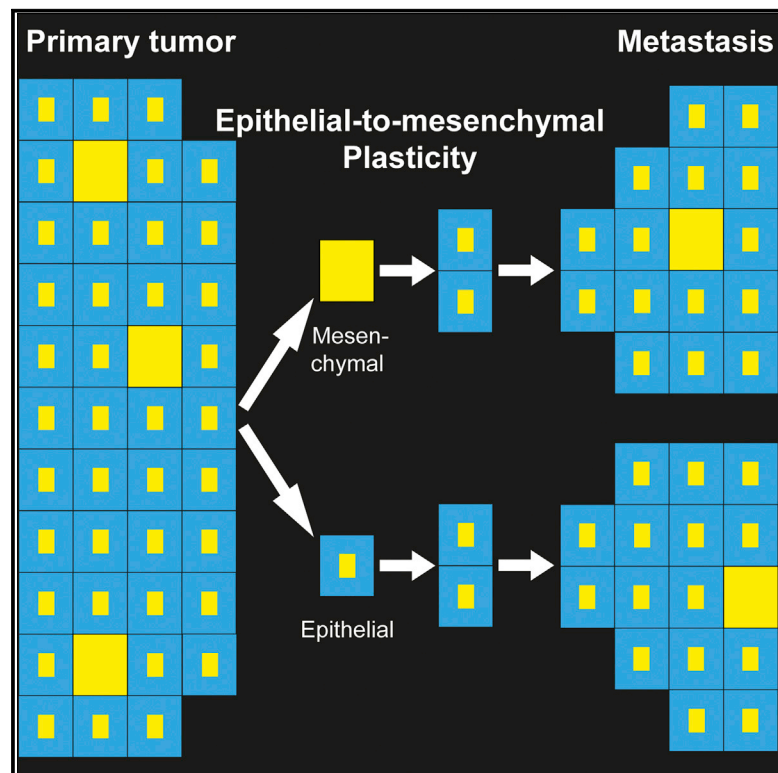


Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity

Graphical Abstract



Authors

Evelyne Beerling, Daniëlle Seinstra, Elzo de Wit, ..., Alexander van Oudenaarden, Nienke Vrisekoop, Jacco van Rheenen

Correspondence

n.vrisekoop@umcutrecht.nl (N.V.), j.vanrheenen@hubrecht.eu (J.v.R.)

In Brief

Beerling et al. identified a previously undetectable pool of cells in epithelial breast tumors that have undergone EMT without experimental induction. These cells are motile when disseminating and revert back to the epithelial state upon metastatic outgrowth. This epithelial-mesenchymal plasticity equalizes metastatic outgrowth potential between epithelial and mesenchymal tumor cells.

Highlights

- Direct evidence of EMT obtained in unperturbed breast tumors by real-time visualization
- EMT exists in breast tumors without experimentally altering EMT inducers
- Tumor cells that underwent EMT are the migratory cells within a tumor
- Outgrowth potential differences between states are irrelevant due to plasticity

Accession Numbers

GSE77107
PRJEB5939



Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity

Evelyne Beerling,^{1,4} Daniëlle Seinstra,^{1,4} Elzo de Wit,^{1,3} Lennart Kester,¹ Daphne van der Velden,³ Carrie Maynard,¹ Ronny Schäfer,¹ Paul van Diest,² Emile Voest,³ Alexander van Oudenaarden,¹ Nienke Vrisekoop,^{1,5,*} and Jacco van Rheenen^{1,*}

¹Cancer Genomics Center-Hubrecht Institute-KNAW & University Medical Centre Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands

²Department of Pathology, University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

³Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands

⁴Co-first author

⁵Present address: Laboratory of Translational Immunology, Department of Pulmonary Diseases, University Medical Centre Utrecht, Lundlaan 6, 3584 EA Utrecht, the Netherlands

*Correspondence: n.vrisekoop@umcutrecht.nl (N.V.), j.vanrheenen@hubrecht.eu (J.v.R.)

<http://dx.doi.org/10.1016/j.celrep.2016.02.034>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

SUMMARY

Forced overexpression and/or downregulation of proteins regulating epithelial-to-mesenchymal transition (EMT) has been reported to alter metastasis by changing migration and stem cell capacity of tumor cells. However, these manipulations artificially keep cells in fixed states, while in vivo cells may adapt transient and reversible states. Here, we have tested the existence and role of epithelial-mesenchymal plasticity in metastasis of mammary tumors without artificially modifying EMT regulators. In these tumors, we found by intravital microscopy that the motile tumor cells have undergone EMT, while their epithelial counterparts were not migratory. Moreover, we found that epithelial-mesenchymal plasticity renders any EMT-induced stemness differences, as reported previously, irrelevant for metastatic outgrowth, because mesenchymal cells that arrive at secondary sites convert to the epithelial state within one or two divisions, thereby obtaining the same stem cell potential as their arrived epithelial counterparts. We conclude that epithelial-mesenchymal plasticity supports migration but additionally eliminates stemness-enhanced metastatic outgrowth differences.

INTRODUCTION

Metastatic growth is the major cause of cancer-associated mortality. To successfully grow metastases, epithelial tumor cells need to acquire invasive properties to disseminate and stem cell properties to grow new tumors at distant sites (Hanahan and Weinberg, 2011). Metastasizing cancer cells have been suggested to hijack a developmental program named epithelial-to-

mesenchymal transition (EMT) (Bill and Christofori, 2015; Kalluri and Weinberg, 2009; Lim and Thiery, 2012). During developmental EMT, cells lose cell-cell contacts and concomitantly decrease the expression of the epithelial adherens junction molecule E-cadherin (E-cad) and gain expression of proteins involved in, e.g., invasion and stemness (Kalluri and Weinberg, 2009; Lim and Thiery, 2012; Thiery and Sleeman, 2006).

The effect of EMT on stemness, as well as the role and even the very existence of EMT during metastasis, are heavily debated (Del Pozo Martin et al., 2015; Fischer et al., 2015; Zheng et al., 2015). For example, contradicting findings were published on the stem cell potential of tumor cells with an epithelial or mesenchymal state. Some studies found that EMT-inducing transcription factors, such as Twist, coincide with the acquisition of stem cell properties, thereby supporting metastatic growth (Mani et al., 2008; Morel et al., 2008; Wellner et al., 2009; Yang et al., 2004). Other studies found that a forced reversion to an epithelial state through Twist knockdown leads to metastasis-initiating abilities (Ocaña et al., 2012; Tsai et al., 2012). Importantly, both experimental approaches may not represent the true in vivo status of cells because they require gene manipulations that artificially force cells into fixed states, while in vivo cells may be able to transiently and reversibly switch between states, a process that from here on is referred to as epithelial-mesenchymal plasticity. Moreover, the non-physiological overexpression or complete loss of EMT-regulators, such as Twist1, may induce expression profiles and subsequently stem cell phenotypes that do not exist under physiological conditions. Finally, EMT regulators can have oncogenic functions independently of their ability to induce EMT, thus observed phenotypes that result from gene manipulation may not be exclusively due to EMT induction (Beck et al., 2015). These data and concerns illustrate the importance of studying EMT in non-manipulated in vivo settings.

Although EMT would best be studied in the physiological in vivo settings, non-experimentally induced EMT during metastasis has yet to be observed. For example, extensive histological examination of human invasive ductal mammary carcinomas

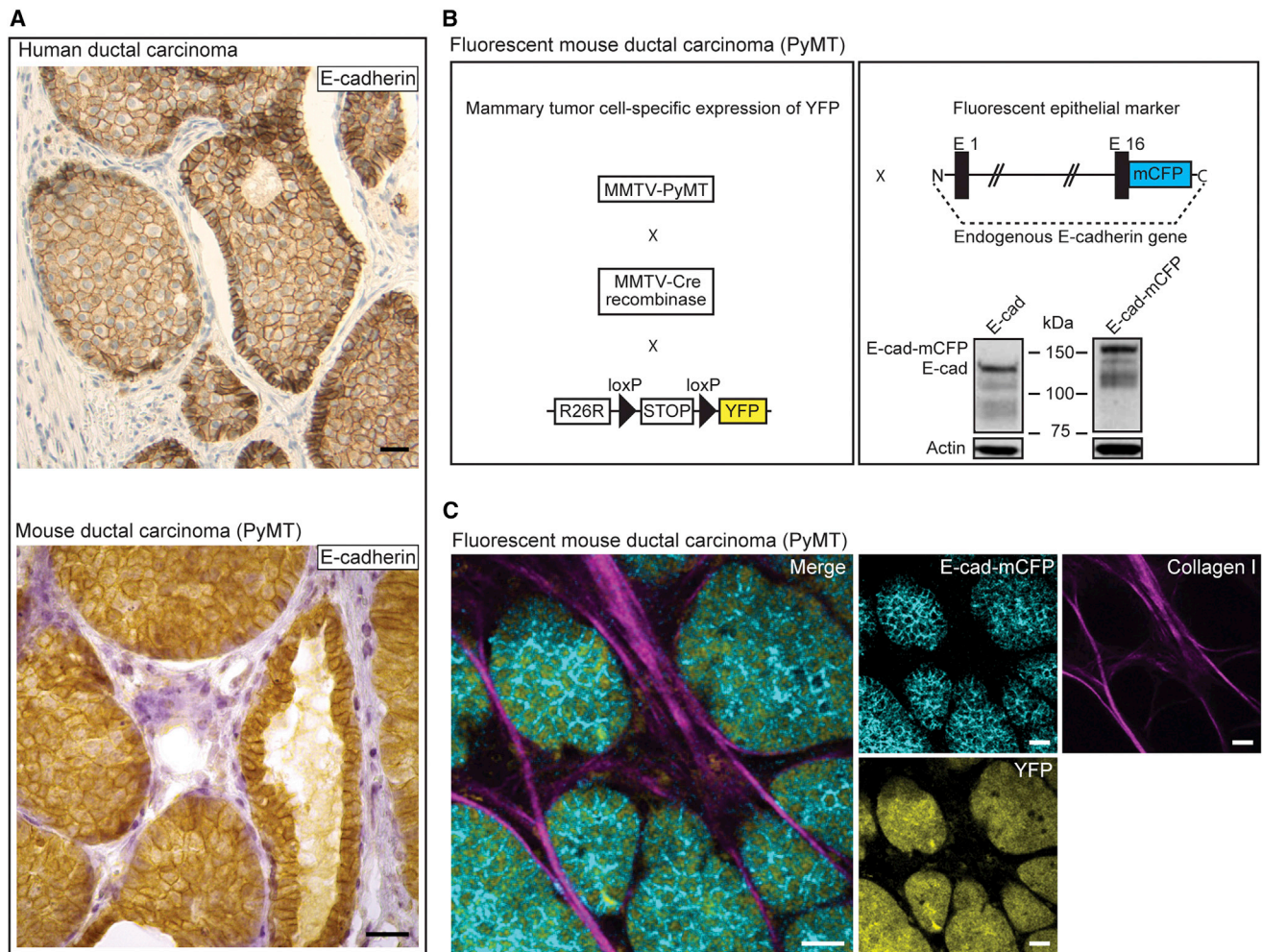


Figure 1. Development of a Fluorescent Mouse Model for Metastatic E-cad-Positive Invasive Ductal Carcinomas

(A) Human invasive ductal carcinoma (upper) and a late-stage MMTV-PyMT tumor (lower), stained for E-cad and counterstained with H&E. Scale bar, 30 μ m. (B) Schematic representation of the fluorescent mouse model in which all tumor cells express YFP and in which the endogenous E-cad is labeled with CFP. The western blot shows wild-type and CFP-tagged E-cad. (C) Multi-photon images of fluorescent PyMT mammary tumors. Scale bars, 30 μ m.

shows that, even in tumors that have metastasized, tumor cells in the primary tumor, as well as the metastases, display an epithelial phenotype (e.g., Bukholm et al., 2000; Jeschke et al., 2007; Kowalski et al., 2003). This means that either EMT does not exist when it is not experimentally induced or EMT remains undetected in these static images because only a small population of cells temporarily adapts a mesenchymal state. Therefore, in addition to the development of models in which EMT can occur without modifying EMT-regulators, techniques are required that are able to reveal and study the potentially rare and undetectable pool of cells going through EMT.

Here, we combine high-resolution intravital imaging, single-cell sequencing, and transplantation techniques to investigate the role of EMT and epithelial-mesenchymal plasticity in metastasis of invasive ductal carcinomas. Our data suggest that epithelial-mesenchymal plasticity supports tumor cell migration and causes metastasis-enhancing stem cell capacity

differences between epithelial and mesenchymal states to be irrelevant.

