HHS Public Access

Author manuscript

Nat Cell Biol. Author manuscript; available in PMC 2015 November 01.

Published in final edited form as:

Nat Cell Biol. 2015 May ; 17(5): 651–664. doi:10.1038/ncb3148.

Identification of Molecular Determinants of Primary and Metastatic Tumor Re-Initiation in Breast Cancer

Jason B. Ross[1], **Doowon Huh**[1], **Lisa B. Noble**[1], and **Sohail F. Tavazoie**[1,*]

[1]Laboratory of Systems Cancer Biology, Rockefeller University

Abstract

Through *in vivo* selection of multiple ER-negative human breast cancer populations for enhanced tumor-forming capacity, we have derived sub-populations that generate tumors more efficiently than their parental populations at low cell numbers. Tumorigenic-enriched (TE) sub-populations displayed increased expression of *LAMA4*, *FOXQ1* and *NAP1L3*—genes that are also expressed at greater levels by independently derived metastatic sub-populations. These genes promote metastatic efficiency. *FOXQ1* promotes *LAMA4* expression, while *LAMA4* enhances clonal expansion upon substratum-detachment *in vitro*, tumor re-initiation in multiple organs, and disseminated metastatic cell proliferation and colonization. *LAMA4*'s promotion of cancer cell proliferation and tumor re-initiation requires β1-integrin. Increased *LAMA4* expression marks the transition of human pre-malignant breast lesions to malignant carcinomas, while tumoral *LAMA4* over-expression predicts reduced relapse-free survival in ER-negative patients. Our findings reveal common features that govern primary and metastatic tumor re-initiation and identify a key molecular determinant of these processes.

Cancer progression is characterized by the formation of tumors in primary organs and subsequent re-formation of tumors in metastatic sites[1–4]. During primary tumorigenesis, malignant cancer cells invade into the surrounding stromal compartment and must survive and proliferate in the absence of their previous attachment to the basement membrane (BM) or other extracellular matrix (ECM) proteins[5]. These early steps of malignant tumor formation can be experimentally modeled by primary xenograft 'tumor re-initiation' assays, which assess the capacity of human cancer cells implanted into a primary organ site to re-

*Corresponding author: Sohail Tavazoie, Leon Hess Associate Professor, Head, Laboratory of Systems Cancer Biology, Rockefeller University, Box 16, 1230 York Avenue, New York, NY 10065 USA, Phone: 212-327-7208, Fax: 212-327-7209, stavazoie@mail.rockefeller.edu.

AUTHOR CONTRIBUTIONS

S.F.T. conceived the project and supervised all research. J.B.R and S.F.T. wrote the manuscript. J.B.R. conducted *in vivo* selection for tumor re-initiation. J.B.R., D.H., and L.B.N. designed, performed and analyzed the experiments.

Accession Numbers

Primary Accessions: the microarray gene expression data generated in this study from the parental MDA-MB-231 and CN34 cell lines and their respective tumorigenic-enriched (TE) and lung-metastatic (LM) derivatives has been deposited and can be obtained from GEO under GSE52629. Referenced Accessions: previously published datasets used in this study were obtained from GEO under GSE41197 (ref. 47), GSE14548 (ref. 49), GSE3893 (ref. 50), GSE7390 (ref. 51), GSE25066 (ref. 52), and GSE2034 (ref. 54).

SUPPLEMENTARY INFORMATION

Information supplementing this article consists eight figures and four tables.

initiate tumors in a secondary host[6]. While comparison of cancer cells with differing tumorigenic capacities has led to the discovery of many important biological mediators of tumor-forming potential[7–9], the relationship of highly tumorigenic cells to metastatic disease has not been systematically explored[10–11], and whether the primary tumor-forming potential of cancer cells is sufficient to also enable the propagation of tumors at distant sites during metastatic progression is a question of considerable interest[10].

