnight. The trypsinzed peptides were desalted on a POROS C18 column before the strong cation exchange (SCX) steps. SCX fractionation was done on a LC Packing nanoHPLC system. A total of 12 fractions were collected for each sample and then desalted with POROS C18 ziptips for mass spectrometry analysis. The capillary nano LC-MS/MS analysis of each SCX fraction were performed using an LTQ-Orbitrap hybrid mass spectrometer interfaced with an Eksigent nanoHPLC system. Samples were loaded onto a pulled fused silica microcapillary column packed with C18 reverse phase resin using an Eksigent autosampler. After loading, samples were separated on a fused silica nano column (20 cm long, 75 µm inner diameter, packed with 5 µm C18 resin) with a three-hour gradient run at 300 nl min^{-1} flow rate. The data dependent MS/MS scans were done using the top seven most abundant ions and collision-induced dissociation (CID) for fragmentation.

Arg-10 and Lys-8 labeled peptides were quantified using area under extracted ion chromatograms (XICs). XICs were found and paired using the previously described methods[21]. The ratio of the areas under the paired XICs was reported as the ratio between heavy and light versions of peptides. MS/MS spectra were searched using IPI version 3.59 mouse protein sequence database. MS/MS database search and quantification were conducted using a 10 ppm precursor mass window using publicly accessible software PVIEW (http://compbio.cs.princeton.edu/pview). MS/MS spectra were filtered to remove peaks caused by isotopes and noise. Up to one missed cleavage was allowed for database search. MS/MS spectra were assigned an amino acid sequence using a high confidence <1% false discovery rate (FDR). FDR was computed using a concatenated and reverse decoy database in which lysine and arginine were swapped to remove precursor mass correlations[22]. For proteins quantified by multiple peptides, the median ratio of all of the peptides was assigned to the protein to eliminate any outlier ratios. Protein ratios were normalized using all detected XIC pairs so the median of their logarithm was zero, correcting for unequal loading of light and heavy sample. Further, MS/MS database search results where the number of lysines and arginines did not match XIC pair spacing were removed.

Secretome analysis. The SILAC-based approach was applied for mass spectrometry analysis of CM samples. Equal numbers of cells were plated and after attachment, cells were washed twice with PBS and overlaid with DMEM without serum and allowed to grow for another 24 h before the media was collected (light label for control and heavy label for Sec23a KD or C1+C2). The media was filtered through a 0.45 μm membrane to remove cellular debris and intact cells to yield the conditioned medium. Mass spectrometry on conditioned media samples was performed as described above. Protein sequences of all detected proteins were inputted into Signal P v3.0 to identify secreted proteins containing signal sequences. Those proteins lacking signal sequences were not used for subsequent analyses.

Electron microscopy. The electron microscopic analysis of endoplasmic reticulum structure was performed on 70% confluent cells (Sec23a KD, C1C2 and their respective vector control lines). Cells grown on tissue culture dishes were fixed at RT for 1 h using 1.6% paraformaldehyde/2.5% glutaraldehyde in a 0.1 M Na Cacodylate buffer, pH 7.4.

Nature Medicine: doi:10.1038/nm.2401

The cells were then removed from the tissue culture dishes by scraping into 1ml of freshly applied fixative, spun at 13,000 g to produce a pellet, and stored for further processing in fixative at 4 °C. Subsequently, routine electron microscopy procedures were followed. Briefly, samples were post-fixed in 1% osmium tetroxide for 1 h at 4 °C, post-fixed again and en bloc stained with 0.5% uranyl acetate O/N at 4 °C, dehydrated, and subsequently embedded in Epon 812 resin for sectioning. 60nm sections were cut, stained with lead citrate - Pb/Uranyl acetate, and profiles were imaged with a Zeiss 912AB electron microscope (Thornwood, NY) operated at 100 Kv.

Gene set enrichment analysis (GSEA). We used GSEA v2.0[23]. Normalized microarray expression data with at least one experimental replicate per phenotype were rank-ordered by expression using the provided ratio of classes (i. e. fold change) metric. Multiple probe matches for the same gene were collapsed into one value, with the median probe reading being used in each case. Gene sets were obtained either from the MSigDB database v2.5 (in the case of GO categories) or from our own ranked lists, with the top 100 genes from a given list being used as a set. Only GO category gene sets with between 15 and 500 gene matches in the ranked list were used (n = 645 of the 1454 in C5 of MSigDB v.2.5). Gene sets were tested for enrichment in rank ordered lists via GSEA using a classic statistics and compared to enrichment results from 1000 random permutations of the gene set to obtain P values. Raw enrichment scores were converted to normalized enrichment scores using default GSEA parameters. Enrichment cores were defined as the members of the gene set that lie before or at the running sum peak (i. e. the enrichment score) of the ranked gene list.

miRNA and mRNA expression levels across NCI-60. The mRNA and miRNA expression data were downloaded from Cellminer (http://discover.nci.nih.gov/cellminer) on May 3rd, 2009. The expression levels were then variance normalized across the 59 samples prior to analysis. The expression of the nine-gene signature was averaged and the resulting values were compared with those of miRNA-200b, c using spearman rank correlation and the resulting ρ and p values were reported.

Multiple regression analysis for association of candidate target genes and secretome genes with RFS. We built a database combining the samples reported in

Nature Medicine: doi:10.1038/nm.2401

GSE11121, GSE7390, GSE2990 and GSE2034 deposited in GEO (http://www.ncbi.nlm.nih.gov/geo/)[24]. These datasets included subsets of individuals who did not receive systemic treatment, thus comprising a truly prognostic dataset[16]. All four databases contain Affymetrix HG-U133A (GPL96) microarrays along with clinical survival information. This dataset contains expression and survival information for 809 individuals with breast cancer who are lymph-node-negative. Among them, 584 (72%) are ER^+ and 253/281/163 (31%/35%/20%) are grade1/2/3. There are a total of 293 (36%) relapse events detected with a median RFS of 92 months. Gene expression values for 18,095 RefSeq transcript IDs across 746 samples were calculated using the "probe-to-transcript ID" mapping provided in GPL96 platform specification. The gene expression values for our eight of nine candidate miR-200 target genes from **Fig. 4a,b** (1 gene was not included in the analysis since it lacked a RefSeq ID) and the 35-gene secretome signature, genes that were significantly reduced in secretion in both Sec23a KD lines (three genes were not included in the analysis since they lacked RefSeq IDs), were extracted and variance-normalized. We then averaged the normalized gene expression values to obtain a candidate miR-200 target gene "signature" and the secretome gene "signature". For Kaplan-Meier plots, the samples were divided into "low-expression" and "high-expression" sets using the median as the threshold. A PERL script was then used to calculate survival rate for the two sets in each unit of time. As a control, 100 random sets of eight- and 35-genes were generated. The 10-year survival graphs were then plotted using the PostScript module and logrank *P* values were calculated using the Log Rank PERL module.

Analysis for association with survival and incidence/sites of metastasis. Kaplan-Meier plots were used to estimate distant metastasis-free survival in the EMC286 dataset[25] stratified into high and low gene expression groups. High expression was defined as tumors expressing the gene of interest within the upper quartile expression range. All other tumors (quartiles 1–3) were considered to have low gene expression. Significance of survival differences between groups was assessed by log-rank test, with $P < 0.05$ being considered significant.

To assess association with incidence/sites of metastasis, gene expression of MSK82 collection of tumors[10] were grouped according to incidence and site of metastasis, and

Nature Medicine: doi:10.1038/nm.2401

differences in group expression means were assessed by two-sided Student's t-tests with the assumption of equal variance. * $P < 0.05$, ** $P < 0.01$.

Nature Medicine: doi:10.1038/nm.2401

SUPPLEMENTARY REFERENCES

1. Santner, S.J., Dawson, P.J., Tait, L., Soule, H.D., Eliason, J., Mohamed, A.N., Wolman, S.R., Heppner, G.H. & Miller, F.R. Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. *Breast Cancer Res Treat* **65**, 101-110 (2001).

2. Chaffer, C.L., Brennan, J.P., Slavin, J.L., Blick, T., Thompson, E.W. & Williams, E.D. Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2. *Cancer research* **66**, 11271-11278 (2006).

3. Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S., Backlund, M.G., Yin, Y., Khramtsov, A.I., Bastein, R., Quackenbush, J., Glazer, R.I., Brown, P.H., Green, J.E., Kopelovich, L., Furth, P.A., Palazzo, J.P., Olopade, O.I., Bernard, P.S., Churchill, G.A., Van Dyke, T. & Perou, C.M. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8**, R76 (2007).

4. Gupta, P.B., Onder, T.T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R.A. & Lander, E.S. Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138**, 645-659 (2009).

5. Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y., Gupta, S., Yang, J., Hartwell, K., Onder, T.T., Gupta, P.B., Evans, K.W., Hollier, B.G., Ram, P.T., Lander, E.S., Rosen, J.M., Weinberg, R.A. & Mani, S.A. 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* **107**, 15449-15454 (2010).