RESULTS AND DISCUSSION

To determine whether EMT occurs without artificial induction, we used polyomavirus middle T antigen (PyMT) mice that develop ductal mammary carcinomas that recapitulate the progression of human mammary adenoma to late carcinoma stages and metastasize primarily to lymph nodes, lungs, and, occasionally, liver (Guy et al., 1992; Lin et al., 2003; Welm et al., 2007). Similar to human ductal carcinomas, these mammary tumors highly express E-cad, even in the late carcinoma stage and metastases (Figure 1A). To visualize EMT in vivo, we crossed these MMTV-PyMT mice with MMTV-Cre and R26R-loxP-stop-loxP-YFP (R26R-YFP) mice (Srinivas et al., 2001) to specifically label all tumor cells with yellow fluorescent protein (YFP). Next, we crossed

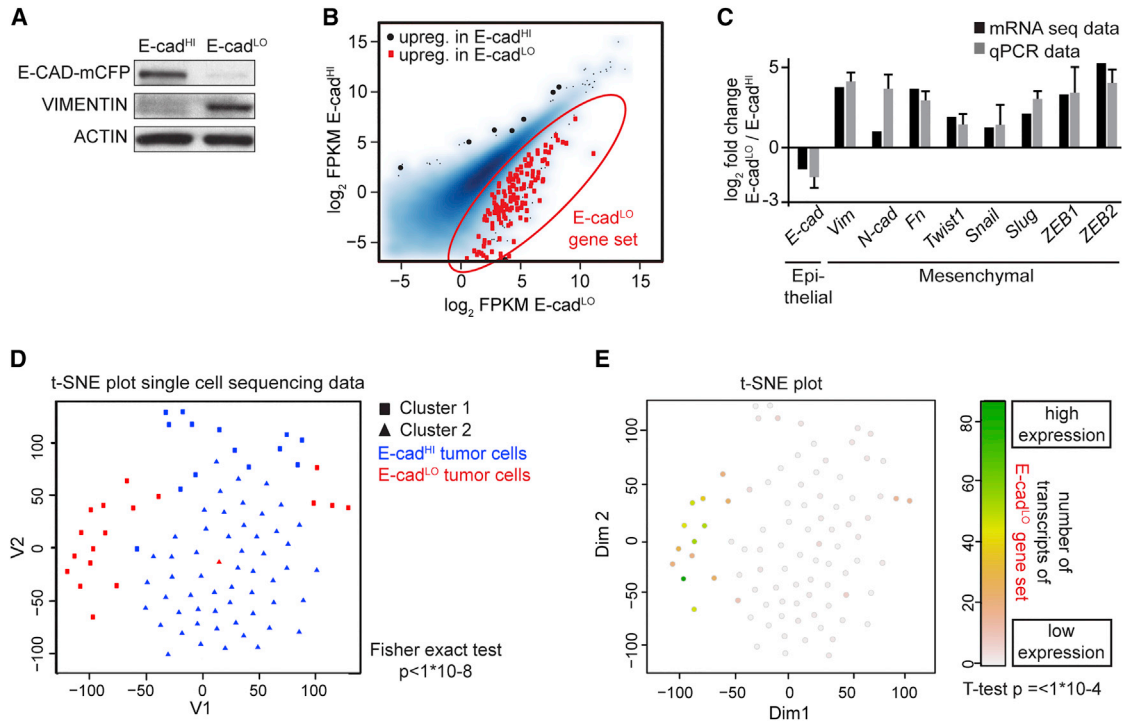


Figure 2. Rare E-cad^{LO} Cells Isolated from Mouse Invasive Ductal Carcinomas Have Undergone EMT

(A) Western blot of indicated samples. n = 3 mice.
 (B) Scatterplot showing relative expression values for E-cad^{HI} and E-cad^{LO} cells. Red dots that are encircled in red represent genes that are significantly upregulated in E-cad^{LO} cells (q value < 0.01).
 (C) The relative mRNA expression of EMT-related genes determined by RNA sequencing (RNA-seq) and RT. n = 4 mice, except for ZEB1, where n = 3 mice.
 (D) T-distributed stochastic neighbor embedding (t-SNE) plot. Using unsupervised K-medoids clustering, two separate clusters were identified indicated as squares and triangles that overlap with E-cad^{HI} (blue) and E-cad^{LO} (red) tumor cells. Fisher exact test p < 1*10⁻⁸.
 (E) t-SNE intensity plot for genes differentially upregulated in (B). T-test p < 1*10⁻⁴.
 Related to [Figures S1, S2, and S3](#) and [Table S1](#).

these animals with *E-cad-mCFP* mice in which a monomeric cyan fluorescent protein (mCFP) is fused to endogenous E-cad ([Snippert et al., 2010](#)), in order to label all endogenous E-cad with mCFP ([Figure 1B](#)). The resultant *MMTV-PyMT; MMTV-Cre; R26R-YFP; E-cad-mCFP* animals develop ductal mammary tumors and metastases in which all tumor cells are YFP-labeled and endogenous E-cad is tagged with mCFP ([Figure 1C](#)). Microscopic inspection ([Figure 1C](#)) and flow cytometry ([Figure S1](#)) of these fluorescent tumors showed that the vast majority of cells appear to have high levels of membrane-localized E-cad (E-cad^{HI} cells).

To test whether these tumors also contain a population of tumor cells that have undergone EMT in which E-cad is not functional by either downregulation of the expression or by decreasing membrane-localized E-cad, we dissociated fluorescent PyMT tumors and exposed the extracellular domain of E-cad to a fluorescently labeled antibody. We sorted YFP-expressing tumor cells, excluding, for example, lymphocytes (see [Figure S1](#)). In contrast to analysis of histological images, careful analysis of the flow cytometry data showed that in addition to the population of E-cad^{HI} tumor cells another, much smaller population of tumor cells could be found. In this population the expression of E-cad-mCFP was low and/or E-cad was non-functional

due to intracellular localization as determined by low extracellular antibody staining (E-cad^{LO}; [Figure S1B](#)). Western blot analysis confirmed that E-cad^{LO} cells have low levels of E-CAD and a concomitant upregulation of VIMENTIN ([Figure 2A](#)), which is consistent with mesenchymal characteristics ([Kalluri and Weinberg, 2009](#)). Moreover, using mRNA deep sequencing, we observed differential expression in the E-cad^{HI} and E-cad^{LO} cells of typical EMT genes, such as *Vimentin*, *Fibronectin*, and *N-cadherin*, and transcription factors that regulate EMT, including *Snail*, *Slug*, *Twist*, *ZEB1*, and *ZEB2*, referred to as the E-cad^{LO} gene set ([Figures 2B and 2C](#); [Table S1](#)). These results were confirmed by qPCR ([Figures 2C and S2A](#)).

These data show that in our system E-cad status can be used to distinguish between epithelial and mesenchymal phenotypes on the population level. To test whether this holds true at the single-cell level, we performed single-cell mRNA sequencing of 72 E-cad^{HI} and 25 E-cad^{LO} cells. When performing unsupervised K-medoids clustering of the individual expression profiles, two separate clusters were identified that overlapped with the E-cad^{HI} and E-cad^{LO} cells ([Figure 2D](#); p < 1⁻⁸). The single E-cad^{LO} cells had higher expression of the E-cad^{LO} gene set from the bulk sequencing data, confirming the mesenchymal identity of the E-cad^{LO} cells on the single-cell level ([Figure 2E](#);

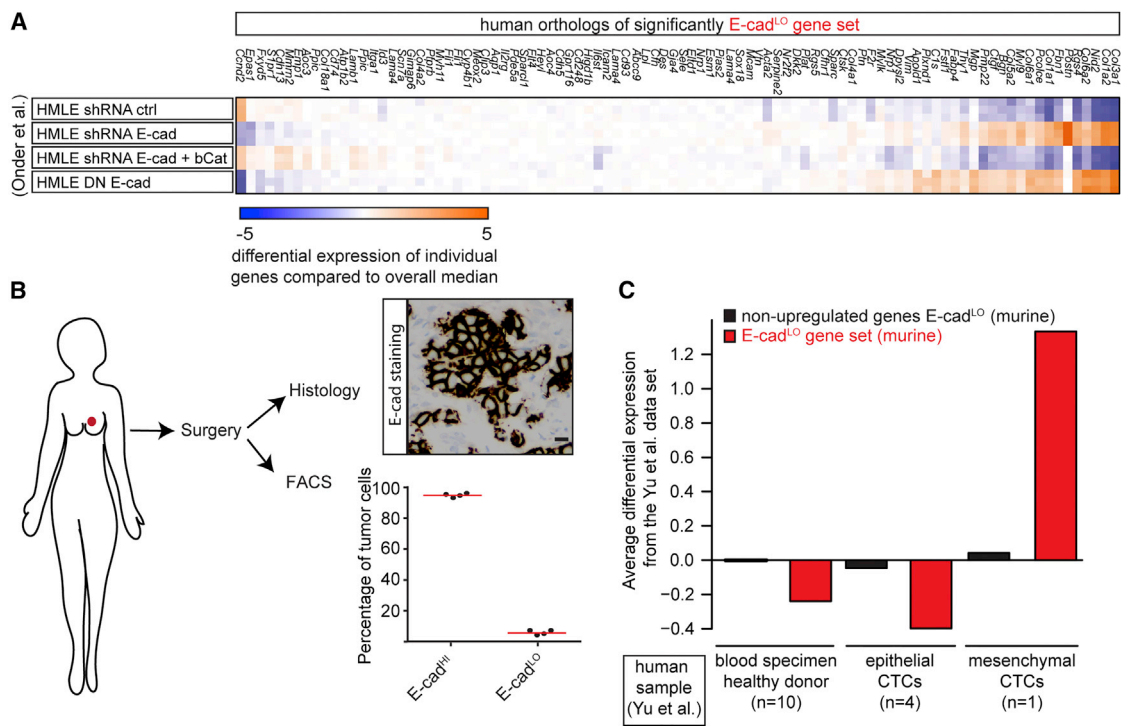


Figure 3. E-cad^{LO} Cells Are Similar to Human Mesenchymal Tumor Cells

(A) From a published dataset (Onder et al., 2008), the expression level in human mammary epithelial cells (HMLE) of the human orthologs of the mouse E-cad^{LO} gene set in Figure 2B was retrieved. Differential expression levels per gene as the deviation of the median across all experiments are shown.

(B) The image shows an E-cad⁺ human invasive ductal carcinoma, and the graph shows the percentage of E-cad^{HI} and E-cad^{LO} cells (n = 4 tumors). Scale bar, 20 μ m.

(C) From published RNA-seq experiments of human breast cancer CTCs (Yu et al., 2013), the relative expression levels of human-mouse orthologs were retrieved. Plots show the average differential expression found in the Yu et al. (2013) dataset for the E-cad^{LO}-upregulated (red bars) or non-upregulated genes (black bars). Expression levels for circulating cells were determined for blood draws from ten healthy donors (left two bars), four blood draws from one patient with CTCs with an epithelial phenotype (middle two bars), and one of the same patient with CTCs with a mesenchymal phenotype (right two bars).

t test $p < 1^{-4}$; Figures S2B and S2C). Furthermore, E-cad^{LO} and E-cad^{HI} cells clustered separately in a heatmap (Figure S2D), and we observed that the expression profiles of E-cad^{HI} cells were more similar to each other (Pearson correlation of 0.45) than to the profiles of E-cad^{LO} cells (Pearson correlation of 0.35, Wilcoxon two-sided test = $p < 10^{-15}$; Figures 2D and S2D). Collectively, these data show that in PyMT tumors at the single-cell level that E-cad status can be used to discriminate between cells with epithelial and mesenchymal features.

Next, we tested whether human tumors also contain E-cad^{LO} cells. As a first indication, we found that the E-cad^{LO} gene set (marked with red dots in Figure 2B) is upregulated in human cells in which EMT is induced upon E-cad knockdown or expression of dominant-negative E-cad (Onder et al., 2008) (Figure 3A, $p < 0.0001$ hypergeometric test). Next, we obtained tumors directly after patients underwent mastectomy and selected four tumors that stained positive for E-cad on histological sections (Figure 3B). We dissociated the tumors into single cells, stained the cells with DAPI to exclude apoptotic cells, EpCAM to select for tumor cells (Yu et al., 2013), and E-cad to distinguish between E-cad^{HI} and E-cad^{LO} cells. By flow cytometry we indeed detected both E-cad^{HI} and E-cad^{LO} cells (Figure 3B). To test whether human E-cad^{LO} and mouse E-cad^{LO} cells are similar,

we used a recently published dataset of gene expression in epithelial- and mesenchymal-circulating tumor cells (CTCs) from breast cancer patients (Yu et al., 2013). Importantly, the mouse E-cad^{LO} gene set was also upregulated in the human mesenchymal CTCs, but not in the human epithelial CTCs or in healthy blood specimens (Figure 3C). Combined, these results show that we have identified a subpopulation of mouse tumor cells (E-cad^{LO}) that is similar to that of human mesenchymal CTCs.