In order to investigate the biological features and molecular determinants governing primary and metastatic tumor re-initiation, we developed an unbiased approach to select for cells with enhanced tumor-forming capacity. Analogous to the previous use of *in vivo* selection to select for and study highly metastatic sub-populations[4,12–17], we sought to select sub-populations of cancer cells that phenotypically demonstrate enhanced tumor-forming capacity. We focused on Estrogen Receptor-negative (ER-negative) breast cancer, an aggressive subset of breast cancer in need of targeted therapies[18]. We subjected multiple ER-negative human breast cancer cell populations to *in vivo* selection for enhanced tumor re-initiation capacity in a xenograft model. This strategy yielded tumorigenic-enriched (TE) populations that demonstrated enhanced tumor re-initiation capacity in multiple organ microenvironments. Transcriptomic profiling of TE sub-populations revealed a set of genes–*LAMA4*, *FOXQ1*, and *NAP1L3*–to be expressed at greater levels in these cells relative to their parental populations. These genes were also expressed at greater levels by highly metastatic cells independently derived from the same parental populations, and each of these genes promoted metastatic efficiency *in vivo*. Further characterization of *LAMA4* revealed it to enhance proliferation during substratum-detachment *in vitro*, tumor re-initiation in multiple microenvironments, the proliferation of disseminated metastatic cells, and the formation of micro-metastatic colonies *in vivo. LAMA4*'s ability to drive cancer cell proliferation and tumor re-initiation requires β1-integrin. In early breast cancer lesions, malignant breast cancer cells express higher levels of *LAMA4* relative to pre-malignant cells, while *LAMA4* expression in established tumors stratifies ER-negative breast cancer patients into those with worse relapse-free survival (*LAMA4* high) and those with improved relapse-free survival (*LAMA4* low). Collectively, our selection for sub-populations of cells with enhanced tumor-forming potential establishes a robust model to interrogate the molecular basis of tumor re-initiation across multiple organ sites. These findings have uncovered a key molecular determinant of these processes in breast cancer, and validate this unbiased approach for discovery of genes and phenotypes that govern re-initiation by malignant cells.

RESULTS

In vivo selection for tumor re-initiation enriches for populations with enhanced tumor-forming capacity

In order to study the biology that governs breast cancer tumor re-initiation, we used *in vivo* selection to select for sub-populations of human breast cancer cells with enhanced tumor-forming capacity. We applied selective pressure for tumor re-initiation at low cell numbers by injecting increasingly limiting numbers of breast cancer cells orthotopically into the mammary fat pads of immunodeficient mice in order to generate xenograft tumors over successive rounds of serial dilution (Fig. 1a). Independent tumorigenic human breast cancer

cell lines, the MDA-MB-231 (MDA-231) line[14,19] and the minimally passaged CN34 line[16], were subjected to *in vivo* selection. These cell lines were selected on the basis of their ER-negative status[20]. Upon injection into the mammary fat pads of immunodeficient mice, both cell lines gave rise to tumors at non-saturating (less than 100-percent) frequencies at the initial cell doses used (10,000 or 20,000 cells, for the MDA-231 or CN34 cell lines, respectively) during the first round of *in vivo* selection (Fig. 1b). Multiple additional rounds of *in vivo* selection yielded tumorigenic-enriched (TE) derivatives MDA-TE3 and CN34-TE2 (Fig. 1b), which were propagated and expanded *in vitro*. TE derivatives gave rise to xenograft tumors at significantly greater frequencies when injected into the mammary fat pads of immunodeficient mice at low cell numbers (100 or 2,000 cells, for the MDA-231 or CN34 cell lines, respectively) relative to their respective parental populations, demonstrating their enhanced orthotopic tumor re-initiation capacity (Fig. 1c,d). Subsequent *in vitro* experiments revealed that the TE derivatives surprisingly proliferated and formed colonies to a lesser extent than their parental populations upon standard adherent cell culture conditions (Supplementary Fig. 1a–d), did not demonstrate significant differences in their capacity to attach to tissue culture plates (Supplementary Fig. 1e,f), and did not recruit a greater number of endothelial cells relative to their parental populations (Supplementary Fig. 1g,h). These findings suggested that the enhanced *in vivo* tumor-forming capacity demonstrated by TE derivatives was independent of multiple *in vitro* phenotypes typically considered to confer a pro-tumorigenic advantage. Additionally, immunophenotypic characterization of the MDA-TE3 or CN34-TE2 derivatives did not reveal enrichment of $CD44^{+}/CD24^{-}$ marker profiles relative to their respective parental populations (Supplementary Fig. 1i,j). Collectively, these results demonstrate that sub-populations of cells with significantly enhanced tumor re-initiation capacity can be derived from human breast cancer populations through *in vivo* selection.