6. Blick, T., Hugo, H., Widodo, E., Waltham, M., Pinto, C., Mani, S.A., Weinberg, R.A., Neve, R.M., Lenburg, M.E. & Thompson, E.W. Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer. *J Mammary Gland Biol Neoplasia* **15**, 235-252 (2010).

7. Giampieri, S., Manning, C., Hooper, S., Jones, L., Hill, C.S. & Sahai, E. Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. *Nat Cell Biol* **11**, 1287-1296 (2009).

8. Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S. & Brabletz, T. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO reports* **9**, 582-589 (2008).

9. Kang, Y., Siegel, P.M., Shu, W., Drobnjak, M., Kakonen, S.M., Cordon-Cardo, C., Guise, T.A. & Massague, J. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537-549 (2003).

Nature Medicine: doi:10.1038/nm.2401

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

11. Bos, P.D., Zhang, X.H., Nadal, C., Shu, W., Gomis, R.R., Nguyen, D.X., Minn, A.J., van de Vijver, M.J., Gerald, W.L., Foekens, J.A. & Massague, J. Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005-1009 (2009).

12. Kaplan, R.N., Riba, R.D., Zacharoulis, S., Bramley, A.H., Vincent, L., Costa, C., MacDonald, D.D., Jin, D.K., Shido, K., Kerns, S.A., Zhu, Z., Hicklin, D., Wu, Y., Port, J.L., Altorki, N., Port, E.R., Ruggero, D., Shmelkov, S.V., Jensen, K.K., Rafii, S. & Lyden, D. VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820-827 (2005).

13. Kim, S., Takahashi, H., Lin, W.W., Descargues, P., Grivennikov, S., Kim, Y., Luo, J.L. & Karin, M. Carcinoma-produced factors activate myeloid cells through TLR2 to stimulate metastasis. *Nature* **457**, 102-106 (2009).

14. Grivennikov, S.I., Greten, F.R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883-899 (2010).

15. Tan, W., Zhang, W., Strasner, A., Grivennikov, S., Cheng, J.Q., Hoffman, R.M. & Karin, M. Tumour-infiltrating regulatory T cells stimulate mammary cancer metastasis through RANKL-RANK signalling. *Nature* **470**, 548-553 (2011).

16. Gyorffy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q. & Szallasi, Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast cancer research and treatment* **123**, 725-731 (2010).

17. Hu, G., Chong, R.A., Yang, Q., Wei, Y., Blanco, M.A., Li, F., Reiss, M., Au, J.L., Haffty, B.G. & Kang, Y. MTDH activation by 8q22 genomic gain promotes chemoresistance and metastasis of poor-prognosis breast cancer. *Cancer Cell* **15**, 9-20 (2009).

18. Camps, C., Buffa, F.M., Colella, S., Moore, J., Sotiriou, C., Sheldon, H., Harris, A.L., Gleadle, J.M. & Ragoussis, J. hsa-miR-210 Is induced by hypoxia and is an independent prognostic factor in breast cancer. *Clin Cancer Res* **14**, 1340-1348 (2008).

19. Gee, H.E., Camps, C., Buffa, F.M., Colella, S., Sheldon, H., Gleadle, J.M., Ragoussis, J. & Harris, A.L. MicroRNA-10b and breast cancer metastasis. *Nature* **455**, E8-9; author reply E9 (2008).

20. Brummelkamp, T.R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science (New York, N.Y* **296**, 550-553 (2002).

Nature Medicine: doi:10.1038/nm.2401

21. Khan, Z., Bloom, J.S., Garcia, B.A., Singh, M. & Kruglyak, L. Protein quantification across hundreds of experimental conditions. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15544-15548 (2009).

22. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**, 1367-1372 (2008).

23. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. & Mesirov, J.P. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550 (2005).

24. Gyorffy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q. & Szallasi, Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* (2009).

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

Nature Medicine: doi:10.1038/nm.2401


Resum
en
català

Joan Oró (Lleida, 1923 - Barcelona, 2004)

1. INTRODUCCIÓ

El càncer és una malaltia molt diversa, intervenen multitud de factors biològics i compren multitud de tipologies diferents. L'evolució del càncer passa per moltes etapes diferents, on cada una d'elles poden ser de gran interès per aplicacions terapèutiques, des del seu origen i creixement com a tumor primari, adquisició de propietats malignes, fins a la seva disseminació a altres òrgans distants. Aquesta disseminació és el que s'anomena metàstasis tumoral i és el procés responsable de més del 90% de morts causades pel càncer. És un punt on la malaltia ja ha evolucionat a un estadi molt avançant i es perd el control de la seva ubicació per part del clínic, per tant complica la bona prognosis del pacient. Evitar aquest procés ha de ser de màxim interès terapèutic per tal d'evitar la progressió de la malaltia, controlar-la i així finalment, curar-la. Per això en l'ultima dècada molts esforços en recerca s'han centrat en entendre i abordar terapèuticament la metàstasis, ja que anteriorment la gran majoria d'esforços en recerca es centraven en entendre els mecanismes de la transformació oncogènica. El propi procés de la metàstasis també consta de varies etapes, des de l'adquisició de característiques per poder sortir del tumor primari, arribar a la circulació, sobreviure dintre del torrent sanguini, "elegir" i extravasar a un determinat òrgan i després poder sobreviure en el nou ambient on arriben, ja sigui de forma latent o creixent ràpidament.

Davant tota aquesta complexitat d'etapes a superar, s'han identificat tota una sèrie de propietats biològiques que la cèl·lula cancerígena ha d'adquirir. Segons Hanahan i Weinberg, aquestes es poden resumir com: proliferació sostinguda i auto-renovació, evasió de senyals supressores, resistència a mort, immortalitat replicativa, angiogènesis, capacitat invasiva, reprogramació metabòlica i evasió del sistema immunitari. Totes aquestes propietats es poden arribar a adquirir gràcies a la inestabilitat genòmica de les cèl·lules cancerígenes i heterogeneïtat tumoral.

1.1 Les etapes de la metàstasis

Com ja s'ha dit, és un procés de múltiples etapes on es requereix de diferents transformacions cel·lulars i participació de multitud de factors extra-tumorals, ja siguin diferents tipus cel·lulars com factors del microambient tumoral. Les etapes les resumirem com:

1- Adquirir el fenotip agressiu per part d'una part de la població tumoral que passarà a proporcionar una nova propietat al tumor, ja sigui per mutacions genètiques com per alteracions epigenètiques, tot provocat per l'entorn.
2- Adquirir capacitat invasiva. Passar de formar part d'un conjunt de cèl·lules no mòbils a adquirir propietats migratoris i invasives per poder rompre el teixit dels voltants.
3- Ser capaç d'intravasar al torrent sanguini i. Això és gràcies a la mobilitat d'aquest tipus de cèl·lules invasives i a més de la capacitat del propi tumor de produir factor angiogènics que faciliten la formació de vasos al voltants del tumor, ja siguin limfàtics o sanguinis.

4- Sobreviure en suspensió. Aquesta és una de les característiques més difícils d'aquirir com a propietat autònoma, ja que els tumors sòlids requereixen d'un substrat sòlid on mantenirse i fomentar el seu creixement. En absència d'ancoratge, les cèl·lules epitelials normalment moren per "anoikis". A més, les cèl·lules han de suportar les pressions físiques del torrent sanguini. Per això, algunes de les cèl·lules metastàtiques tenen l'habilitat de protegir-se envoltant-se d'escuts de plaquetes.
5- Han de reconèixer senyals dels endotelis de diferents òrgans per parar-se i penetrar dintre dels òrgans. En alguns casos la vasculatura és tan estreta que els propis grups de cèl·lules metastàtiques queden atrapats i això ajuda a extravasar, com per exemple en els pulmons.
6- Un cop dintre del teixit distant, degut a l'hostilitat dels nous teixits i diferents condicions respecte de l'original, molt poques cèl·lules tindran la capacitat d'adaptar-se, sobreviure i addicionalment poder iniciar el creixement d'un nou tumor.

Per tant, des del punta de vista de la cèl·lula cancerígena, una ínfima part de les cèl·lules que passen pel torrent sanguini arribaran a un teixit distant, i alhora, molt poques cèl·lules que arriben a un teixit distant seran capaces de sobreviure-hi i formar un tumor. Tot això fa que el procés de la metàstasis requereixi d'una multitud de dràstiques transformacions i alliberació de moltíssimes cèl.lules, per tal de poder habilitar que unes poques arribin al destí final.