Since tumors are genetically very heterogeneous, the E-cad^{LO} CTCs from breast cancer patients and from our mouse model may either adapt a permanent mesenchymal state by, e.g., mutations in EMT-regulators, or represent a transient reversible mesenchymal state. To test whether the mesenchymal state is reversible, we first generated organoids from the *MMTV-PyMT*; *MMTV-Cre*; *R26R-YFP*; *E-cad-mCFP* carcinomas. We stimulated these organoids with transforming growth factor beta (TGF-beta) and hepatocyte growth factor (HGF) and indeed observed an increase in the number of E-cad^{LO} cells (Figure S3A), showing that the mesenchymal state of E-cad^{LO} cells can be stimulated. This state can also be lost, since orthotopic transplantation of E-cad^{LO} cells always resulted in mammary tumors containing tumor cells with a predominantly epithelial

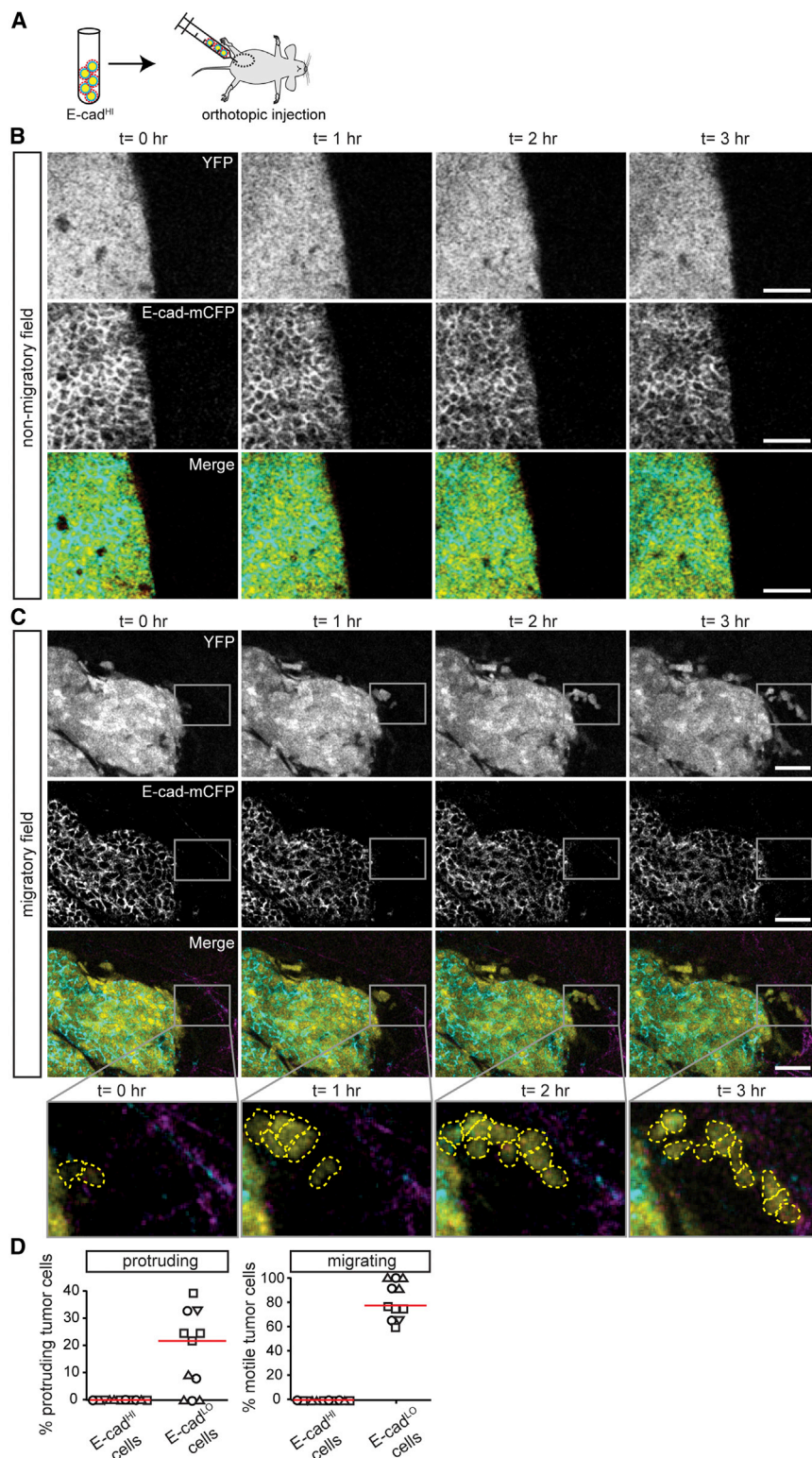


Figure 4. Behavioral Characterization of Rare E-cad^{LO} Tumor Cells in Mouse Mammary Carcinomas that Highly Express E-cad

(A) Cartoon of the experimental setup. (B and C) Intravital images of PyMT tumors containing non-motile (B) and migratory (C) tumor cells. The rectangular box highlights migrating E-cad^{LO} cells. Scale bars, 50 μ m. (D) The percentage of protruding (left) and motile cells (right) plotted against E-cad status. Red lines indicate the median. The graph represents data from imaging fields with moving cells (11 out of 45 imaging fields from four mice, where symbols represent different mice). Related to [Figure S4](#) and [Movies S1](#) and [S2](#).

CellTracker dye to visualize cell division, reverted to an epithelial state both before ([Figure S3C](#), upper) and after cell division ([Figure S3C](#), lower). Combined, these results show that the mesenchymal state of E-cad^{LO} cells is plastic and can be gained and lost by tumor cells.

Similar to previously identified invasive signatures of tumor cells ([Wang et al., 2004, 2007](#)), many categories of E-cad^{LO} gene set were related to tissue development, morphogenesis, migration, and adhesion ([Figures S3D–S3F](#)). This result prompted us to visualize the behavior of these E-cad^H and E-cad^{LO} tumor cells in vivo using multi-photon microscopy. To exclude YFP expression in non-epithelial lineages, tumors were imaged that developed upon transplantation of E-cad^H tumor cells into the mammary glands of wild-type mice ([Figure 4A](#)). In addition to endogenous mCFP-labeled E-cad and YFP, we visualized type I collagen by imaging the second harmonic generation signal. As reported previously ([Wyckoff et al., 2007](#)), we found that the migratory behavior of tumor cells in these genetic PyMT tumors is very heterogeneous: while no migratory cells were found in the majority of imaging fields ([Figure 4B](#); [Movie S1](#)), we found many migratory cells in some imaging fields ([Figure 4C](#); [Movie S2](#); in [Figure S4A](#) we demonstrate that cell motility is not due to Z-drift of the focal plane). The tumor cells migrated either individually or as streams in which single cells appeared to follow each other's migration path ([Figure 4C](#)), as has been demonstrated

before in other tumor models ([Patsialou et al., 2013](#); [Roussos et al., 2011](#)), but collective migration of cohesive epithelial clusters was not observed in this model. While on average E-cad^H

phenotype as indicated by E-cad expression ([Figure S3B](#)). This reversibility is not necessarily dependent on cell division, since sorted E-cad^{LO} cells plated into a 3D matrix and stained with a

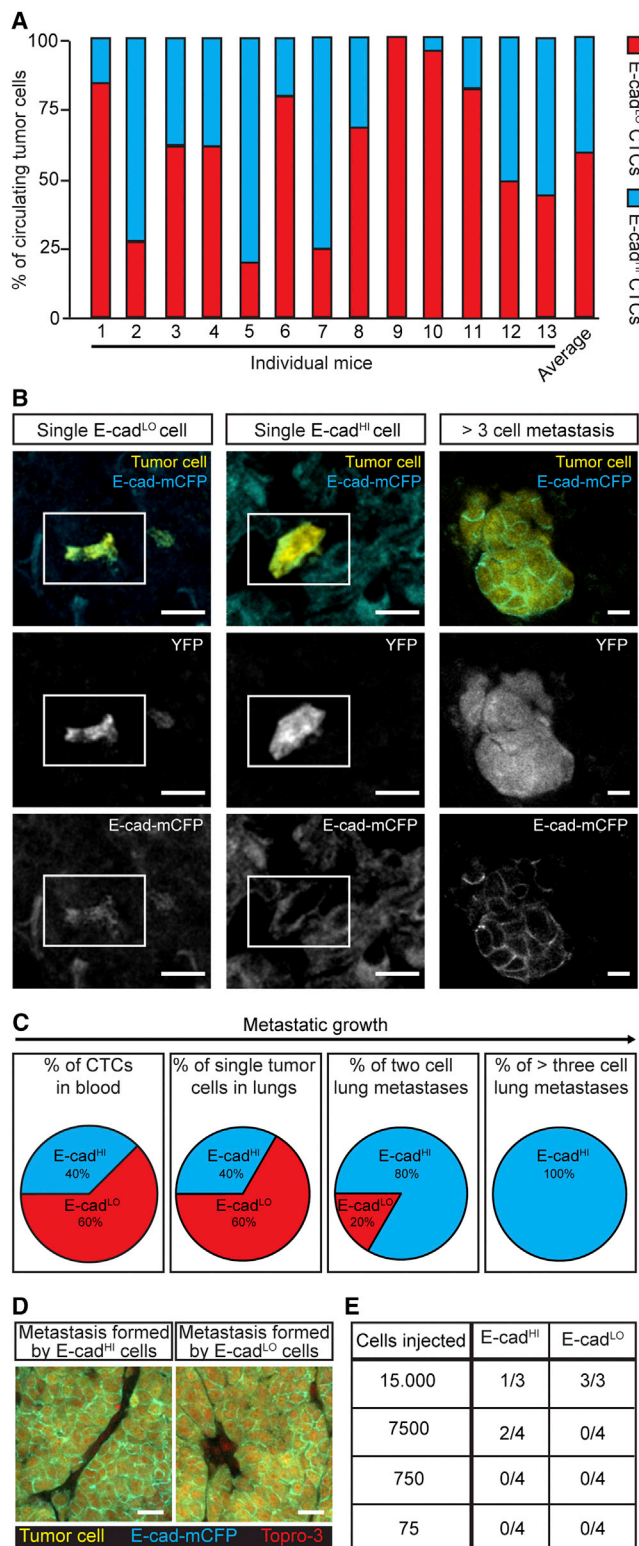


Figure 5. Epithelial-Mesenchymal Plasticity Renders Potential Stem Cell Differences Irrelevant for Metastatic Outgrowth

(A) The percentage of E-cad^{LO}- and E-cad^{HI}-circulating tumor cells. n = 13 mice.

cells were non-motile, the rare E-cad^{LO} cells were either protruding or displaying migratory behavior (Figure 4D). These migratory E-cad^{LO} cells do not relate to the CK14-positive cells (Figures S4B and S4C) that lead collective migration in organoids (Cheung et al., 2013). Collectively, our data show that E-cad^{LO} cells represent a rare population of motile cells that have undergone spontaneous EMT without experimental induction, within otherwise non-motile epithelial tumors.

Since we found that the small population of migratory cells in tumors has undergone EMT, we questioned whether these tumor cells enter the circulation in this mesenchymal state only or whether they can revert and/or enter in an epithelial state. Despite a large variation between mice regarding the total number of CTCs and the percentage of E-cad^{HI} and E-cad^{LO} CTCs, both types of CTCs were present in the blood of tumor-bearing MMTV-PyMT mice (Figure 5A). Next, we tested whether the E-cad^{LO}-circulating tumor cells show the same expression profile as the E-cad^{LO} tumor cells in the primary tumor. We performed single-cell sequencing and observed that the circulating E-cad^{LO} and E-cad^{HI} cells cluster into two populations and that the circulating E-cad^{LO} cells show the same expression profile as the primary E-cad^{LO} tumor cells (Figures S5A and S5B). Since in this mouse model tumor cells do not stay long enough in the circulation to switch to another state (99.99% of IV-injected tumor cells get cleared from circulation within 30 s [Figure S5C]), we can conclude from our data that disseminating tumor cells that enter the circulation are in a mesenchymal, but also an epithelial, state.

Next, we investigated the epithelial and mesenchymal state of endogenous spontaneous metastases to the lung. In line with the percentages of E-cad^{HI} and E-cad^{LO} cells in the blood, 40% of single metastasized tumor cells appeared to be E-cad^{HI} cells, and 60% were E-cad^{LO} cells (Figures 5B and 5C). In contrast to findings in prior studies in which EMT was induced (Stoletov et al., 2010), our data suggest that naturally occurring EMT does not influence the arrival and extravasation of the CTCs at the site of metastatic outgrowth. To investigate the cells that grow out to metastases, we examined endogenous metastases with a size of two and more than three cells. Although 20% of the two-cell micrometastases were E-cad negative, all metastases larger than three cells were E-cad positive (Figures 5C and 5D).