Tumorigenic-enriched populations exhibit enhanced colonization of multiple ectopic organs relative to their parental populations

The relationship between primary tumor-forming potential and metastatic activity is a fundamental question of great interest in cancer biology[10,21,22]. Although highly tumorigenic cells have been posited to be more able to metastasize[23], some studies have suggested these processes to be independent[8,24]. As a first step to address this issue, we examined whether TE derivatives were more effective at re-initiating tumors in a metastatic organ relative to their parental populations. When injected directly into the lung parenchyma, TE derivatives gave rise to a greater number of macroscopic colonies in the lungs relative to their respective parental populations (Fig. 2a–d), demonstrating that they were enriched in ectopic tumor re-initiation capacity in addition to orthotopic tumor re-initiation capacity. This finding prompted us to ask if the enhanced tumor-forming potential of TE cells was sufficient to enhance their colonization of organs during metastasis. Because organ-specific features such as the endothelial tight-junctions of the lungs represent significant barriers to cell extravasation[25–26], we hypothesized that the TE derivatives might demonstrate an increased capacity to metastasize to an organ with less stringent requirements for trans-endothelial migration, whereby colonization is primarily driven by tumor-forming capacity. When we assessed the capacity of TE cells to metastasize to the liver, an organ whose fenestrated sinusoidal capillaries more readily permit cell seeding[27],

TE cells gave rise to a significantly greater number of macroscopic colonies in the liver relative to their respective parental populations (Fig. 2e–h). Taken together, these findings indicate that *in vivo* selection for orthotopic tumor re-initiation is sufficient to co-select for enhanced colonization of multiple organs.

LAMA4, *FOXQ1*, and *NAP1L3* display increased expression levels in tumorigenic-enriched and highly metastatic cells and promote metastatic efficiency

Having selected for TE derivatives with enhanced tumor re-initiation capacity, we next used a systematic approach to identify genes that might govern this process. We first focused on genes whose expression was greater in TE cells relative to their parental populations. Whole-genome transcriptomic profiling of TE derivatives from the MDA-231 and CN34 cell lines revealed a large set of genes (169, Supplementary Table 1) that displayed at least 1.5-fold greater expression in these cells relative to their respective parental populations (Fig. 3b). Interestingly, this list contained several genes that were expressed at greater levels in highly lung-metastatic (LM) cells (derived through *in vivo* selection of the MDA-231 and CN34 cell lines)[14], relative to the same parental populations, an overlap that was highly statistically significant ($P=2.91\times10^{-6}$). We posited that the set of genes expressed at greater levels in both tumorigenic-enriched and highly metastatic derivatives might represent general regulators of tumor re-initiation whose function was not restricted to i) the organ of selection or ii) features only specific to primary tumor formation or metastasis (Fig. 3a). Comparison of the set of genes that displayed greater expression in the TE derivatives (Supplementary Table 1) and in LM sub-lines (Supplementary Table 2) that were previously derived from the MDA-231[14] and CN34[16] cell lines revealed 8 candidate genes that displayed a greater than 1.5-fold increased expression in all *in vivo* selected derivatives relative to their respective parental populations (Fig. 3b, Supplementary Table 3). Quantitative real-time PCR validation of this set yielded three genes–*LAMA4*, *FOXQ1*, and *NAP1L3*–that displayed significantly greater levels of expression in TE and LM derivatives from both cell lines relative to their respective parental populations (Fig. 3c,d). *LAMA4* encodes the secreted alpha chain isoform protein laminin-α4, an ECM component mainly present in basement membranes[28]. *FOXQ1* is a member of the FOX family of transcription factors with roles in development[29] and cancer progression[30]. *NAP1L3* encodes a poorly characterized member of the nucleosome assembly protein (NAP) family[31].

The increased expression of *LAMA4*, *FOXQ1*, and *NAP1L3* in tumorigenic-enriched and highly metastatic derivatives relative to their parental populations led us to speculate that these genes might regulate tumor re-initiation. As a first step to assess the function of these genes, we asked if they regulated the formation of macroscopic nodules by highly metastatic cells. Knockdown of *LAMA4*, *FOXQ1*, or *NAP1L3* in highly metastatic MDA-LM2 and CN34-LM1a derivatives significantly decreased lung-metastatic colonization as quantified by bioluminescence imaging (Fig. 3e,f, Supplementary Fig. 2a–c), and histological analysis of lungs from these mice revealed fewer macro-metastases formed upon knockdown of each of these genes (Fig. 3g–h). Collectively, these findings reveal a set of genes (*LAMA4*, *FOXQ1*, and *NAP1L3*) whose expression is increased in sub-populations independently selected for enhanced tumorigenic and metastatic activity, and whose individual expression promotes metastatic efficiency *in vivo*.