A més, en tot això, existeix una sèrie de mecanismes d'especificitat a l'hora de colonitzar altres teixits. Determinats tipus de càncer colonitzen determinats tipus de teixits, ja sigui per atraccions quimiotàctiques per factors difusibles o per condicionaments físics d'accés a l'òrgan, com per compatibilitat amb la pròpia arquitectura i senyals del teixit de destí. Tot i que la proximitat física o vies directes a partir de la circulació sanguínia a un determinat lloc poden afavorir la metàstasis, no és suficient si no existeix compatibilitat entre els teixits diana i les cèl·lules tumorals. Alguns exemples: els tumor de mama solen metastatitzar a òs, pulmons, fetge i cervell; els de pròstata, principalment a òs però també a fetge i pulmó; els colorectals, a fetge i pulmons; alguns melanomes, a fetge i els sarcomes, a pulmó. Normalment és requereixen de mecanismes de retroalimentació positiva entre el teixit i les cèl·lules, on s'estableixen cercles virtuosos que empenyen el creixement metastàtic.

1.2 Heterogeneïtat i origen tumorals

En els tumors sòlids existeixen diferents subpoblacions tumorals amb diferents capacitats i característiques, es a dir, heterogeneïtat intratumoral. La gran heterogeneïtat intratumoral representa un dels grans reptes per comprendre els mecanismes i la dinàmica tumorals, i per tant, per tal d'abordar teràpies antitumorals més específiques. Per això resulta molt important saber determinar com s'origina el tumor i de quin tipus cel·lular es composen els diferents subtipus de tumors. Avui en dia els dos grans models per explicar la propagació tumoral són: (A)

el model d'evolució clonal i (B) el model/hipòtesis de cèl·lules mare cancerígenes, CSCs (sigles de l'anglès "Cancer Stem Cells").

El model d'evolució clonal és un model on la transformació maligna a partir d'acumulació de mutacions d'una cèl·lula original (o alteracions epigenètiques) produeix tota una varietat de noves generacions i característiques cel·lulars on aquelles més competitives seran les que lideraran el creixement tumoral. Es basa en un model evolutiu i estocàstic de deriva genètica.

Les CSCs són un petit grup de cèl·lules tumorals que tenen la capacitat d'iniciar el tumor i sostenir el seu creixement, ja que tenen la capacitat d'auto-renovar-se indefinidament al mateix temps que poden donar lloc a altres tipus cel·lulars (relativa pluripotència) més diferenciats que formarien la resta de poblacions del tumor. Aquestes sorgeixen de cèl·lules mare adultes en els teixits que degut a alteracions genètiques i/o epigenètiques perden el seu control homeostàtic i passen a créixer de forma descontrolada. O també poden sorgir de cèl·lules més diferenciades que al sofrir una transformació maligna retrocedeixen des-diferenciant-se i així adquireixen la capacitat d'auto-renovació. Aquest model també porta implícit una jerarquia. A més, explica com les CSCs poden ser les cèl·lules més àgils a l'hora de colonitzar un òrgan distant i també els casos de recurrència després de tractament, al ser resistents a molts agents quimioterapèutics.

Aquests dos paradigmes no són mútuament exclusius. És lògic pensar que les CSCs durant el curs tumoral segueixen un model d'evolució clonal i a la inversa: les CSCs que inicien un tumor poden anar adquirint mutacions o alteracions epigenètiques que les faran més competitives i rellevants per la progressió del tumor; al contrari, les cèl·lules malignes resultants d'una evolució clonal que derivin i adquireixin propietats auto-renovadores i pluripotents, tindran avantatges selectives per créixer sobra la resta de la població tumoral. Per altre part, això introdueix un component estocàstic al model de CSCs a més de processos de reversibilitat, ofereixin així una versatilitat i dinamisme al tumors que incrementa la seva complexitat i determinació de l'heterogeneïtat.

1.3 Cèl·lules mare i CSCs

Per entendre la biologia de les CSCs és necessari comprendre els mecanismes i propietats de les cèl·lules mare normals. D'entrada les cèl·lules mare es divideixen en cèl·lules mare embrionàries (ESCs, sigles de "Embryonic Stem Cells" en anglès) i les cèl·lules mare somàtiques (o adultes) (SSCs, sigles de "Somatic Stem Cells" en l'anglès). Totes elles tenen capacitat auto-renovadora, immortalitat, divisió asimètrica i control homeostàtic de la diferenciació a diferents tipus cel·lulars. Les ESCs són les veritablement pluripotents amb capacitat per donar lloc a qualsevol tipus de teixit, mentre que les SSCs només poden donar lloc als determinats tipus cel·lulars específics del teixit on resideixen. Una altre diferència important és que les ESCs es divideixen de forma molt més ràpida que les SSCs, que tenen una molt baixa taxa de proliferació [28]. A més, estan modulades per diferents vies de senyalització i expressen diferents factors de transcripció. En el cas de les ESCs, aquestes expressen

de forma general, factors de pluripotència com OCT4 (POU5F1), TCF3, SOX2, NANOG, KLF4 i MYC, entre d'altres. En els cas de les SSCs els complements gènics expressats són més variables i específics del teixit on es troben.

La presumpta absoluta irreversibilitat del procés de diferenciació es va veure trencada quan al 2006 al laboratori de Yamanaka van aconseguir reprogramar fibroblasts cap a un estat més pluripotent capaç de donar lloc a altres tipus cel·lulars totalment diferents, com ara cèl·lules cardíaques, musculars o òssies. Per reprogramar les cèl·lules a un estat induït de pluripotència (iPSCs, sigles de "induced Pluripotent Stem Cells" en anglès), al laboratori de Yamanaka van introduir 4 factors de transcripció de pluripotència: SOX2, KLF4, MYC i OCT4. Tot i aquest i altres avenços, la reprogramació permanent a un estat més pluripotent no deixa de ser de gran dificultat, ja que l'eficiència de reprogramació sol ser molt baixa i reversible. La capacitat de reprogramar de forma estable i eficient serà el que proporcionarà importants avanços en el camp de la biomedicina, tant en regeneració de teixits com en el camp de l'oncologia.

D'aquesta manera podem entendre millor la biologia de les CSCs, tot i que la seva caracterització tampoc està ben definida degut a la alta variabilitat dels marcadors i de la heterogeneïtat dinàmica intratumoral, que canvia al llarg de la progressió del tumor. Per exemple, el compartiment de CSCs en el tumor pot variar entre el 0.1% fins el 30%, depenent del tipus de tumor. Les CSCs en alguns casos s'han considerat com a poc proliferatives però recentment la majoria d'estudis aposten per considerar-les com a altament proliferatives, aquesta i altres similituds fan que les CSCs es considerin biològicament més pròximes a les ESCs que les SSCs.

Una altre característica important de les CSCs que comparteixen amb les cèl·lules mare normals, és la resistència a agents quimiterapèutics i a la radioteràpia. Comparteixen mecanismes comuns que eviten l'acumulació d'aquests. A més, gràcies a la seva plasticitat, poder superar amb relativa facilitat, la pressió selectiva que puguin tenir determinats agents quimioterapèutics i esdevenir-ne resistents.

1.4 El nínxol

El nínxol on resideixen les cèl·lules mare és l'altre component que determina la homeòstasis i característiques d'aquestes. Les cèl·lules mare requereixen d'un nínxol específic amb una arquitectura determinada on rebran diferents senyals del seu entorn més pròxim i mantindran la seva dinàmica de diferenciació. Paral·lelament, les CSCs també depenen d'un nínxol per mantenir les seves propietats, però aquest, al igual que les pròpies CSCs, resulta alterat, desorganitzat i a més és clau en el procés tumorigènic. Així, el nínxol pot regular l'auto-renovació, proliferació, diferenciació i resistència a mort de les CSCs. Està compost de cèl·lules estromals com ara cèl·lules mesenquimals, immunitaris, vasculars, així com de factors solubles i de factors de la matriu extracel·lular. Aquesta dependència es veu ressaltada quan la pèrdua del nínxol acaba en la pèrdua de les pròpies CSCs.

A part de la seva importància pel tumor primari, s'ha vist que el nínxol també és determinant per la formació de metàstasis secundàries a distància. Fins i tot existeixen mecanismes pels quals senyals enviades pel propi tumor primari, acomoden el nínxol a un punt distant abans de establir la seva metàstasis a aquesta localització secundària, mitjançant el reclutament de cèl·lules que provenen del moll de l'òs.

1.5 La metàstasis i les CSCs

Les noves tècniques de l'era post-genòmica permeten diferents aproximacions i sobretot de forma més global en l'estudi de xarxes transcripcionals característiques de determinats estadis del tumor. D'aquesta manera s'han descobert perfils transcripcionals comuns entre ESCs i CSCs que tenen la seva representació en diferents tipus de carcinomes. S'ha pogut comprovar com aquestes xarxes gèniques pròpies de CSCs i també de ESCs estan més actives en tumors epitelials més agressius i metastàsics, i mostren una associació directa amb la mortalitat dels pacients. Aquests resultats reforcen la idea que les CSCs, a més del fet que tenen les propietats idònies com per acomodar-se en un teixit distant i iniciar un nou tumor, són la subpoblació tumoral responsable del procés metastàtic.