Since our histological analysis shows that all metastases larger than three cells contain E-cad^{HI} cells, we hypothesized that either only E-cad^{HI} cells are able to grow metastases or E-cad^{LO} cells convert to an epithelial state during the first cell divisions. Interestingly, E-cad^{HI} and E-cad^{LO} cells do not differ in their proliferative capacity (Figures S5D and S5E). To further

(B) Representative images of single E-cad^{LO} and E-cad^{HI} cells and a multi-cellular metastasis in the lung. White rectangle highlights single cells. Scale bars, 20 μm.

(C) Percentages of E-cad^{LO} and E-cad^{HI} tumor cells in blood and lungs. Blood: n = 13 mice; lungs: n = 143 metastases in 16 mice.

(D) Representative images of liver metastases grown from E-cad^{HI} cells (left) and E-cad^{LO} cells (right). Scale bars, 40 μm.

(E) Table indicating the metastatic outgrowth potential of E-cad^{LO} and E-cad^{HI} cells. Tumor-initiating cell frequency as tested by the Elda-limiting dilution test: E-cad^{HI} cells 1/21,228; E-cad^{LO} cells 1/17,545, p = 0.82.

Related to Figure S5.

test the impact of each state, and especially epithelial-mesenchymal plasticity, on metastatic potential, we investigated the ability of E-cad^{HI} and E-cad^{LO} cells to initiate liver metastases. In contrast to previous reports with fixed states (Fantozzi et al., 2014; Ocaña et al., 2012; Shibue and Weinberg, 2009, 2011; Tsai et al., 2012), the potential to grow metastases from plastic E-cad^{HI} cells and E-cad^{LO} cells is approximately equal (Figure 5E). The outgrowth of epithelial metastases from the E-cad^{LO} cells strongly suggest that at least a significant fraction of the spontaneous metastases as found in Figure 5B are grown from mesenchymal E-cad^{LO} cells that have converted to an epithelial state during the first few cell divisions. Considering all these data together, we conclude that, although intrinsically epithelial and mesenchymal cells may differ in their stem cell potential, this difference does not provide a large metastatic outgrowth advantage as mesenchymal cells adapt an epithelial state after the first few cell divisions, thereby abolishing any potential initial differences in stem cell properties.

Collectively, our data provide evidence for the existence of EMT in vivo without experimentally altering EMT-inducers. Artificial interference of EMT regulators does not reflect the moderate fluctuations of expression levels that occurs under physiological conditions and therefore is likely to lead to more extreme phenotypes. Moreover, these manipulations artificially keep cells in fixed states, whereas we here show that cells adapt transient and reversible states. Our data support the notion that temporal acquisition of the mesenchymal state is important for migration, but not for entering the circulation. We observed that mesenchymal cells that arrive at the secondary site adapt an epithelial state after a few cell divisions. These cells therefore acquire the same stemness properties as their epithelial counterparts. Thus, due to epithelial-mesenchymal plasticity, any differences in stemness between epithelial and mesenchymal states will be lost and become irrelevant for metastatic outgrowth. In conclusion, we have demonstrated plasticity between epithelial and mesenchymal states, thereby ruling out a critical role for differential stemness capacities and ultimately the potential to grow metastases.

EXPERIMENTAL PROCEDURES

Mice

All experiments were carried out in accordance with the guidelines of the Animal Welfare Committee of the Royal Netherlands Academy of Arts and Sciences, the Netherlands. For more details, see the [Supplemental Experimental Procedures](#).

Human Material

Human tissues were obtained in compliance with Dutch law that does not require informed consent when leftover materials are used anonymously.

Flow Cytometry on Mouse Material

After putting cells through a 70- μ m strainer cap (BD Falcon), cells were sorted on a fluorescence-activated cell sorting Ariall special-ordered research product (BD Biosciences). The sort strategy is illustrated in [Figure S1B](#). For more details, see the [Supplemental Experimental Procedures](#).

Intravital Imaging

Imaging was performed on an inverted Leica TCS SP5 AOBS multi-photon microscope with a chameleon Ti:Sapphire pumped Optical Parametric Oscillator (Coherent). For more details, see the [Supplemental Experimental Procedures](#).

ACCESSION NUMBERS

The accession number for the mRNA sequencing data reported in this paper has been uploaded to European Nucleotide Archive: PRJEB5939. The accession number for the single-cell mRNA sequencing data reported in this paper has been uploaded to GEO: GSE77107.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, one table, and two movies and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2016.02.034>.

AUTHOR CONTRIBUTIONS

E.B., D.S., N.V., and J.v.R. designed the experiments. E.B., D.S., C.M., R.S., and P.v.D. performed the experiments. E.B., D.S., and N.V. analyzed the data. E.d.W. analyzed the mRNA deep-sequencing data. L.K. performed single-cell sequencing experiments. A.v.O. and L.K. analyzed the single-cell sequencing experiments. D.v.d.V., P.v.D., and E.V. helped to obtain patient material. E.B., D.S., E.d.W., N.V., and J.v.R. wrote the manuscript. J.v.R. conceived of the conceptual ideas and supervised the study.

ACKNOWLEDGMENTS

We would like to thank Anko de Graaff, Stefan van der Elst, Reinier van der Linden, and Rinske Drost for technical assistance, Hans Clevers and Hugo Snippet for mice, and the NKI/AVL Core Facility Molecular Pathology & Biobanking (CFMPB) of the NKI/AVL Biobank for material and lab support. This work was supported by NWO (91710330, 700.10.402, 175.010.2007.00, and 834.11.002), ERC (648804), and KWF (HUBR 2009-4621).

Received: September 16, 2015

Revised: December 16, 2015

Accepted: February 2, 2016

Published: March 3, 2016

REFERENCES

- Beck, B., Lapouge, G., Rorive, S., Drogat, B., Desaedelaere, K., Delafaille, S., Dubois, C., Salmon, I., Willekens, K., Marine, J.C., and Blanpain, C. (2015). Different levels of Twist1 regulate skin tumor initiation, stemness, and progression. *Cell Stem Cell* 16, 67–79.
- Bill, R., and Christofori, G. (2015). The relevance of EMT in breast cancer metastasis: correlation or causality? *FEBS Lett.* 589, 1577–1587.
- Bukholm, I.K., Nesland, J.M., and Børresen-Dale, A.L. (2000). Re-expression of E-cadherin, alpha-catenin and beta-catenin, but not of gamma-catenin, in metastatic tissue from breast cancer patients [seecomments]. *J. Pathol.* 190, 15–19.
- Cheung, K.J., Gabrielson, E., Werb, Z., and Ewald, A.J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell* 155, 1639–1651.
- Del Pozo Martin, Y., Park, D., Ramachandran, A., Ombrato, L., Calvo, F., Chakravarty, P., Spencer-Dene, B., Derzsi, S., Hill, C.S., Sahai, E., and Malanchi, I. (2015). Mesenchymal cancer cell-stroma crosstalk promotes niche activation, epithelial reversion, and metastatic colonization. *Cell Rep.* 13, 2456–2469.
- Fantozzi, A., Gruber, D.C., Pisarsky, L., Heck, C., Kunita, A., Yilmaz, M., Meyer-Schaller, N., Cornille, K., Hopfer, U., Bentires-Alj, M., and Christofori, G. (2014). VEGF-mediated angiogenesis links EMT-induced cancer stemness to tumor initiation. *Cancer Res.* 74, 1566–1575.
- Fischer, K.R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S.T., Choi, H., El Rayes, T., Ryu, S., Troeger, J., et al. (2015). Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 527, 472–476.

- Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol. Cell. Biol.* *12*, 954–961.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646–674.
- Jeschke, U., Mylonas, I., Kuhn, C., Shabani, N., Kunert-Keil, C., Schindlbeck, C., Gerber, B., and Friese, K. (2007). 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* (4A), 1969–1974.
- Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* *119*, 1420–1428.
- Kowalski, P.J., Rubin, M.A., and Kleer, C.G. (2003). E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res.* *5*, R217–R222.
- Lim, J., and Thiery, J.P. (2012). Epithelial-mesenchymal transitions: insights from development. *Development* *139*, 3471–3486.
- Lin, E.Y., Jones, J.G., Li, P., Zhu, L., Whitney, K.D., Muller, W.J., and Pollard, J.W. (2003). Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am. J. Pathol.* *163*, 2113–2126.
- Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* *133*, 704–715.
- Morel, A.P., Lièvre, M., Thomas, C., Hinkal, G., Ansieau, S., and Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE* *3*, e2888.
- Ocaña, O.H., Córcoles, R., Fabra, A., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., and Nieto, M.A. (2012). Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* *22*, 709–724.
- 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.
- Patsialou, A., Bravo-Cordero, J.J., Wang, Y., Entenberg, D., Liu, H., Clarke, M., and Condeelis, J.S. (2013). Intravital multiphoton imaging reveals multicellular streaming as a crucial component of in vivo cell migration in human breast tumors. *Intravital* *2*, e25294.
- Roussos, E.T., Balsamo, M., Alford, S.K., Wyckoff, J.B., Gligorijevic, B., Wang, Y., Pozzuto, M., Stobezki, R., Goswami, S., Segall, J.E., et al. (2011). Mena invasive (Mena^{INV}) promotes multicellular streaming motility and transendothelial migration in a mouse model of breast cancer. *J. Cell Sci.* *124*, 2120–2131.
- Shibue, T., and Weinberg, R.A. (2009). Integrin beta1-focal adhesion kinase signaling directs the proliferation of metastatic cancer cells disseminated in the lungs. *Proc. Natl. Acad. Sci. USA* *106*, 10290–10295.
- Shibue, T., and Weinberg, R.A. (2011). Metastatic colonization: settlement, adaptation and propagation of tumor cells in a foreign tissue environment. *Semin. Cancer Biol.* *21*, 99–106.
- Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., and Clevers, H. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* *143*, 134–144.
- Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* *1*, 4.
- Stoletov, K., Kato, H., Zardoujian, E., Kelber, J., Yang, J., Shattil, S., and Klemke, R. (2010). Visualizing extravasation dynamics of metastatic tumor cells. *J. Cell Sci.* *1*, 2332–2341.
- Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nat. Rev. Mol. Cell Biol.* *7*, 131–142.
- Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatio-temporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* *22*, 725–736.
- Wang, W., Goswami, S., Lapidus, K., Wells, A.L., Wyckoff, J.B., Sahai, E., Singer, R.H., Segall, J.E., and Condeelis, J.S. (2004). Identification and testing of a gene expression signature of invasive carcinoma cells within primary mammary tumors. *Cancer Res.* *64*, 8585–8594.
- Wang, W., Wyckoff, J.B., Goswami, S., Wang, Y., Sidani, M., Segall, J.E., and Condeelis, J.S. (2007). Coordinated regulation of pathways for enhanced cell motility and chemotaxis is conserved in rat and mouse mammary tumors. *Cancer Res.* *67*, 3505–3511.
- Wellner, U., Schubert, J., Burk, U.C., Schmalhofer, O., Zhu, F., Sonntag, A., Waldvogel, B., Vannier, C., Darling, D., zur Hausen, A., et al. (2009). The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* *11*, 1487–1495.
- Welm, A.L., Sneddon, J.B., Taylor, C., Nuyten, D.S., van de Vijver, M.J., Hasegawa, B.H., and Bishop, J.M. (2007). The macrophage-stimulating protein pathway promotes metastasis in a mouse model for breast cancer and predicts poor prognosis in humans. *Proc. Natl. Acad. Sci. USA* *104*, 7570–7575.
- Wyckoff, J.B., Wang, Y., Lin, E.Y., Li, J.F., Goswami, S., Stanley, E.R., Segall, J.E., Pollard, J.W., and Condeelis, J. (2007). Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res.* *67*, 2649–2656.
- Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* *117*, 927–939.
- Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., et al. (2013). Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* *339*, 580–584.
- Zheng, X., Carstens, J.L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C.C., LeBleu, V.S., and Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* *527*, 525–530.