LAMA4 promotes tumor re-initiation in multiple organ microenvironments

Having identified a set of genes (*LAMA4*, *FOXQ1*, and *NAP1L3*) expressed at greater levels in cells with enhanced tumorigenic and metastatic potential and that individually enhanced metastatic efficiency, we focused on one of these genes, *LAMA4*, for further study. Our focus on *LAMA4* was based on i) its greater fold-expression increase in *in vivo* selected derivatives relative to *FOXQ1* and *NAP1L3*, and ii) its encoding of a secreted ECM protein that could potentially be therapeutically targeted. Consistent with its extracellular role, we detected greater levels of *LAMA4* protein, laminin-α4, in conditioned media obtained from TE cells relative to their parental populations (Supplementary Fig. 3a-c). We also observed that laminin-α4 was distributed throughout xenograft tumors, with greater levels in tumors derived from TE cells relative to tumors derived from parental cells (Supplementary Fig. 3d,e). *LAMA4*-depletion in a third independent ER-negative breast cancer cell line also led to a significant reduction in macroscopic lung colonization (Supplementary Fig. 4a–b), further supporting the potential importance of this gene in ER-negative breast cancer.

We next asked if *LAMA4* was sufficient to enhance the tumor re-initiation capacity of parental breast cancer populations. Over-expression of *LAMA4* in parental populations significantly increased (4-fold) the number of tumors formed after injection of low numbers of cells (10 cells/injection) into the mammary fat pads of immunodeficient mice (Fig. 4a). In the context of injecting a high number of cells (5×10^5 cells/injection), overexpression or depletion of *LAMA4* did not significantly alter tumor growth, as there were no significant differences in tumor size (Fig. 4b,f), and *LAMA4*-depletion had no effect on the frequency of tumors formed (Supplementary Fig. 4f). Over-expression of *LAMA4* in both the MDA-231 and CN34 parental populations also led to a significant increase in metastatic colonization of the lungs upon tail-vein injection of cancer cells (Fig. 4c, Supplementary Fig. 4c), resulting in a significant increase in the number of macroscopic colonies formed (Fig. 4d, Supplementary Fig. 4d). Taken together, these results demonstrate that over-expression of *LAMA4* in parental populations is sufficient to enhance both their tumorigenic and metastatic potential at limiting cell numbers.

We next assessed the role of endogenous *LAMA4* in TE derivatives and found that *LAMA4*-depletion in TE3 cells significantly decreased the numbers of tumors formed upon orthotopic injection of limiting numbers of cells (10 cells/injection) into the mammary fat pads of immunodeficient mice (Fig. 4e, Supplementary Fig. 4e). *LAMA4*-depletion in TE3 cells also significantly decreased the numbers of macroscopic colonies formed upon ectopic injection into the lungs (Fig. 4g,h, Supplementary Fig. 4g,h). Collectively, these findings reveal that *LAMA4* is sufficient to enhance the tumor-forming and metastatic potential of parental populations, that endogenous expression of *LAMA4* promotes tumor re-initiation in multiple organ microenvironments, and that the impact of *LAMA4* on tumor re-initiation emerges in the physiological context of limiting cell numbers.

LAMA4 promotes cancer cell proliferation in the absence of substratum-attachment *in vitro* in a β1-integrin dependent manner

An important feature of malignant cells is their capacity to proliferate and survive in the absence of attachment to an underlying matrix during multiple stages of cancer