1.6 La transició epiteli-mesènquima en tumors

La transició epiteli-mesènquima o EMT (sigles de "Epithelial to Mesenchymal Transition"), es considera tot un programa genètic que produeix una sèrie de transformacions i remodelacions del citoesquelet cel·lular per tal d'adquirir un fenotip més migratori i invasiu. És un procés molt important durant el desenvolupament de l'embrió i té participació directa en la formació del mesoderma, cresta neural i altres teixits mesenquimals. A l'any 2000, dos laboratoris d'Espanya, paral·lelament, van descriure la seva important implicació en càncer. Els tres canvis fonamentals d'una EMT són:

- Canvis morfològics de pèrdua de polaritat apico-basal de les cèl·lules epitelials, passant a una morfologia fibroblastoide, dispersa i estirada amb fil·lopodis i pseudopodis.
- Reestructuració del citoesquelet i pèrdua de molècules de adhesió epitelial. La pèrdua de la E-cadherina és considera com la més clàssica i important per passar a un fenotip mesenquimal.
- Posta en marxa de mecanismes de migració i invasió cel·lular amb la capacitat de degradar molècules de la matriu extracel·lular.

La EMT és un procés reversible, ja que en alguns casos poden retornar al fenotip epitelial original un cop realitzada la seva funció, si les condicions ho requereixen, procés que es coneix com transició mesènquima-epiteli o MET (sigles de Mesenchymal to Epithelial Transition). La MET s'ha descrit tant en processos que es produeixen durant el desenvolupament embrionari com en processos tumorals, bàsicament en la metàstasis.

La EMT és un procés altament regulat per multitud de senyals extracel·lulars, components de la matriu extracel·lular i factors solubles. A més, existeix un cert entrecreuament de senyals o "crosstalk" entre diferents vies amb resultats finals comuns com la repressió de la E-cadherina. Per això, aquestes vies normalment acaben induint l'expressió d'una sèrie de factors de transcripció que entre d'altres funcions tenen la capacitat de reprimir la E-cadherina, ja sigui de forma directa, com SNAI1, SNAI2, ZEB1, ZEB2, E47, o KLF8, o indirecta com TWIST1, TWIST2, FOXC2, E2.2 o Goosecoid. A més, aquests factor de transcripció tenen la capacitat de modular-se entre ells, per exemple, SNAI1 indueix indirectament l'expressió de ZEB1 i ZEB2.

Els factors Snail pertanyen a la família més clàssica d'inductors de EMT, i s'han descrit en tot els processos de EMT estudiats. SNAI1 és l'inductor més potent de EMT dintre i fora de la familia, ja que exerceix una forta repressió de la E-cadherina i a més controla molts dels altres factors de EMT. De forma jeràrquica, SNAI1 seria el que inicia la EMT i el altres, SNAI2, ZEBs, TWISTs, s'encarreguen de mantenir el fenotip mesenquimal invasiu. La repressió la porten a terme gràcies a dominis de reconeixement d'elements E-Box del ADN situats a la regió C'-terminal. Els E-box són seqüències curtes en l'ADN (consens 5'-CACCTG-3'), de les que el promotor del gen de la E-cadherina en té varis. A l'extrem N-terminal tenen el domini SNAG a través del qual Snail recluta diferents cofactors que executen remodelacions de la cromatina associada a la regió promotora i reprimeixen transcripcionalment l'expressió del gen diana com ara la E-cadherina. A més de la E-cadherina, els factors Snail també regulen i reprimeixen, directa o indirectament, altres gens epitelials, així com altres gens per inhibir la proliferació, activen marcadors mesenquimals, i gens que codifiquen per proteïnes implicades en invasió i degradació de matriu extracel·lular i supervivència. Una altre família important són els factors ZEB, que estan controlats per la família miR-200 de microRNAs, i, a la vegada, els factors ZEB reprimeixen els miRs-200, establint així un circuit de control recíproc. Aquest circuits no deixen d'estar influenciats per l'SNAI1, ja que aquest inhibeix els miR-200s i per tant afavoreix l'activació dels factors ZEB.

El fenotip mesenquimal arrel d'una EMT no només proporciona una major capacitat invasiva, també implica una pèrdua de la polaritat; capacitat per escapar de la mort cel·lular programada, "apoptosis"; mecanismes de resistència a la quimioteràpia, tals com oxaliplatí, paclitaxel o adriamicina, ja que aquest estan dissenyats per matar cèl·lules molt proliferatives però aquestes no són altament proliferatives; mecanismes de resistència al sistema immunitari; i, finalment, s'ha descrit que en alguns casos poden generar propietats de cèl·lula mare (tema que serà discutit més avall).

1.7 Els mecanismes invasius

Durant el procés metastàsic, determinades cèl·lules canvien de lloc, la qual cosa implica l'adquisició de capacitat per migrar, gràcies a modificacions de les molècules adhesió cèl·lula a cèl·lula i de cèl·lula a matriu extracel·lular, i a reordenaments del citoesquelet amb extensió de fil·lopodis. A part d'adquirir mobilitat, les cèl·lules tumorals que metastatitzen necessiten obrir-se pas entre el

teixits que s'interposen al seu camí de sortida del lloc del tumor primari, per la qual cosa es fan servir d'activitats proteolítiques, per exemple, MMPs (metal·lo-proteases de matriu) per poder degradar matriu extracel·lular.

Des del punt de vista físicomecànic, una cèl·lula individual, a l'hora de migrar, inicia un procés cíclic que comença per l'elongació d'un pseudòpod en la direcció de migració; adhesió d'aquest i generació de força; establiment de punts de proteòlisis al voltant de la cèl·lula; contracció de l'actina-miosina a partir de l'activació de small GTPases i finalment retracció de l'uròpod de la cèl·lula. Aquest procés repetit tant cops com sigui necessari permet a la cèl·lula avançar entre els teixits. Però no tota invasió requereix que es faci a partir de cèl·lules individuals, i la migració/invasió pot donar-se de forma col·lectiva quan les adhesions entre cèl·lules es mantenen. D'aquesta forma un grup de cèl·lules pot avançar entre els teixits de forma que a la punta del grup hi hagi cèl·lules líder guiant a la resta i amb activitat proteolítica, obrint així una petita obertura que la resta de cèl·lules s'encarreguen d'acabar d'obrir mitjançant pressió mecànica des del darrera, organitzant-se en forma incisiva i procurant activitat proteolítica als marges del grup. Aquesta modalitat de migració col·lectiva es pot interpretar com la seguida per una única mega-cèl·lula.

Aquestes dues modalitats de invasió, individual i col·lectiva, poden presentar multitud de variants. La plasticitat cel·lular permet que la combinació d'elles es vagi modulant al llarg del camí invasiu segons el requeriments dels obstacles trobats. Així la EMT pot ser necessària en les cèl·lules de la punta d'una invasió col·lectiva, que, al mateix temps es pot acompanyar d'invasió ameboide per tal de passar per barreres estretes de la matriu extracel·lular. Al llarg del camí, a través de mediadors, hi haurà tota una modulació recíproca entre l'estroma i les cèl·lules tumorals, que fan que el procés sigui dinàmic i plàstic, de manera que la cèl·lula tumoral que arriba a un punt final haurà sofert una sèrie de transformacions que la faran fenotípicament diferent d'aquella que va iniciar el procés.

Evidentment, al llarg de tot aquest procés invasiu, sigui individual o col·lectiu, intervenen molt factors. Els receptors de matriu extracel·lular participen activament durant el procés migratori, ja que constitueixen un sistema de mecanotransducció entre els components de la matriu que connecten amb el citoesquelet, activant cascades de senyalització que engeguen mecanismes per generar forces d'adhesió i tracció, com bé fan les integrines. Per altre banda, les molècules d'adhesió cèl·lula a cèl·lula transmeten al citoesquelet les forces d'adhesió entre cèl·lules, a més de mantenir la cohesió entre cèl·lules, afavorint així la invasió col·lectiva, tot i que també poden intervenir en la invasió individual. La E-cadherina seria una de les molècules més rellevants en aquests processos, ja que s'encarrega de mantenir les cèl·lules unides, i també evita la seva migració al inhibir Rac1, estabilitzant així els teixits epitelials. La pèrdua d'expressió o funció de la E-cadherina és un dels requisits per la EMT i invasió individual, i a més té efectes directes sobre la senyalització i modulació del citoesquelet. Com a conseqüència de la pèrdua de la E-cadherina, la β-catenina és alliberada i si la via de Wnt està activa, pot traslocar-se al nucli i juntament amb els factors transcriptionals TCF activar gens implicats en migració. Un altre conseqüència de la pèrdua de la E-cadherina és que també s'allibera p120, que

activa Rac1 i Cdc42, els quals modulen el citoesquelet d'actina induint la formació de fil·lopodis i lamel·lipodis. Fins ara s'havia considerat la E-cadherina com un supressor de la progressió tumoral però s'ha demostrat que pot tenir un paper rellevant al coordinar els moviments en la invasió col·lectiva de determinats càncers, i per tant es pot considerar també com un inductor de la progressió tumoral, i més si tenim en compte que afavoreix la supervivència en la circulació i la colonització de teixits distants.