Cell Reports, Volume 14

Supplemental Information

**Plasticity between Epithelial
and Mesenchymal States Unlinks EMT
from Metastasis-Enhancing Stem Cell Capacity**

Evelyne Beerling, Daniëlle Seinstra, Elzo de Wit, Lennart Kester, Daphne van der Velden, Carrie Maynard, Ronny Schäfer, Paul van Diest, Emile Voest, Alexander van Oudenaarden, Nienke Vrisekoop, and Jacco van Rheenen

Supplemental figures and information

Figure S1. Overview of experimental setup and FACS strategy

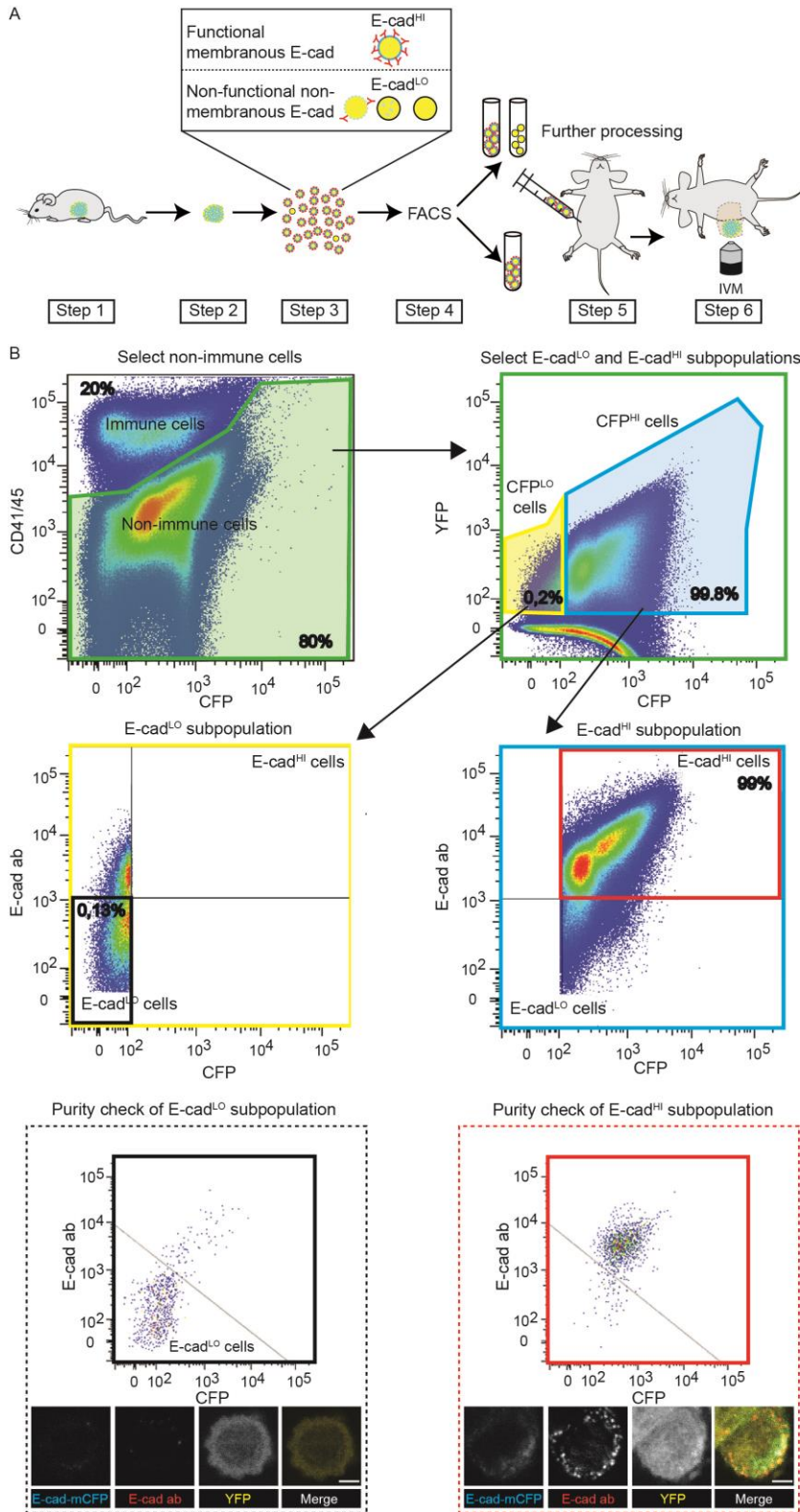


Figure S1. Overview of experimental setup and FACS strategy (related to Figure 2)

(A) Schematic overview of experimental setup: (Step 1) *MMTV-PyMT;MMTV-Cre;R26R-YFP;E-Cad-mCFP* mice with primary mammary tumor. (Step 2) Isolation of primary tumor. (Step 3) Preparation of single cell suspensions and labeling with E-cad antibody (red dotted lines). E-Cad^{HI} cells, cells with high membranous E-cad, will be labeled with the E-cad antibody. E-cad^{LO} cells, cells with low membranous E-cad due to low expression and/or intracellularly localized E-cad, will not be stained by the E-cad antibody. (Step 4) Single cell sorting of E-cad^{HI} and E-cad^{LO} tumor cells using FACS, based on E-cad-mCFP expression/E-cad antibody binding. (Step 5) Transplantation of E-cad^{HI} tumor cells and (Step 6) subsequent IVM.

(B) FACS strategy for E-cad^{LO} and E-cad^{HI} tumor cell sorting. A broad FSC SSC gate was followed by a gate excluding doublets, after which immune cells and megakaryocytes were excluded from the single cell population based on the expression of CD41/45 (upper left plot). YFP⁺ tumor cells (upper right plot) were subdivided in E-cad-mCFP^{HI} (blue frame) and E-cad-mCFP^{LO} (yellow frame) using very stringent gating, percentages are indicated in the frame. Finally, YFP⁺mCFP^{LO} tumor cells were additionally stringently gated for absence of E-cad antibody expression (black frame; E-cad^{LO}) and YFP⁺mCFP^{HI} tumor cells for high E-cad antibody expression (red frame; E-cad^{HI}). The sorted cell pools are then checked for purity (lower plots) and imaged using a confocal microscope. Separate channels and merged image are shown for both populations of tumor cells. Scale bar, 2 μ m.

Figure S2. The relative mRNA expression of E-cad^{lo} gene set

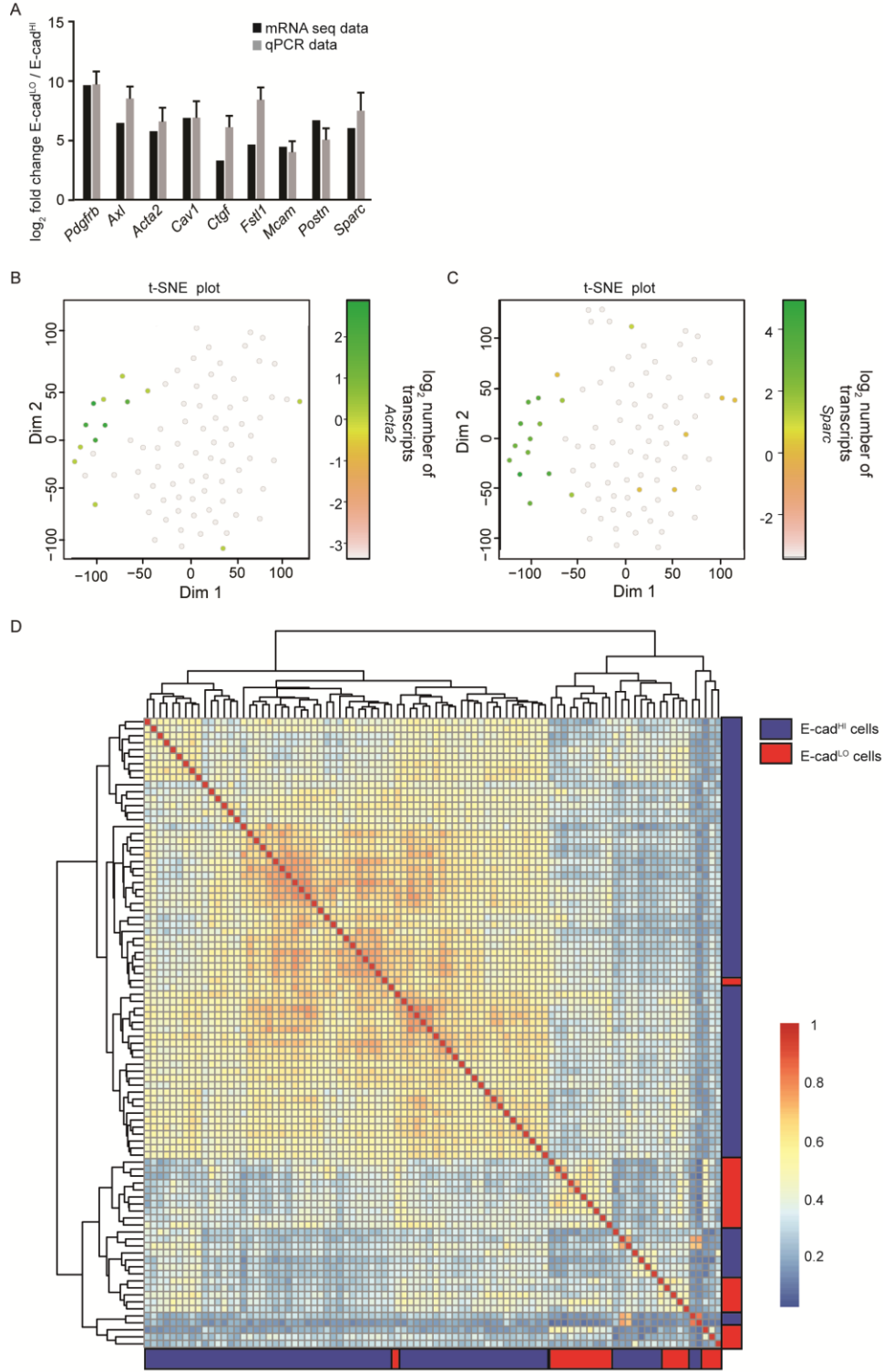


Figure S2. The relative mRNA expression of E-cad^{LO} gene set (related to Figure 2)

(A) Shown are the relative expressions of genes as determined by mRNAseq (black bars) and confirmed by reverse transcriptase (RT) qPCR (grey bars). Mean \log_2 fold change is shown of 3 mice, \pm SEM.

(B) Single cell sequencing data from E-cad^{HI} and E-cad^{LO} cells represented in a T-distributed stochastic neighbor embedding (t-SNE) intensity plot for Acta2. The scale bar for t-SNE plots is \log_2 .

(C) Single cell sequencing data from E-cad^{HI} and E-cad^{LO} cells represented in a t-SNE intensity plot for Sparc. The scale bar for t-SNE plots is \log_2 .

(D) Heat map was made based on Pearson's correlation between individual cells after filtering out genes that had less than 5 transcripts in at least 1 cell. The expression profiles of E-cad^{HI} cells show a Pearson correlation of 0.45 to each other and a Pearson correlation of 0.35 to E-cad^{LO} cells (Wilcoxon 2-sided test = $p < 10^{-15}$).

Figure S3. E-cad^{LO} tumor cells show plasticity

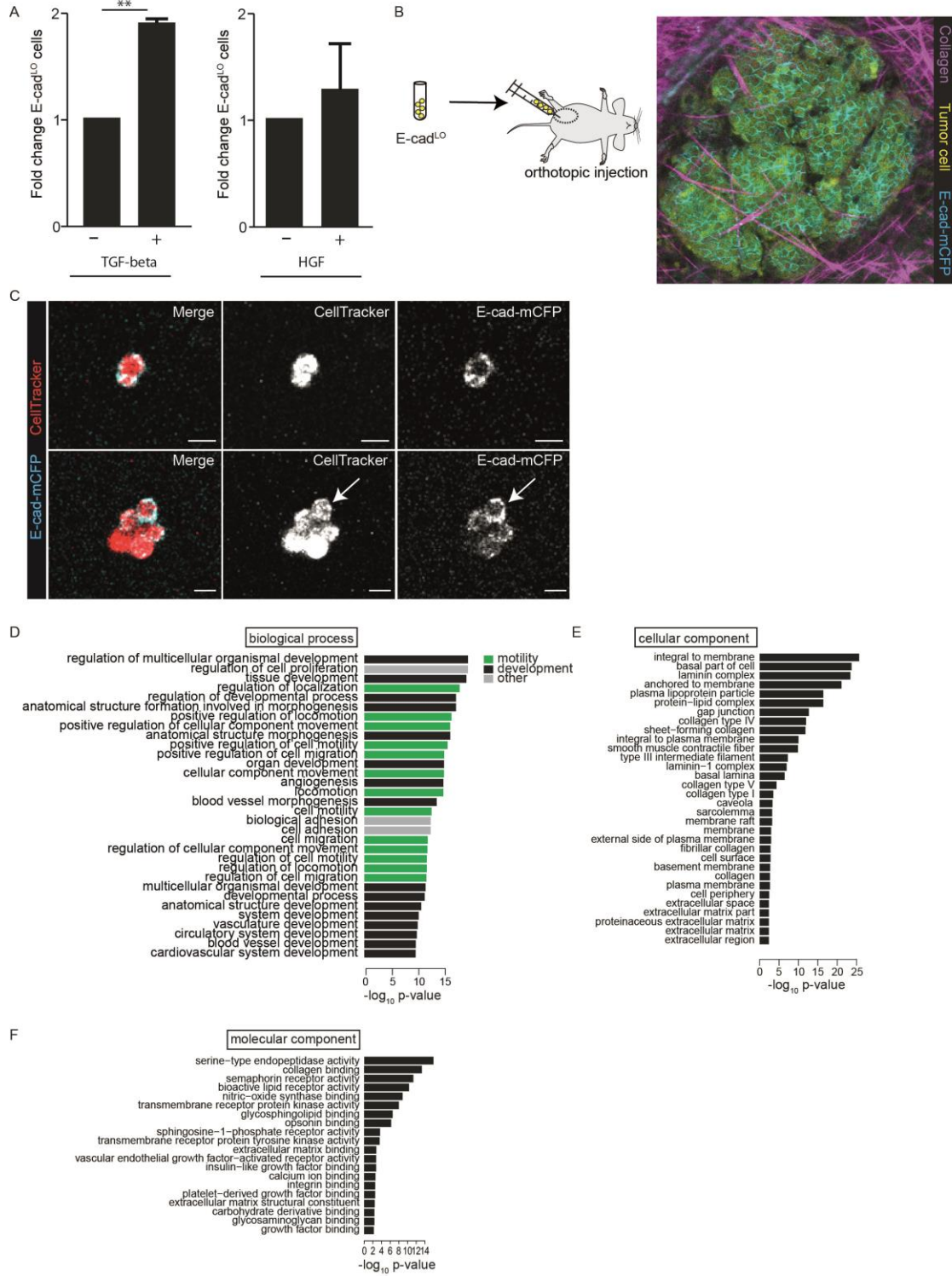


Figure S3. E-cad^{LO} tumor cells show plasticity (related to Figure 2)

(A) Organoids were cultured in collagen and stimulated with either HGF or TGF-beta. Fold change in the percentage of E-cad^{LO} cells compared to unstimulated organoids. Asterisks indicate a p value < 0,01 and n=3 independent replicates. Values are shown as SEM.