Les proteases i citocines també juguen un paper molt important durant la invasió cel·lular. Les proteases degraden matriu extracel·lular, alliberen components de la matriu que tenen activitat per promoure la invasió, promouen senyals per afavorir angiogènesis, tallen molècules d'adhesió, afavoreixen la renovació del receptor cel·lulars Hi ha diferents tipus de proteases i alhora diferent tipus de inhibidors de les proteases, ja que són proteïnes que requereixen d'un estricte control per mantenir la homeòstasis tissular, doncs la seva desregulació pot comportar alteracions fisiològiques i patologies fibròtiques, immunitàries i de progressió tumoral. Recentment, s'ha vist que durant la progressió tumoral participa tota una xarxa de molècules que, col·lectivament, aconsegueixen degradar el teixit circumdant els tumors. Les proteases es classifiquen segons el seu mecanisme catalític: aspàrtic-; cisteïna-; metal·lo-; serina- i treonina-proteases. I els seus inhibidors són les cistatines, serpines i TIMPs (sigles de "Tissue Inhibitors Metalloproteinases). El balanç de la xarxa proteolítica pot determinar l'esdevenir del teixit. En el cas de les citocines són les que s'encarreguen de la comunicació a llarga distància, poden activar invasió, motilitat de cèl·lules estromals, reclutament de tot un ventall de cèl·lules immunitàries que contribuiran en la progressió del tumor (macròfags,...), mobilització de cèl·lules del moll de l'os, establiment de nínxols pre-metastàtics i altres funcions.

1.8 Regulació epigenètica i RNAis no codificants en càncer

La progressió tumoral no sols depèn de les mutacions gèniques que s'adquireixen al llarg de les divisions cel·lulars, sinó que també els mecanismes epigenètics que regulen l'expressió gènica poden alterar de forma estable l'expressió de gens implicats en la progressió. Els mecanismes epigenètics que alteren l'expressió estan molt bé regulats però quan aquesta regulació és alterada tenen el poder de causar canvis importants i de forma duradora en l'expressió gènica. Aquests mecanismes són:

- La metilació de la seqüència de l'ADN, on les anomenades illes CpG (compostes de nucleòtids de citosina i guanidina) properes a promotors gènics són metilades per les DNMTs (sigles de "DNA Methyltransferases") i així reprimeixen l'activitat del promotor en qüestió.
- Remodelació de la cromatina. La cromatina és aquella estructura que conformen, juntament amb les histones, un ADN ordenadament empaquetat, però que necessàriament requereix d'un control dinàmic d'obertura-compactació, ja que per l'expressió gènica es necessita una cromatina oberta per permetre l'accés de la maquinària transcripcional al lloc

de transcripció. Així, els factors remodeladors de cromatina tenen la capacitat de modificar les càrregues i la conformació de les histones de tal forma que relaxen o compacten la cromatina d'una determinada regió, així activant o reprimint la transcripció. Per exemple, la metilació de la lisina 27 de la histona H3, causa una repressió transcripcional del gen sota el control del promotor enriquit en aquesta modificació d'histones, i aquesta és una típica marca induïda per l'SNAI1 a través del grup de proteïnes polycomb per reprimir la transcripció de la E-cadherina durant la EMT.

- Pèrdua de impromta. Es refereix a l'activació o repressió per alteracions epigenètiques d'un determinat al·lel de la línia germinal, i que per tant es transmetrà a les cèl·lules somàtiques de la descendència.

D'aquesta manera, els canvis epigenètics poden integrar multitud de senyals del microambient per acabar regulant sencers programes gènics de les cèl·lules tumorals. Jeràrquicament, aquests mecanismes són epistàsics respecte d'altres grups de gens reguladors del fenotip cel·lular. Les modificacions epigenètiques van modulant i marcant el camí, juntament amb les mutacions, de les cèl·lules malignes. Hi ha evidències que alteracions en els mecanismes epigenètics poden estar a l'origen del fenotip neoplàsic en molts casos , de forma al menys tan prominent com l'aparició de mutacions genètiques.

Els microRNAs són ARNs no codificants d'uns 21-23 parelles de bases. Aquests curts fragments d'ARN tenen seqüències d'uns 7-8 nucleòtids de reconeixement sobre el 3'-UTR de l'ARNm (ARN missatger) de determinats gens. Aquests, al unir-se als ARNm provoquen un bloqueig de la seva transcripció o la seva traducció, o afavoreixen la seva degradació mitjançant el complex RISC, de tal manera que causen una disminució dels nivells del missatger diana i de la seva proteïna codificada. Durant el últims anys s'han descobert molts microRNAs que participen en la regulació de gens claus en diferents punts de la cascada metastàsica. Alguns autors han batejat aquests microRNAs amb el nom de "metastamirs". Poden regular el fenotip CSCs, la inducció de EMT, l'apoptosis, l'angiogenèsis i la colonització metastàtica. Una de les famílies més estudiades de microRNAs implicats en metàstasis són la família dels miR-200s, dianes prominents dels quals són els ARNm de ZEB1 i ZEB2, causant la seva pèrdua d'expressió i d'aquesta manera mantenint el fenotip epitelial al impedir la repressió de la E-cadherina per part dels ZEBs. De manera recíproca, els factors ZEB i SNAI1 poden inhibir l'expressió dels miR-200, desplaçant així l'equilibri cap a la repressió de la E-cadherina. Resulta interessant que els miR-200 estiguin altament expressats en aquelles línies més metastàsiques, perquè això implica que el fenotip epitelial és important per la colonització d'un òrgan distant.

1.9 El microambient tumoral

Un tumor no només consta de cèl·lules malignes. Podríem entendre els tumor amb un compartiment de cèl·lules epitelials neoplàsiques (l'equivalent del parènquima d'òrgans normals) i un compartiment amb cèl·lules no tumorals que conformen l'estroma del tumor. Evidentment, la coexistència de tots aquests tipus

cel·lulars té una incidència directa sobre la progressió del tumor i existeix una intensa comunicació entre tots ells, tumorals i no tumorals. Unint aquest panorama amb el model d'origen i progressió tumoral basat en CSCs, resulta cada cop més evident que les senyals del microambient tumoral marquen la dinàmica i evolució de les CSCs. Així, experimentalment s'ha vist que al introduir CSCs o cèl·lules metastàtiques en un ambient diferent al d'origen del tumor, poden evolucionar de forma molt diferent i la progressió del tumor pot canviar radicalment.

Els diferents tipus cel·lulars que majoritàriament composen l'estroma d'un tumor són cèl·lules endotelials; perícits; cèl·lules del sistema immune, que participen de forma molt important en el creixement dels tumors, angiogènesis i reclutament d'altres cèl·lules a la zona tumoral, entre d'altres més coses; els fibroblasts associats al càncer, que són fibroblasts més reactius del normal i també tenen un paper molt important en la comunicació tumoral i alliberació de factors de creixement; i cèl·lules mare mesenquimals, que també intervenen en creixement, diferenciació i metàstasis de les cèl·lules neoplàsiques.

Altres factors del microambient que afecten el desenvolupament i progressió tumorals poden ser varis, sent un dels més rellevants la hipòxia. La falta d'oxigen desencadena tota un sèrie d'esdeveniments que promou la progressió del tumor. A través de l'activació del HIF-1, es promou l'angiogènesis, canvis metabòlics davant la falta d'oxigen afavorint la via glucolítica, proliferació, i també pot arribar a activar la transició epiteli-mesenquima (EMT) per tal d'afavorir la mobilització i trobar noves fonts d'oxigen. Per tant, la hipòxia resulta ser un fort inductor de la metàstasis, i per aquest mateix motiu molts del fàrmacs destinant a inhibir l'angiogènesis han fracassat, ja que al impedir l'abastiment d'oxigen a les cèl·lules tumoral aquestes posen en marxa mecanismes adaptatius que les converteixen en molt més agressives. El pH del microambient evidentment també influeix en el metabolisme i comoditat de les cèl·lules tumorals, de tal manera que l'acidificació del medi tumoral també pot afavorir l'angiogènesis i activació de proteases.

1.10 Controvèrsia i dinàmica entre els fenotips CSC i EMT

En cèl·lules no tumorals sembla clar que els estats de pluripotència i mesenquimals serien mútuament exclusius i contraposats. Al 2010 Li *et al*. i en paral·lel Samavarchi-Tehrani *et al*. varen demostrar com la transició mesènquima-epiteli era un requisit indispensable per reprogramar fibroblasts a iPSCs. Recíprocament, van demostrar que els factors de pluripotència de Yamanaka, KLF4, OCT4, SOX2 i MYC, inhibeixen la senyalització del TGFβ-SNAI1 i afavoreixen l'expressió de la E-cadherina. A més, al 2011 es va demostrar que la E-cadherina pot reemplaçar OCT4 i que és un requisit per la reprogramació a un estat més pluripotent. Per altre banda, també s'ha demostrar que SNAI1 indueix diferenciació de les ESCs i per tant pèrdua de pluripotència.

La dissociació d'aquests dos estats no s'ha demostrat de forma clara en càncer. L'article publicat al 2008 per Mani *et al*. va postular que la EMT genera propietats de cèl·lula mare a cèl·lules de carcinoma de mama. A partir d'aquest

altres estudis sorgeixen i reafirmen aquest solapament entre característiques mesenquimals i de cèl·lula mare. Ara bé, altres evidències difícilment concorden amb aquestes conclusions. Per una banda, és ben conegut que moltes metàstasis presenten un fenotip i marcadors epitelials. Per altre banda, és també evident que la EMT fa disminuir la capacitat proliferativa de les cèl·lules, cosa que no concorda amb l'alta activitat proliferativa pròpia de les CSCs. Altres estudis suggereixen directament que la EMT, tot i incrementant la invasió, suprimeix la capacitat metastàtica i que per això és necessària d'altres tipus de poblacions no mesenquimals per acabar produint una metàstasis. Així, miR-200s promouen la metàstasis precisament pel fet de induir un fenotip epitelial i de reduïda secreció de factors anti-metastàtics. Estudis de perfils transcriptòmics i epigenètics, posen de manifest que determinades signatures de gens implicats en ESCs estan directament relacionats i expressant en tumors de fenotip epitelials altament metastàtics.

Des del punt de vista biològic, sembla més lògic que les característiques de CSC i EMT estiguin representades per diferents poblacions no solapants en els tumors, que es puguin interconvertir les unes amb les altres però que l'adquisició d'un programa impliqui la supressió de l'altre per tal d'augmentar l'eficiència de la cèl·lula a l'hora d'assolir una funció determinada, ja que és difícil pensar que un cèl·lula altament proliferativa, que manté el creixement tumoral, estigui al mateix temps fent funcions d'invasivitat i migració. En un tumor donat existeixen diferents poblacions tumorals amb diferents capacitats, que duen a terme funcions diferents, possiblement complementàries, que repercuteixen en l'èxit col·lectiu de del conjunt del tumor. Per exemple, les cèl·lules tumorals amb característiques més mesenquimals es poden especialitzar en la tasca d'infiltrar teixits i obrir pas per la resta de població, mentre que les cèl·lules amb fenotip CSC serien aquelles que duen a terme l'auto-renovació de la població tumoral així com la generació dels diferents tipus de cèl.lules tumorals, tant al lloc inicial com a localitzacions distants o metàstasis. Aquestes capacitats estarien afavorides per CSC epitelials més que a les subpoblacions tumorals més mesenquimals perquè un fenotip epitelial potencia estats més pluripotent, evita la secreció de factors anti-metastàtics i permet viatjar agrupades per torrent sanguini per arribar en "equip" al teixit distant.

2. OBJECTIUS

1) Identificació de programes gènics diferencialment activats entre cèl·lules altament metastàtiques i cèl·lules poc metastàtiques d'un mateix tumor.

2) Identificació de mecanismes moleculars que regulen l'agressivitat de poblacions tumorals.

3) Identificació de gens d'agressivitat en mostres de tumors humans.

3. RESUM GLOBAL: RESULTATS, DISCUSSIÓ I CONCLUSIONS

Aquesta tesi es basa en l'observació, caracterització i anàlisis de diferències entre línies cel·lulars altament metastàtiques i línies poc metastàtiques. Tot i que en tots els casos es tractaria de cèl·lules metastàtiques, les diferències d'eficiència entre elles són enormes. S'han utilitzat models cel·lulars dual de càncer de pròstata, càncer de bufeta i càncer de mama, per tal d'entendre millor els mecanismes moleculars que defineixen les diferències pel que fa a capacitat metastàtica al llarg de les diferent etapes de la metàstasis entre el tipus cel·lulars estudiats. Aquest seria el principal objectiu global de la tesi. Presentada per dos articles, el primer article es centra en les interconversions fenotípiques i equilibri entre poblacions que promouen o suprimeixen les capacitats metastàtiques, i així quins són els mecanismes que regulen aquestes transicions. L'article 2, està més centrat en mecanismes concrets de regulació gènica a diferents nivells que influencien i determinen la capacitat metastàtica en els últims passos, en concret la colonització d'un teixit distant.

L'article 1, explica els resultats obtinguts de la comparació en models cel·lular duals com PC-3/Mc vs. PC-3/S i TSU-Pr1-B2 vs. TSU-Pr1. Ambdós models mostren una interessant dissociació entre dues de les capacitats fonamentals per metastatitzar que són la capacitat invasiva i la capacitat de créixer sense substrat sòlid i auto-renovació *in vitro* i metàstasis *in vivo*. En els dos casos de model cel·lulars, cada un dels subtipus provenen d'una mateixa línia parental. Les PC-3/Mc són un clon que provenen de les PC-3/M que són una subpoblació, provinent de les PC-3, seleccionada *in vivo* com a altament metastàtiques i les PC-3/S són un clon que provenen directament de les PC-3 parentals. Les TSU-Pr1-B2 provenen de la selecció *in vivo* també en funció de la seva capacitat metastàtica, provinents de les TSU-Pr1.

Llavors els resultats inicials obtinguts amb aquestes són interessants, ja que, aquelles que són altament metastàtiques *in vivo*, les PC-3/Mc o les TSU-Pr1-B2, són molt poc invasives en els nostres assaigs *in vitro*, però en canvi, i correlacionant amb la seva capacitat metastàtica, són altament formadores de colònies sense substrat sòlid (esferoides) i auto-renovadores. L'altre cara de la moneda són les PC-3/S i les TSU-Pr1, les quals són poc metastàtiques, poc formadores d'esferoides però molt invasives. Per tant, tenim en aquests models una important dissociació entre dues capacitats fonamentals per metastatitzar, perdent importància la capacitat invasiva degut als resultats finals obtinguts *in vivo*. L'anàlisi dels seus perfils transcriptòmics fou clau per tal d'entendre perquè aquestes cèl·lules es comportaven de tal forma. Així les PC-3/Mc expressen molt gens propis de ESCs, d'auto-renovació, pluripotència i conjuntament gens epitelials com la E-cadherina i altres. L'expressió d'aquestes signatures gèniques correlacionen amb el fet de formar més esferoides i de créixer i colonitzar millor teixits distants. En canvi, les PC-3/S expressen molts gens implicats en la transició epiteli-mesènquima (EMT) i secreció de diversos factors com les citocines.

Actualment, hi ha molts estudis centrats en aquests tipus de fenotips i un dels paradigmes més actuals, principalment en estudis de càncer de mama és que

quan les cèl·lules sofreixen una EMT també adquireixen propietats de cèl·lula mare cancerosa (CSC o TIC) i són més metastàtiques, tot i que els mecanismes pels quals adquireixen aquestes propietats no estan aclarits. En canvi, els nostres resultats amb els models de pròstata i bufeta utilitzats demostren justament el contrari, és a dir, que aquelles cèl·lules que han sofert una EMT, tot i ser més invasives, són menys metastàtiques, menys auto-renovadores.

Amb la intenció de caracteritzar millor el model de PC-3 i esbrinar quin tipus de dinàmica de poblacions segueixen, les PC-3/Mc sembla que es poden definir com $CD44^{+}$, $CD24^{+}$, $CD40^{-}$, $CD71^{+}$ i $CDH1^{+}$ i les PC-3/S com $CD44^{+}$, $CD24^{-}$, $CD40^{+}$, $CD71^{-}$ i $CDH1^{-}$. En models de càncer de mama són considerades com CSC-like aquelles poblacions definides com $CD44^{high}/CD24^{low}$, també s'han caracteritzat en altres models com en càncer de pròstata, però existeix molta controvèrsia al respecte perquè en altres models com el d'ovari, colorectal, pàncreas i també en pròstata es defineixen com $CD44^{high}/CD24^{high}$. A més en estudis de perfils transcripcionals propis d'ESCs lligats a agressivitat tumoral de càncer de mama, es descriu el CD24 com a altament expressat. Per tant, el camp dels biomarcadors per aïllar poblacions de CSC sembla ser massa ambigu com per tenir seguretat. En el nostra cas, hem utilitzat la E-cadherin (CDH1) per aïllar cèl·lules tipus PC-3/Mc a partir de les PC-3 parentals, d'aquesta manera aconseguim enriquir poblacions cel·lular amb cèl·lules tipus PC-3/Mc, tot i que l'eficiència no és suficient per tal d'aconseguir exactament la mateixa població, però almenys el seu comportament en assajos funcionals i d'expressió gènica *in vitro* mostren una important tendència cap a les PC-3/Mc.