(B) An image of a PyMT tumor that developed upon orthotopic injection of E-cad^{LO} cells. Scale bar, 30 μ m.

(C) Representative images of E-cad^{LO} tumor cells plated in collagen one day after sorting. Shown are merged images (left panel) and individual channels in which the middle panel represents the tumor cells stains with CellTracker red and the right panel represents E-cad-mCFP. The white arrow indicates a cell that has divided. Scale bar, 20 μ m.

(D) Gene ontology list of biological processes of the E-cad^{LO} gene set divided into motility (green bars), developmental (black bars), and other (grey bars) categories. Barplots show Benjamini-Hochberg corrected p-values from WebGestalt (Zhang et al., 2005).

(E) Gene ontology of list cellular components of the E-cad^{LO} gene set.

(F) Gene ontology of list molecular components of the E-cad^{LO} gene set.

Figure S4. *In vivo* behavioral characterization of E-cad^{HI} and E-cad^{LO} tumor cells

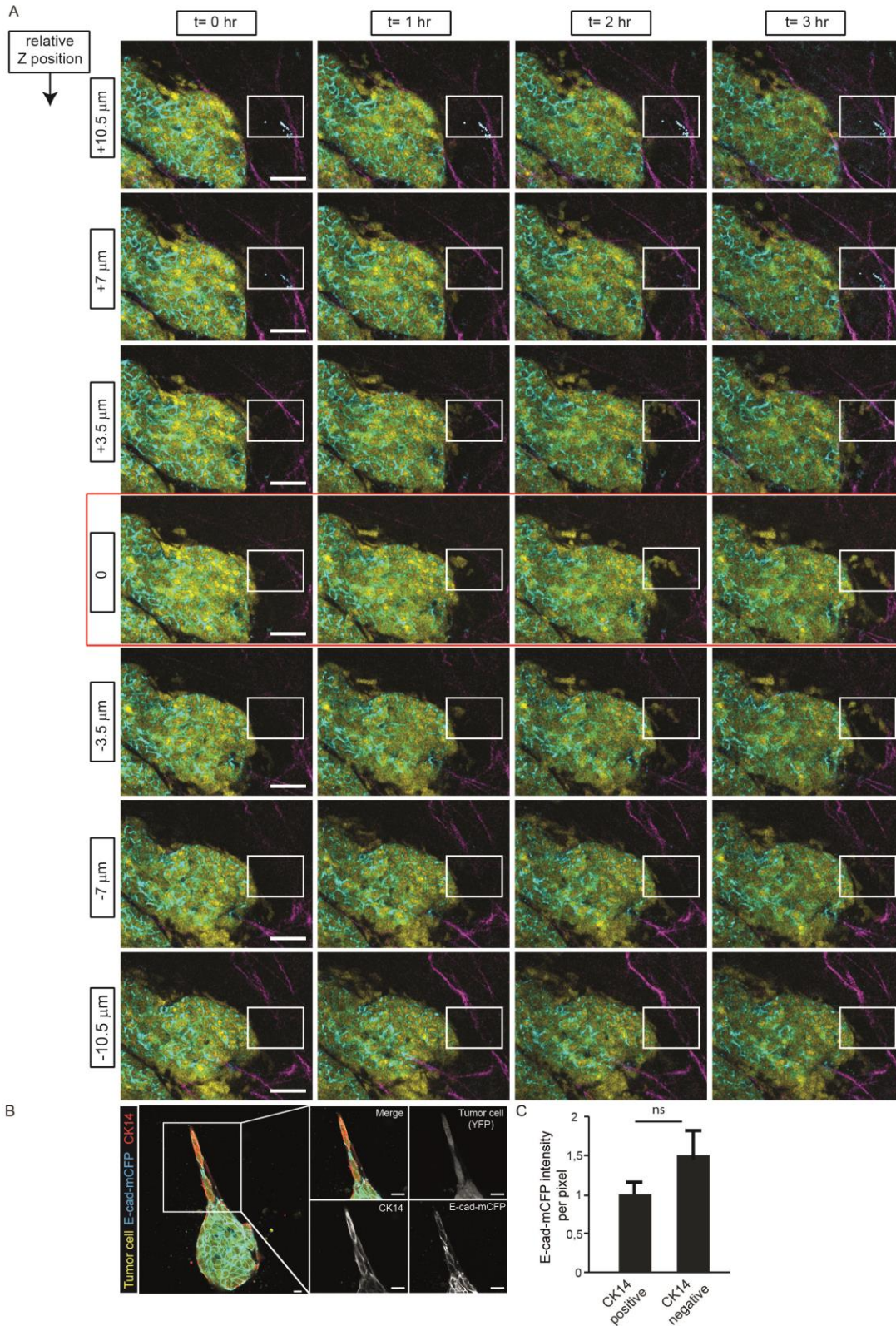


Figure S4. *In vivo* behavioral characterization of E-cad^{HI} and E-cad^{LO} tumor cells (related to Figure 4)

(A) Shown are the Z-stacks of the images in Figure 4C. The red box illustrates the Z-position shown in Figure 4C (Z-level 0). In the Z-levels 3.5 μm above and below, the cells are still visible, whereas the Z-levels 7 and 10.5 μm above and below show no cells in the circles. This observation excludes Z-drift as the source for cell movement and confirms true tumor cell movement. Scale bars, 50 μm .

(B) A representative image of an organoid stained with CK14 antibody. Left panel shows a merged image in which yellow represent tumor cells, cyan represent E-cad-mCFP and red is CK14 antibody staining. The right panels show a zoom of the protrusion in the organoids with in the left top a merged image and the individual channels shown in gray. Each channel is indicated in the figure. Shown is a single Z-level of 2.5 μm . Scale bar, 20 μm .

(C) Quantification of the E-cadherin intensity per pixel in CK14 positive and negative areas. SEM, n=8 organoids, 2 z-levels per organoid.

Figure S5. Circulation and proliferation of tumor cells

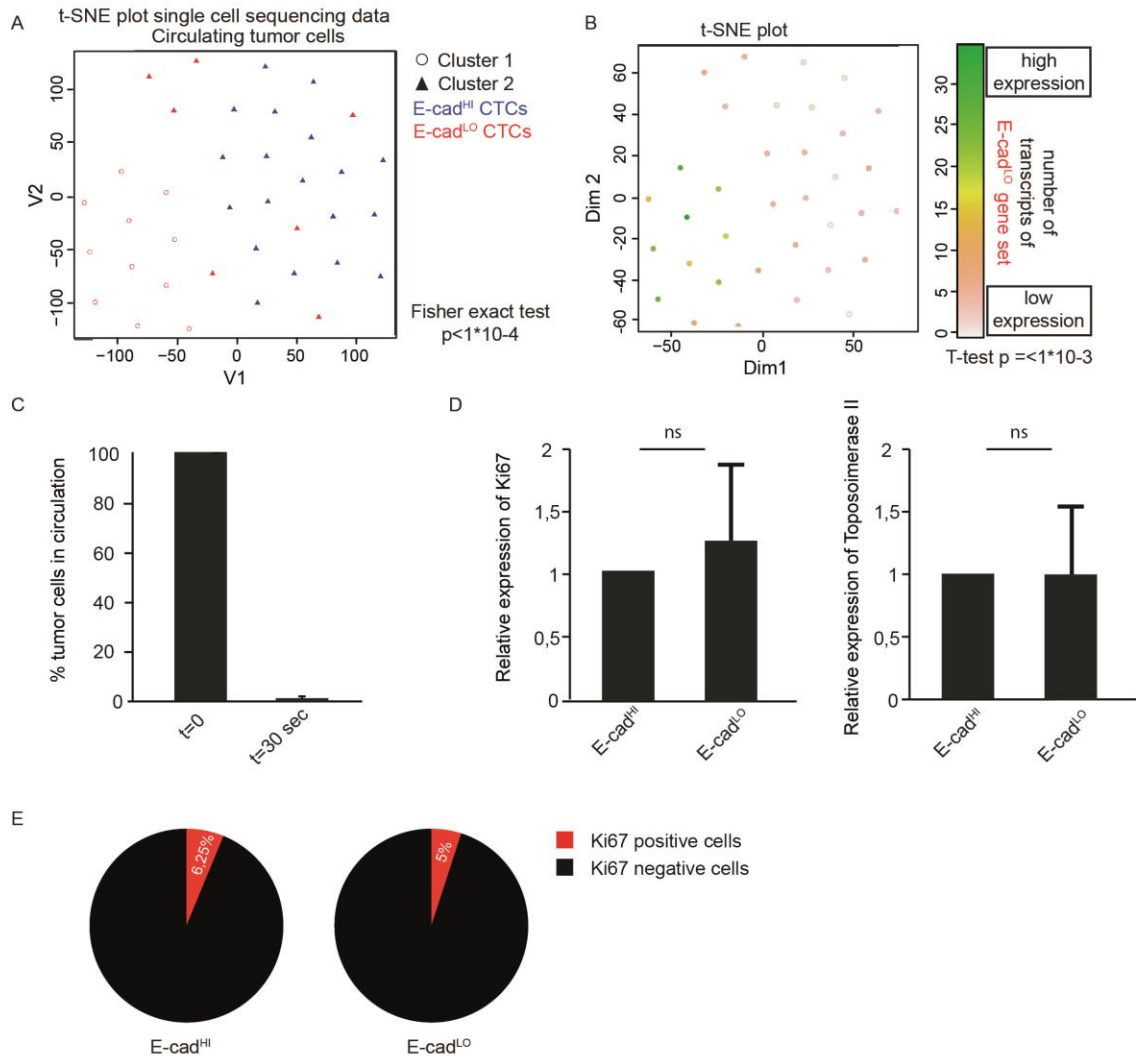


Figure S5. Circulation and proliferation of tumor cells (related to Figure 5)

(A) t-SNE plot of circulating tumor cells (CTCs). Using unsupervised K-medoids clustering, two separate clusters of CTCs were identified indicated as squares and triangles that overlap with E-cad^{HI} (blue) and E-cad^{LO} (red) tumor cells.

(B) t-SNE intensity plot of CTCs for the E-cad^{LO} gene set of Figure 2B. (C) Percentage of tumor cells found back in the circulation 30 seconds after tail vein injection. SEM, n=3.

(D) Relative expression of Ki67 and topoisomerase II in E-cad^{LO} cells by qPCR. SEM, n=4.

(E) Percentage of Ki67 positive tumor cells with immunohistochemistry staining after sorting by flow cytometry (n=120 cells).

Table S1. Significantly upregulated genes between E-cad^{HI} and E-cad^{LO} cells (related to Figure 2)

See separate Excel sheet.

Supporting Videos

Video S1. A three hour intravital timelapse of a lobe of a PyMT-induced tumor with non-motile cells (related to Figure 4). All tumor cells express YFP and fusion protein of mCFP to the endogenous E-cad. Shown are the separate channels for CFP (E-cad-mCFP, left), YFP (YFP, middle) and the merged images (Merge, right). The still images of this video are shown in Figure 4B. Scale bars, 50 μm .