Amb la intenció d'explorar millor les hipòtesis que la capacitat auto-renovadora i metastàtica va lligada a un fenotip epitelial hem (1) induït EMT a les PC-3/Mc sobreexpressant factors inductors d'EMT com l'SNAI1 (2) hem fet Knock-downs d'aquest factor d'EMT a les PC-3/S, (3) hem fet knock-down de la E-cadherin a les PC-3/M i recíprocament (4) sobreexpressat E-cadherina a les PC-3/S, (5) hem fet knock-downs del factors involucrats en auto-renovació i pluripotència en les PC-3/Mc, i hem (6) sobreexpressat SOX2 a les PC-3/S. Tots aquest experiments complementaris demostren la dissociació entre els fenotips d'EMT i epitelial-CSC. Els punts 1, 3, i 5 suprimeixen el fenotip epitelial-CSC i donen lloc a EMT disminuint la seva capacitat metastàtica, en canvi els punts 2, 4 i 6 suprimeixen l'estat EMT i activen el fenotip epitelial-CSC, augmentant la capacitat metastàtica. Alguns d'aquests punts també han estat aplicats al model de bufeta, com el punt 1 i el 5, i els resultats són igual o més espectaculars. Per tant, es demostra que aquests dos fenotips són mútuament excloents en els nostres models cel·lulars. A més resulta molt interessant que el fenotip epitelial resulti estar tan lligat al fenotip CSC (o TIC). Això, tot i anar en contra del model de càncer de mama, té molt sentit i correlaciona molt bé amb models d'estudis de pluripotència en fibroblasts. Aquests necessiten perdre els gens mesenquimals i guanyar E-cadherina i altres per poder reprogramar-se a un estat pluripotent amb l'expressió de gens com SOX2, KLF4, OCT4 o MYC que resulten ser tant rellevants en els nostres models. A més, un altre estudi en fibroblast ensenya com la E-cadherina pot arribar a substituir l'OCT4 alhora de reprogramar aquestes cèl·lules a un estat pluripotent. Per tant, l'E-cadherina pivota entre els dos estats afavorint la transició cap a la reprogramació i pluripotència. Els

nostres resultats mostren com la E-cadherina sola no és suficient per adquirir pluripotència però és clau per aconseguir l'estat de pluripotència acompanyant la regulació creuada (crosstalk) dels factors de Yamanaka (SOX2, KLF4, OCT4 i MYC) i afavorint les condicions físiques d'unió entre cèl·lules per arribar a aquest estat.

El nostres models també són recolzats per processos que tene lloc durant el desenvolupament embrionari, com és la formació de la cresta neural on l'activació de mecanismes migratoris suprimeix l'expressió de SOX2. La inducció d'EMT també s'ha descrit que disminueix la proliferació. A més altres suggereixen que l'EMT tot hi incrementar la invasió pot suprimir metàstasis. Però també existeixen molts més estudis descrivint el contrari. Aquestes diferències es postulen ser degudes a mecanismes intrínsecs dels propis models i contextos cel·lulars, per tant, la validació d'un no té per què negar l'existència de l'altre. Altres dades importants a tenir en compte és que les mostres histològiques de metàstasis de pròstata i altres, tenen un fenotip epitelial determinat per immunohistoquimica. Per tant, és possible que aquestes cèl·lules hagin inicialment sofert una EMT i després una MET per colonitzar el teixit o que les cèl·lules epitelials procedents del tumor primari que varen seguir el camí obert per aquelles més invasives són les que finalment estableixen la colonització i creixement del tumor secundari. Una altre possibilitat és que les dues coses ocorreixin al mateix temps.

Per això i pel fet que aquestes cèl·lules provenen de les mateixes línies cel·lulars originals, vàrem iniciar una sèrie d'experiments *in vitro* i *in vivo*. Co-cultivant aquestes cèl·lules en assajos d'invasivitat es demostra com el contacte entre PC-3/Mc (poc invasives) i PC-3/S (molt invasives), incrementa la capacitat metastàtica d'aquestes. Aquest increment podria ser per invasió passiva darrera les PC-3/S o per activació de mecanismes invasius. El fet que el medi condicionat del les PC-3/S sobre les PC-3/Mc ja és suficient per augmentar parcialment la invasió, per tant, dona a entendre que a part dels mecanismes invasius passius d'alguna forma hi ha una activació per augmentar la invasió. Això es confirma quan al mirar l'expressió gènica de les PC-3/Mc que han estat en contacte amb les PC-3/S, es veu com aquestes perden els gens propis del fenotip CSC i epitelial amb el contraposat guany d'alguns gens propis d'EMT. A més la formació d'esferoides també es veu disminuïda i en experiments *in vivo* es veu com la tumorigenicitat local és disminuïda, però la capacitat metastàtica es veu accelerada, cosa molt interessant. El fet de co-injectar PC-3/Mc amb PC-3/S, unes amb capacitats de créixer sense un substrat sòlid d'auto-renovació i les altres amb capacitat invasiva, demostra que tenen la capacitat de cooperar i complementar les seves capacitats per augmentar l'eficiència de la metàstasis. Així les cèl·lules invasives obren el camí a través dels teixits i són seguides per les PC-3/Mc que arribades al punt distant són les que tenen la capacitat de formar i iniciar nous tumor secundaris. A més d'aquesta sortida passiva de les PC-3/Mc també hi ha la sortida activa d'aquella part de població de PC-3/Mc que ha sofert una EMT transitòria al contactar amb les PC-3/S, tal com passa *in vitro*, precisament per això disminueix la tumorigenicitat local. Llavors aquestes, al no patir una EMT absoluta, un cop arribades al teixit distant podrien revertir fent una MET is sumat-se a les PC-3/Mc que havien sortit passivament. Els nostres resultats també demostren que les PC-3/Mc són les úniques que formen aquets tumor metastàtics,

ja que la presència de PC-3/S mai ha estat detectada. És interessant veure com les PC-3/Mc no necessiten les PC-3/S per colonitzar el moll de l'os, ja que els seus capil·lars són porosos, en canvi sí són necessàries per colonitzar els pulmons, ja que aquests no tenen els capil·lars porosos. A més, cal mencionar que la co-injecció intracardiaca produeix metàstasis a la glàndula adrenal cosa que mai passa quan les cèl·lules són injectades per separat.

Els resultats amb mostres humanes de pròstata mostren una important correlació de la seva agressivitat amb el fenotip de les PC-3/Mc. La signatura de gens que vàrem establir de les PC-3/Mc juntament amb una signatura d'ESCs, mostra com aquesta signatura és altament expressada en aquelles mostres de tumors primaris que finalment metastatitzen diferenciant-los d'aquells que no arriben a metastatitzar. Per tant, és una signatura que podria predir el comportament metastàtic. A més, els anàlisis d'immunohistoquímica mostren com aquelles mostres que derivaran en més agressives expressen més SOX2 i que algunes mostres de metàstasis el 100% de les cèl·lules expressen SOX2 conjuntament amb E-cadherina.

Tot aquest conjunt de resultats senyalen com el fenotip epitelial sol anar lligat al fenotip CSC i així confereix una gran potència metastàtica. Per altre banda, la inducció d'EMT pot causar la supressió de metàstasis en els nostres models estudiats. A més, aquests fenotips representats per diferents poblacions cel·lular dintre de l'heterogeneïtat tumoral poden cooperar per tal d'incrementar l'eficiència metastàtica. I que les poblacions altament invasives semblen tenir una influència sobre la resta induint part de la població tumoral cap a un estat més mesenquimal, equilibrant les poblacions tumoral i augmentant les possibilitats d'establir cooperacions entre capacitats invasives i capacitats colonitzadores. Els nostres resultats amb mostres clíniques mostren tenir una correlació bastant directa amb l'agressivitat i metàstasis, per tant, són fenotips i transformacions cel·lulars a tenir en compte alhora de dissenyar dianes terapèutiques.

L'article 2 es basa en un estudi sobre els mecanismes pels quals la família de microRNAs miR-200s incrementen la capacitat metastàtica. Es veu com l'expressió d'aquests està incrementada en mostres metastàtiques. El seu mecanisme de funcionament consisteix en reprimir els factors inductors d'EMT, ZEB1 i ZEB2, els quals són potents repressors directes de la E-cadherina. La sobreexpressió d'aquests miR-200s confereix a les cèl·lules 4TO7 un fenotip epitelial altament metastàtic tot i disminuir la seva capacitat invasiva. En aquest altre model cel·lular també s'observa aquesta dissociació entre la capacitat invasiva i la capacitat metastàtica. El model cel·lular és un model dual de càncer de mama de ratolí, on es comparen les 4T1 (epitelials i metastàtiques) amb les 4TO7 (mesenquimals i poc metastàtiques). Aquest model es suma doncs als models de pròstata i bufeta de l'article 1, on la invasió i la capacitat metastàtica no van lligades.

Els mir-200s indueixen el fenotip epitelial permetent l'expressió d'E-cadherina, mecanisme molt ben caracteritzat a través dels Zebs. Anàlisis transcriptòmics realitzats després de la sobreexpressió de miR-200s demostren que són capaços de generar molts canvis d'expressió gènica, induint molts gens que

afavoreixen el fenotip epitelial. Incrementen també l'expressió de gens com EpCAM of Msln, que són gens epitelials implicats en formació d'esferoides i autorrenovació, a través de mecanismes encara desconeguts. Després d'això i d'observar que el knock-down o expressió ectòpica de la E-cadherina no afecta de forma significativa la capacitat metastàtica, sembla que els miR-200s exerceixen les seves funcions metastàtiques a través de influir molt altres gens. Es interesant destacar que l'expressió de miR-200s tot i disminuir la capacitat de disseminació de les 4TO7 finalment augmenta molt la seva capacitat invasiva.

Intentant cercar possibles dianes dels miR-200s, a partir d'anàlisi genòmics i proteòmics, vàrem determinar el gen Sec23a com un diana directa dels miR-200s. Aquests al reprimir l'expressió del Sec23a, que és una proteïna implicada en la secreció de proteïnes des de l'aparell de golgi a l'exterior, es veu afectada la secreció de moltes altres proteïnes. Els anàlisi funcionals *in vitro* i *in vivo* mostren com la repressió del Sec23a augmenta la capacitat metastàtica de les cèl·lules i disminueix la seva invasió. A part Sec23a es trobava menys expressat en cèl·lules altament metastàtiques. Aquestes dades són recolzades per l'anàlisi de mostres humanes metastàtiques humanes, les quals expressen nivells inferiors de Sec23a que els tumors primaris.

Anàlisi d'espectrometria de masses ens permeteren identificar les proteïnes més afectades, és a dir, aquelles que es deixaven de secretar després de la repressió del Sec23a. Del llistat de proteïnes menys secretades, la baixa expressió d'algunes d'elles en mostres clíniques correlaciona amb mal pronòstic del pacient, en concret els factors Igfbp4 i Tinagl1. La seva caracterització funcional *in vitro* i *in vivo* ens va revelar que poden exercir funcions de supressors metastàtics, ja que el knock-down d'aquest augmentava considerablement la colonització pulmonar de les 4TO7.

Per tant, en aquest segon estudi es descriu com els miR-200s promouen un fenotip epitelial metastàtic al suprimir les propietats invasives i induir gens epitelials i reprimir directament el Sec23a que conseqüentment provoca un disminució dels supressors tumorals Igfpb4 i Tinagl1.

En resum, els dos articles junts comparteixen importants nous punts de vista del llarg procés de la metàstasis, però principalment centrant l'atenció en els últims passos de colonització de teixits distants. Aquests últims passos, arrel d'aquest estudis presentats, semblen ser més crítics alhora de metastatitzar que no els passos inicials on les capacitats invasives juguen un paper fonamental. Així suggereixen que el fenotip epitelial-CSC no secretor té una alta potència per provocar metàstasis en comparació amb fenotips invasius-mesenquimals. Això pot ser degut a que el fenotip EMT no resulta ser tant limitant com l'epitelial-CSC, ja que en el tumor primari les epitelials-CSC poden derivar fàcilment a EMT transitòries, tal i com es demostra en el primer article on les PC-3/S poden induir a les PC-3/Ms a patir una EMT. Així aquestes de fenotip invasiu poden ajudar a les epitelial-CSC a arribar, de forma passiva o activa, al lloc distant i són realment aquestes les que tenen capacitat de complir amb la difícil tasca de colonitzar i iniciar un nou tumor en un lloc a distància.

A més, si el fenotip epitelial-CSC resulta tenir una menor secreció de supressors metastàtics encara potencia més la capacitat metastàtica d'aquestes.

Conclusions

- Utilitzant models cel·lulars de pròstata, bufeta i mama hem demostrat una dicotomia funcional entre dos processos fonamentals per la metàstasis, invasió i auto-renovació, cada una associada a diferents poblacions tumorals.

- En els nostres models, fenotips epitelials van lligats a propietats CSC-TIC, sent molt més eficients en creixement tumoral i metàstasis que els tumors cel·lulars mesenquimals.

- La transició epiteli-mesènquima i el programa d'auto-renovació poden ser mútuament exclusius en determinats tipus de cèl·lules tumorals. Llavors la inducció d'EMT en cèl·lules epitelials de pròstata i bufeta pot suprimir el seu programa d'auto-renovació i les seves capacitats de tumorigenicitat i metàstasi.

- En el nostre model de càncer de pròstata, demostrem que interaccions entre subpoblacions tumorals tenen conseqüències significants per l'eficiència de la metàstasis. Per tant, cèl·lules de tipus mesenquimal indueixen una EMT en cèl·lules epitelials iniciadores de tumors o CSC, augmentant així la seva capacitat invasiva de forma transitòria i reduint les propietats d'auto-renovació in vitro, i augmentant la seva sortida del lloc de formació tumoral amb el conseqüent augment d'acceleració de la metàstasis.

- L'expressió de SOX2 està significativament associada a agressivitat en càncer de pròstata.

- L'expressió de la família miR-200 de microRNAs està significativament associada amb metàstasis del càncer de mama.

- MiR-200s indueixen un fenotip epitelial-no secretor altament metastàtic en cèl·lules de càncer de mama, a través de la directa repressió dels Zebs i del Sec23a.

- Igfbp4 i Tinagl1 són identificats com a supressors de la metàstasis, regulats pel Sec23a i així pels miR-200s.

- Els dos estudis suporten la idea que les cèl·lules tumorals requereixen d'un fenotip epitelial actiu per finalment poder metastatitzar.

4. RESUM ARTICLE 1: La transició epiteli-mesènquima pot suprimir propietats rellevants de cèl·lules epitelials iniciadores de tumor.

La progressió maligna en càncer requereix de poblacions iniciadores de tumors (TICs) que tenen capacitat auto-renovadora il·limitada, supervivència baix estrès i d'establir metàstasis a distància. Addicionalment, l'adquisició de propietats invasives per transició epiteli-mesènquima (EMT) és crítica per l'evolució neoplàsica en la conversió a poblacions metastàtiques. Aquí caracteritzem models cel·lulars derivats de línies de càncer de pròstata i bufeta, en les quals unes subpoblacions expressen un programa epitelial enriquit en cèl·lules metastàtiques iniciadores de tumors (TICs), mentre que una segona població presenta característiques mesenquimals i és pobre en cèl·lules iniciadores de tumors (TICs). La sobreexpressió constitutiva del factor de transcripció *Snai1* en les poblacions enriquides en cèl.lules epitelials-iniciadores de tumors, indueix el fenotip mesenquimal i suprimeix l'auto-renovador i metastàtic. D'altra banda, el knock-down de factors d'EMT en la subpoblació mesenquimal va provocar el guany de propietats epitelials i de cèl·lules iniciadores de tumors. Aquestes dues subpoblacions de càncer de pròstata poden cooperar entre elles. La subpoblació mesenquimal augmenta la seva invasió *in vitro* de les epitelials-TICs i *in vivo* promou la seva sortida del lloc de tumor primari accelerant així la colonització metastàtica. Els nostres models proveeixen nous coneixements en com la dinàmica entre programes epitelials auto-renovadors i mesenquimals, determinen la plasticitat de les cèl·lules epitelials iniciadores de tumors.

5. RESUM ARTICLE 2: Sec23a és diana directe dels miR-200s i així influencien el secretoma de les cèl·lules canceroses i promou colonització metastàtica

Tot i que el paper dels miR-200s en la regulació de l'expressió de la E-cadherina i la transició epiteli-mesènquima està molt ben establerta, la seva influència en colonització metastàtica encara genera controvèrsia. Aquí utilitzem models clínics i experimentals de metàstasis de càncer de mama per descobrir el paper pro-metastàtic dels miR-200s més enllà de la seva regulació de l'E-cadherina i el fenotip epitelial. La sobreexpressió dels miR-200s està associada a un elevat risc de metàstasis en càncer de mama i promou colonització metastàtica en models de ratolí, fenotip que no és recapitulat per la E-cadherina sobreexpressada sola. Anàlisis genòmics i proteòmics revelen canvis globals en expressió gènica al sobreexpressar els miR-200s i la conseqüent obtenció de cèl·lules altament metastàtiques. Els miR-200s promouen la colonització metastàtica parcialment a partir de la repressió directa del Sec23a, el qual mediatitza la secreció de proteïnes supressores de la metàstasis, incloent Igfbp4 i Tinagl1, aquí validades funcionalment i en estudis clínics. En general, aquests descobriments suggereixen un paper pleiotròpic dels miR-200s en promoure la colonització metastàtica influenciant propietats epitelials dependents de l'E-cadherina i el secretoma de cèl·lules tumorals que mediatitza el Sec23a.