Video S2. A three hour intravital timelapse of a lobe of a PyMT-induced tumor where the tumor edge is broken and cells migrate into the surrounding stroma (related to Figure 4). All tumor cells express YFP and fusion protein of mCFP to the endogenous E-cad. Shown are the separate channels for CFP (E-cad-mCFP, left), YFP (YFP, middle) and the merged images (Merge, right). The still images of this video are shown in Figure 4C. Scale bars, 50 μm .

Supplemental Materials and Methods

Human material

Four μm thick sections were cut from the paraffin block of an invasive ductal carcinoma and stained for E-cad using the Ventana BenchMark Ultra autostainer according to the manufacturer's instructions. Appropriate negative and positive controls were used throughout. Human tissues were obtained in compliance with Dutch law that does not require informed consent when leftover materials are used anonymously.

Mice

MMTV-PyMT and *MMTV-Cre* mice were purchased from Jackson Laboratory, Sacramento, CA, USA. *E-cad-mCFP* mice were a gift from Hans Clevers and *R26-loxP-stop-loxP-YFP* (*R26R-YFP*) mice a gift from Jacqueline Deschamps. Experiments were performed with *MMTV-PyMT;MMTV-Cre;R26R-YFP;E-Cad-mCFP* mice on a mixed and a pure FVB genetic background. Immunocompetent mice were housed under standard laboratory conditions and non-obese diabetic SCID IL-2 receptor gamma chain knockout (NSG) mice (own colony) were housed under IVC conditions. Mice received food and water ad libitum. All experiments were carried out in accordance with the guidelines of the Animal Welfare Committee of the Royal Netherlands Academy of Arts and Sciences, The Netherlands.

Flow cytometry on mouse material

From all mammary glands of *MMTV-PyMT;MMTV-Cre;R26R-YFP;E-Cad-mCFP* mice, tumors were collected and minced manually on ice using sterile scalpels. The tumor mass was digested in PBS supplemented with 25 $\mu\text{g}/\text{ml}$ DNase I (Roche) and 5 Wünsch units TH Liberase /ml (Roche) at 37 $^{\circ}\text{C}$ for 35 minutes, followed by mashing through a 70 μm filter (BD Falcon) while adding DMEM/F12 + GlutaMAX (GIBCO, Invitrogen Life Technologies) supplemented with 5% (v/v) fetal bovine serum (Sigma), 100 $\mu\text{g}/\text{ml}$ streptomycin and 100 U/ml penicillin (Invitrogen Life Technologies), 5 ng/ml insulin (I0516-5ML Sigma, St. Louis, MO, USA), 5 ng/ml EGF (Invitrogen) and 25 $\mu\text{g}/\text{ml}$ DNase. After spin down (4 minutes at 500 RCF at RT) the pellet was resuspended in 6 ml 5mM EDTA/PBS, after which a Ficoll gradient (Histopaque-1077, Sigma) was used to select for live cells (30 minutes at 400 RCF at RT, break 1). Cells were washed once in 5mM EDTA/PBS and centrifuged (4 minutes at 500 RCF at RT) before proceeding with antibody labeling.

Blood was collected via intracardiac injection, the red blood cells were depleted by NH₄Cl treatment. The remaining circulating tumor cells and immune cells were spun down (4 minutes 500 RCF at RT). Tumor cells and blood cells were blocked in 80% FACS buffer (5 mM EDTA in PBS supplemented with 5% fetal calf serum) / 20% serum mix (50/50 normal goat serum (monx10961, Monosan) and Fc γ II/III receptor blocking serum 2.4G2 (kind gift from Kiki Tesselaar, UMCU, The Netherlands)) for 10 minutes on ice before labeling with the following antibodies: rat anti-E-cad (DECMA-1 U 3254, Sigma) or E-cad-eFluor660 (DECMA-1, eBioscience), rabbit-anti-rat linker antibody (Southern Biotech), biotin-conjugated anti-mouse CD41 (clone eBioMWRreg30, eBioscience) and anti-CD45 (clone 30-F11, eBioscience). Secondary labeling was performed using goat-anti-rabbit AF647 (Invitrogen) and Streptavidin-PerCP (Biolegend). After putting cells through a 70 μm strainer cap (BD Falcon) cells were sorted on a FACS AriaII Special Ordered Research Product (BD Biosciences). The sort strategy is illustrated in Figure S1B. A broad FSC SSC gate was followed by a gate excluding doublets, after which immune cells and megakaryocytes were excluded in a dump channel. YFP⁺ tumor cells were subdivided in E-cad-mCFP^{HI} and E-cad-mCFP^{LO} using very stringent gating. Finally, YFP⁺mCFP^{LO} tumor cells were additionally stringently gated for absence of E-cad antibody expression (E-cad^{LO}) and YFP⁺mCFP^{HI} tumor cells for high E-cad antibody expression (E-cad^{HI}).

Flow cytometry on human material

Female breast cancer patients diagnosed with invasive ductal carcinoma who underwent mastectomy in the Netherlands Cancer Institute were included in the study.

The tumor material was received after pathological examination and kept on ice from that point onwards. Tumors were minced using sterile scalpels and digested in PBS supplemented with 25 $\mu\text{g}/\text{ml}$ DNase I (Roche) and 15 Wünsch units TH Liberase /ml (Roche) at 37 $^{\circ}\text{C}$ for 35 minutes, followed by mashing through a 70 μm filter (BD Falcon) while adding DMEM/F12 + GlutaMAX (GIBCO, Invitrogen Life Technologies), 1% HEPES, 1% Pen-Strep. Tumor cells were blocked in 80% FACS buffer (5 mM EDTA in PBS supplemented with 5% fetal calf serum) / 20% serum mix (50/50 normal goat serum (monx10961, Monosan) and Fc γ II/III receptor blocking serum 2.4G2 (kind gift from Kiki Tesselaar, UMCU, The Netherlands)) for 10 minutes on ice before labeling with E-cad-eFluor660 (DECMA-1, eBioscience) and EpCAM PE (1B7, eBioscience) labeled antibodies for 40 minutes on ice. DAPI was added as a life/death marker, cells were sorted on a FACS AriaII Special Ordered Research Product (BD

Biosciences) or on a FACS Jazz (BD biosciences). A broad FSC SSC gate was followed by a gate excluding doublets, after which the DAPI negative cells were selected. EpCAM positive tumor cells were subdivided in E-cad antibody positive and E-cad antibody negative cells.

Imaging sorted E-cad^{LO} cells

Cells were spun down at 0.2 RCF for 4 minutes. The cells were dissolved in medium with CellTracker Red CMPTX (Life technologies) as described by the manufacturer's instructions. After incubation cells were spun down and dissolved in rat-tail collagen I (1mg/ml) (Gibco, 3mg/ml) and plated in a glass bottom wilko dish. The cells were maintained with medium described below for organoids supplemented with RhoKinase Inhibitor (Abmole). Images were acquired using a Leica SP5 confocal microscope (Mannheim, Germany). All images were collected in 12 bit with a 25x (HCX IRAPO N.A. 0.95 WD 2.5 mm) water objective. CFP was excited at 458nm and emission was collected at 455-495 nm. CellTracker Red was excited at 633 nm and emission was collected at 640-700nm.

Ki67 staining sorted E-cad^{HI} and E-cad^{LO} cells

Cells were spun down after the FACS and fixated and permeabilized with BD cytofix/cytoperm (BD biosciences) according to manufacturer's instructions. Incubated with Ki67-Alexafluor 647 (clone B57, BD biosciences) for 30 minutes at 4 degrees. After washing cells were put on a glass slide and imaged at a Leica SP5 confocal microscope (Mannheim, Germany). All images were collected in 12 bit with a 25x (HCX IRAPO N.A. 0.95 WD 2.5 mm) water objective. Cells were imaged with DIC and Alexa647 was excited at 633 nm and emission was collected at 640-700nm. Cells were manually scored for positive staining.

Intravital imaging

For the intravital imaging experiments, YFP-expressing E-cad^{HI} tumor cells were isolated by flow cytometry and orthotopically transplanted in non-obese diabetic SCID IL-2 receptor gamma chain knockout mice to exclude potential YFP-labeling of non-epithelial lineages. For this, 1×10^5 sorted E-cad^{HI} tumor cells were injected in the right inguinal mammary gland of female non-obese diabetic SCID IL-2 receptor gamma chain knockout mice at 10-20 weeks of age. Mice bearing tumors of $\leq 500 \text{ mm}^3$ were used for intravital imaging. Mice were sedated using isoflurane inhalation anesthesia (1.5% to 2% isoflurane/O₂ mixture). The imaging site was surgically exposed, and the mouse was placed with its head in a facemask within a custom designed imaging box. The isoflurane was introduced through the facemask, and ventilated by an outlet on the other side of the box. The temperature of the imaging box and microscope were constantly adjusted to keep the mice between 36 and 37°C by a climate chamber that surrounds the whole stage of the microscope including the objectives. Imaging was performed on an inverted Leica TCS SP5 AOBS multi-photon microscope (Mannheim, Germany) with a chameleon Ti:Sapphire pumped Optical Parametric Oscillator (Coherent Inc. Santa Clare, CA, USA). The microscope is equipped with four non-descanned detectors: NDD1 (<455 nm), NDD2 (455-505 nm), NDD3 (505-550 nm), and NDD4 (560-650 nm). CFP was excited at 820 nm and detected in NND 2 and 3. YFP and Texas Red were excited at 960 nm. Second harmonic generation was detected in NDD1, YFP in NDD3, and Texas Red in NDD4. All images were collected in 12 bit and acquired with a 25x (HCX IRAPO N.A. 0.95 WD 2.5 mm) water objective. All images were processed using ImageJ software; pictures were converted to RGB, corrected for bleed through (if necessary), smoothed (if necessary), cropped (if necessary), rotated (if necessary) and contrasted linearly. Videos were corrected for XY and Z drift using custom-written software (codes on request available from J.v.R.).

Imaging analysis and quantification

Imaging positions showing migrating tumor cells were selected and single migrating and protruding tumor cells were counted manually. For every tumor cell, the E-cad-mCFP status (level and localization) was determined by 3 independent researchers, after which the cells were counted as mCFP^{HI} or mCFP^{LO}.

PyMT tumor organoids

PyMT tumor organoids were established as previously described (Nguyen-Ngoc et al., 2012). In short, tumors were harvested and enzymatically digested using trypsin (from bovine pancreas, Sigma) and collagenase A (Roche). The digested tumors were spun down in several steps until only the cell fragments of 200 to 1000 cells were left. These organoids were embedded in BME (RGF BME type 2 pathClear). Organoids were maintained in medium consisting of DMEM/F12 Glutamax supplemented with Hepes (1M Gibco), Penicillin-streptomycin, FGF (Life technologies) and B27 (50x Gibco). For experiments the organoids were plated in rat-tail collagen 1 (Gibco, 3mg/ml) in a concentration of 1mg/ml and the medium was further supplemented with 5 ng/ml HGF(RnD) or 100ng/ml TGF-beta protein (eBioscience).

Flow cytometry organoids

Organoids were harvested, spun down at 0.8 RCF for 3 minutes, digested in PBS supplemented with 25 µg/ml DNase I (Roche) and 5 Wünsch units TH Liberase /ml (Roche) at 37 °C for 20 minutes. Organoid cells were blocked in 80% FACS buffer (5 mM EDTA in PBS supplemented with 5% fetal calf serum) / 20% serum mix (50/50 normal goat serum (monx10961, Monosan) and FcγII/III receptor blocking serum 2.4G2 (kind gift from Kiki Tesselaar, UMCU, The Netherlands)) for 10 minutes on ice before labeling with E-cad-eFluor660 (DECMA-1, eBioscience). The cells were analyzed on FACS Jazz (BD biosciences) using the same strategy as the tumor cells.

Immunohistochemistry on organoids

Organoids were fixed in periodate-lysine-paraformaldehyde (PLP) buffer (2.5 ml 4% PFA in PBS + 0.0212 g NaIO₄ + 3.75 ml L-Lysine + 3.75 ml P-buffer (81 ml of 0.2 M Na₂HPO₄ and 19 ml of 0.2 M NaH₂PO₄ added to 100 ml demi water (pH 7.4)) for 1,5 hours at 4°C in a 4 compartment glass wilko dish, washed with P-buffer and incubated in 30% sucrose in P-buffer for 2 hours at 4°C. Samples were washed with PBS and blocked and permeabilized with PBS containing 1% Triton-X (Sigma), 10% FBS (Life technologies) and 1% BSA (Roche) for 3 hours at 4°C. CK14 Antibody (Covance, PRB-155P) was diluted in permeabilization buffer O/N at 4°C. Samples were washed with PBS and secondary anti-rabbit Alexa647 diluted in 1% BSA in PBS was incubated for 4 hours at 4°C and washed with PBS.

Images were acquired using a Leica SP5 confocal microscope (Mannheim, Germany). All images were collected in 12 bit with a 25x (HCX IRAPO N.A. 0.95 WD 2.5 mm) water objective. CFP was excited at 458nm and emission was collected at 455-495 nm. YFP was excited at 514nm and emission was collected at 520-560 nm. Alexa-647 was excited at 633 nm and emission was collected at 640-700nm. At least ten organoids with protrusions were analyzed using ImageJ software.

Western blot

Cells were lysed in 1% SDS buffer and equal amounts of protein were loaded onto an 8% SDS/PAGE gel. Antibodies against the following proteins were used: mouse-anti E-cad (clone 36/E-Cadherin, BD Biosciences), guinea pig-anti Vimentin (Fitzgerald), mouse-anti ZO-1 (Invitrogen), rabbit-anti N-cadherin (clone GC-4, Takara), mouse-anti Twist (Santa Cruz Biotechnology), mouse-anti β-actin (clone AC-15, Sigma). Immunoreactive bands were detected by HRP-conjugated secondary antibody incubation (mouse NA931V; rabbit NA934V, GE Healthcare, guinea pig ab 97155 Abcam) and ECL treatment (Thermo Scientific) according to the manufacturer's instructions.

RNA isolation and mRNA-sequencing

RNA was isolated using Trizol reagent (Invitrogen Life Technologies) according to the manufacturer's protocol and stored at -80°C. The amount and purity of isolated RNA was analyzed by the Nanodrop spectrophotometer (Wilmington, DE, USA). RNAseq libraries were prepared using Clontech's SMARTer Ultra Low RNA Kit and sequenced on an Illumina HiSeq.

mRNA-sequencing analysis

Paired end sequencing reads were mapped to mouse mm9 using TopHat2/Bowtie2 (Kim et al., 2013). Differential expression was calculated using CuffDiff2 (Trapnell et al., 2013) using UCSC genome annotation. A complete list of significantly upregulated genes between E-cad^{HI} and E-cad^{LO} tumor cells can be found in Table S1 and the data is deposited in the European Nucleotide Archive (accession number is: PRJEB5939). To compare the gene expression profile from our mouse model with human datasets we used a human-mouse ortholog database from the UCSC genome browser.

HMLE expression data was taken from (Onder et al., 2008) (GEO accession: GSE9691). Affymetrix data was normalized using RMA (Irizarry et al., 2003) from the affy Bioconductor package. Both the E-cad shRNA and the E-cad dominant negative data were compared to the wild-type data. Differential expression in Figure 3A was defined as the deviation from the median over times 3 replicate experimental and 3 replicate control experiments. Single molecule RNAseq data was taken from GEO accession GSE41245 (Yu et al., 2013). We performed quantile normalization on the raw expression values and calculated the differential expression values between EpCAM⁺ cells and the corresponding IgG⁺ cells. Expression values were averaged per gene over multiple experiments belonging to the same class (i.e. "blood specimen healthy donor", "epithelial CTC" and "mesenchymal CTC"). The set of upregulated genes was analyzed using the WebGestalt tool (Zhang et al., 2005) for functional classification. The set of upregulated RNAs was used as the input set; as a background set we used all the genes in the mouse genome.

Single cell mRNA sequencing

Single E-cad^{HI} or E-cad^{LO} cells were sorted into trizol. The RNA was extracted using chloroform and precipitated with iso-propanol. The RNA-pellet was then processed using the CEL-seq protocol as described in (Hashimshony et al., 2012; Grun et al., 2014) and sequenced on an Illumina Nextseq using 75bp paired end sequencing. After sequencing, read 1 was aligned to the mm10 RefSeq mouse transcriptome downloaded from the UCSC genome browser using with default parameters. Read 2 contains a barcode identifying the sample from which the read originated. CEL-seq only sequences the most 3' prime end of a transcript and generates one read per transcript. Data analysis was done in R. Samples with less than 1000 transcripts were discarded after which the remaining samples were downsampled to 1000 transcripts. T-distributed stochastic neighbor embedding (t-SNE) was performed on the Euclidean distances between cells. Heat map was made based on Pearson's correlation between cells after filtering out genes that had less than 5 transcripts in at least 1 cell.

cDNA preparation and qPCR

Complementary DNA was synthesized using the Avian Myeloblastosis Virus Reverse Transcriptase kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. Sequences of used primers can be found below. qPCR was performed using Power SYBR Green PCR Master Mix (Applied Biosystems). Thermal cycle conditions used for all qPCR reactions were as follows: 5 min at 95°C, followed by 40 cycles consisting of denaturation for 30 sec at 95°C, annealing for 30 sec at 58°C, and extension for 1 min at 72°C. PCR reactions were concluded with incubation for 10 min at 72°C to complete the extension of all synthesized products. Relative quantification values were calculated using the ddCt method and mean values were plotted with SD using GraphPad Prism v5.0 (GraphPad Software, Inc.).

The following primers were used for qPCR: E-cadherin (Fwd: 5'- GCT TCA GTT CCG AGG TCT AC -3', Rev: 5'-GCC AGT GCA TCC TTC AAA TC -3'), Twist-1 (Fwd: 5'- CGG AGA CCT AGA TGT CAT TGT TT -3', Rev: 5'-CGC CCT GAT TCT TGT GAA TTT G -3'), Vimentin (Fwd: 5'- GAG GAG ATG CTC CAG AGA GA -3', Rev: 5'- TCC TGC AAG GAT TCC ACT TT -3'), N-Cadherin (Fwd: 5'- GTG GAG GCT TCT GGT GAA AT -3', Rev: 5'- GGC TCG CTG CTT TCA TAC T -3'), Snail1 (Fwd: 5'- CTG CAC GAC CTG TGG AAA -3', Rev: 5'- GGC ACT GGT ATC TCT TCA CAT C -3'), Fibronectin (Fwd: 5'- GGT GTC CGA TAC CAG TGT TAC -3', Rev: 5'- TCT CCG TGA TAA TTA CTT GGA CAG -3'), ZO-1 (Fwd: 5'- GTC ACG ATC TCC TGA CCA AC -3', Rev: 5'- CCA GGT TTA GAC ATT CGC TCT -3'), Pdgfrb (Fwd: 5'- AGC GAG AAG CAA GCC TTA AT -3', Rev: 5'- GAT CAC CGT ATC GGC AGT ATT C -3'), Axl (Fwd: 5'-AAA CTC CAG GCC TGA ACA AG- 3', Rev: 5'- TTG GCA TTG TGG GCT TCA - 3'), Acta2 (Fwd: 5'-TAA GGC CAA CCG GGA GAA-3', Rev: 5'-GGG ACA TTG AAG GTC TCA AAC A- 3'), Cav1 (Fwd: 5'- CAA CAT CTA CAA GCC CAA CAA C - 3', Rev: 5'- TCC CTT CTG GTT CTG CAA TC - 3'), Ctgf (Fwd: 5'- AAC CGC AAG ATC GGA GTG - 3', Rev: 5'- TGC TTT GGA AGG ACT CAC C - 3'), Fstl1 (Fwd: 5'- GGC TGG AAG CTG AGA TCA TT -3', Rev: 5'- CAG GTG AGA GTC GCC ATT AT -3'), Mcam (Fwd: 5'- CCC ATT CCT CAA GTC CTA TGG - 3', Rev: 5'- CCC ATT CCT CAA GTC CTA TGG - 3'), Postn (Fwd: 5'- CAG CTC CTG TAA GAA CTG GTA TC -3', Rev: 5'- ATA TAG CCA GGG CAG CAT TC - 3'), Sparc (Fwd: 5'- GCT GGA TCA GCA CCC TAT T -3', Rev: 5'- TGT CTA GGT CAC AGG TCT CA -3'), Topoisomerase II (Fwd: 5'- AAC GAG AGA CAC ATC ATT GTC AG-3', Rev: 5'- TCA CCT TCC CTA TCA CAG TCC-3'), Ki67 (Fwd: 5'- ACC GTG GAG TAG TTT ATC TGG-3', Rev: 5'- TGT TTC CAG TCC GCT TAC TTC T -3')

Mouse tumor and tissue processing for histology

Tumors and other tissues were isolated from the mice at the end of the experiment and fixed in periodate-lysine-paraformaldehyde (PLP) buffer (2.5 ml 4% PFA + 0.0212 g NaIO₄ + 3.75 ml L-Lysine + 3.75 ml P-buffer (pH 7.4)) O/N at 4°C. The following day, the fixed tumors and tissues were washed twice with P-buffer and placed for at least 6 hours in 30% sucrose at 4°C. The tumors and tissues were then embedded in tissue freezing medium (Leica Microsystems, Nussloch, Germany) and stored at -80°C before cryosectioning.

Human immunohistochemistry

E-cad staining on 4 µm paraffin-embedded tumor slices was performed using mouse anti-E-cad antibody (clone 36/E-Cadherin, BD Biosciences) overnight, followed by an anti-mouse horseradish peroxidase labeled secondary antibody (Dako, Envision+ System- HRP labeled Polymer anti-mouse).

Analyzing metastatic outgrowth in the lungs

Tumor or tissue cryosections (20 μm thick tumor sections and 100-150 μm thick lungs sections) were rehydrated for 10 min in PBS and embedded in Vectashield mounting medium (hard set; Vector Labs, Burlington, Ontario, Canada). When indicated, cryosections were counterstained with TO-PRO3 (Invitrogen molecular probes, Paisley, UK) to visualize the nuclei. Images were acquired using a Leica SP5 or SP8 confocal microscope (Mannheim, Germany) equipped with 10x NA 0.3 and 20x NA 0.7 dry objectives. CFP was excited at 405 nm or at 458nm and emission was collected at 455-495 nm for CFP. YFP was excited at 514nm and emission was collected at 520-560 nm. TO-PRO-3 was excited at 633 nm and emission was collected at 650-670nm. At least 4 frozen sections per mouse were analyzed.

Determination of the time in circulation for tumor cells

PyMT tumor organoids were cultured as described above and processed until single cells as described with flow cytometry for organoids. The cells were counted on a hemocytometer and 10,000 cells were diluted in the appropriate amount of PBS for a tail vein injection. The mice were anesthetized using isoflurane. The tumor cells were injected into the tail vein of the mouse and after 30 seconds blood was harvested via cardiac injection of the right heart chamber. The blood was processed for flow cytometry as described above.

Intrahepatic injection of tumor cells

From all mammary glands of *MMTV-PyMT;MMTV-Cre;R26R-YFP;E-Cad-mCFP* mice, tumors were collected and processed to single cell suspension as described for flow cytometry (see above). After spin down (4 minutes at 500 RCF at RT) the pellet was suspended in PBS. For the formation of hepatic tumors 5×10^5 tumor cells were injected. The appropriate amount of cells was spun down (4 minutes at 500RCF at RT) and suspended in Matrigel (Corning Matrigel Basement Membrane Matrix Growth Factor Reduced, Phenol Red Free). NSG mice were sedated using isoflurane inhalation anesthesia (1.5% to 2% isoflurane/O₂ mixture) and received subcutaneous analgetica of 3mg/100ul buprenorphine (Temgesic). After a midline incision, the liver was mobilized and the cells were injected superficially into the liver. After coagulation of potential bleeding caused by the injection, the liver was placed back and the abdomen was closed. Tumor cells were injected into the liver of different recipient mice for five subsequent rounds to prime the tumor material for growing out in the liver. After the final round, tumors were harvested and processed for FACS as described above.

Cells obtained from cell sorting were spun down (4 minutes at 500 RCF at RT) and suspended in Matrigel (Corning Matrigel Basement Membrane Matrix Growth Factor Reduced, Phenol Red Free) and injected into the liver of three NSG mice. Tumors developed within 4-8 weeks and the sets of mice were sacrificed by cervical dislocation when one of the three showed signs of tumor development. The tumors and organs were then processed as described above.

Statistical analyses

Statistical analyses were performed using Graphpad software, We used a unpaired t-test with Welsh correction for the E-cadherin intensity in CK14 positive and negative areas, a one-sample t-test for the induction of EMT with TGF-beta or HGF in organoids. The proliferation difference in E-cad^{LO} cells compared to E-cad^{HI} cells was performed with a one-sample t-test compared to hypothetical value one.

Tumor-initiating cell frequency was tested by Elda-limiting dilution test on this website <http://bioinf.wehi.edu.au/software/elda/> (Hu et al. 2009).

Supplemental references:

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell reports* 2, 666-673.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* 14, R36.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology* 31, 46-53.

Grun, D., Kester, L., and van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nature methods* 11, 637-640

Nguyen-Ngoc, K.V., Cheung, K.J., Brenot, A., Shamir, E.R., Gray, R.S., Hines, W.C., Yaswen, P., Werb, Z., and Ewald, A.J. (2012). ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2595-2604.

Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research* 33, W741-748.

Hu, Y, and Smyth, GK (2009). ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *Journal of Immunological Methods* 347, 70-78.