progression[5,32–33]–conditions that can lead to protracted cell-cycle arrest or programmed cell death[34]. We hypothesized that the cell-autonomous expression of *LAMA4* by cancer cells might enable them to proliferate independent of attachment to an underlying substratum. When TE3 cells were seeded at clonal density (one cell per well) into low-attachment plates that prevent substratum-attachment (Supplementary Fig. 5a), *LAMA4*-depleted cells proliferated at a reduced level relative to cells transduced with a control hairpin (Fig. 5a). We next asked if tumorigenic-enriched and highly metastatic derivatives, which physiologically express elevated levels of *LAMA4* relative to their parental populations, might demonstrate an enhanced ability to proliferate in the absence of substratum-attachment. TE3 or LM2 cells seeded at clonal density into low-attachment plates proliferated more extensively than their parental population (Fig. 5b). In addition, after several days a greater fraction of wells seeded with TE3 and LM2 cells contained multiple cells relative to wells seeded with the parental population, which had a greater fraction of wells that still contained only one cell (Supplementary Fig. 5b). This effect was *LAMA4*-dependent, as a smaller fraction of wells containing *LAMA4*-knockdown cells contained multiple cells relative to wells containing control cells (Supplementary Fig. 5c). These observations suggested that *LAMA4* might regulate cell-cycle dynamics during conditions when cells are detached from substratum. Modification of this assay to enable quantitative cell-cycle analysis (Supplementary Fig. 5d) revealed that *LAMA4*-depletion in the absence of substratum-attachment increased the fraction of cells in G0/G1 (Fig. 5c, Supplementary Fig. 5e). TE3 and LM2 populations cultured under the same conditions contained a smaller fraction of cells in G0/G1 relative to their parental populations (Fig. 5d). Taken together, these findings demonstrate that in the absence of substratum-attachment *in vitro*, *LAMA4* represses the fraction of cells in G0/G1 and promotes the proliferation of these breast cancer cells.

We next performed more detailed studies to understand how *LAMA4* impacts cell proliferation. The accumulation of cells in G0/G1 in the absence of substratum attachment *in vitro* suggested that cell cycle inhibitory proteins, such as p21 or p27, might be impacted[35–37]. We found that *LAMA4*-depletion led to increased levels of p27 in the absence of substratum attachment *in vitro* (Fig. 5e, Supplementary Fig. 5l). *LAMA4*–depletion also decreased the fraction of Ki-67-positive TE3 cells in the absence of substratum attachment *in vitro* (Fig. 5f, Supplementary Fig. 5f), consistent with these findings. We next asked if *LAMA4* acts through integrins, which are known to be receptors for laminins[38–41] and have well-described roles in cell proliferation[42–43]. Inhibition of β1-integrin using a blocking antibody decreased the proliferation of TE cells in the absence of substratum attachment *in vitro* (Fig. 5i, Supplementary Fig. 5g,k). Inhibition of an independent integrin, αVβ3-integrin, did not have a significant effect (Fig. 5i). Similar results were found in MDA-parental cells that over-expressed *LAMA4* (Fig. 5g, Supplementary Fig. 5i). Inhibition of β1-integrin also prevented the ability of recombinant *LAMA4*-containing protein (laminin-411) to promote the proliferation of TE3 *LAMA4*-knockdown cells (Fig. 5h, Supplementary Fig. 5h,j). Finally, extension of these findings *in vivo* revealed that inhibition of β1-integrin also prevented the increase in ectopic tumor re-initiation by *LAMA4* over-expressing cells (Fig. 5j,k). These results demonstrate that *LAMA4*'s promotion of cancer cell proliferation and tumor re-initiation requires β1-integrin.

LAMA4 promotes metastatic proliferation and incipient micro-metastasis formation *in vivo*

During metastasis, cancer cells detach from the primary tumor and subsequently enter a foreign microenvironment with a distinct ECM composition acutely devoid of proliferative and survival cues that permit the effective formation of colonies[4,44]. In addition, studies have shown that introduction of cancer cells into the intravenous circulation leads to the pulmonary seeding of mainly solitary cells that initially lack contact with each other[45]. These features represent barriers that can prevent the ability of cancer cells to regain proliferative potential and to re-initiate metastatic colonies, and they share several similarities with the seeding of single cells in non-attachment conditions *in vitro*. Based on our findings that *LAMA4* promotes the proliferation of cancer cells under such *in vitro* conditions, we hypothesized that *LAMA4* may also regulate the initial proliferation of disseminated cancer cells and their subsequent formation of multi-cellular colonies *in vivo*. To address this, we seeded highly metastatic LM1a cells with or without *LAMA4*-knockdown into the lungs of mice through intravenous injection. After a latency period of several days, lung-bioluminescence signal from mice injected with LM1a-control cells began to increase while signal from mice injected with LM1a-*LAMA4*-knockdown cells showed persistently reduced signal (Fig. 6a). Measurement of cancer cell-derived caspase-3/7 activity through an *in vivo* reporter during this same period revealed no significant differences between animals injected with control cells relative to those injected with *LAMA4*-knockdown cells, demonstrating that the differences we observed in early metastatic outgrowth were unlikely to be secondary to apoptosis (Supplementary Fig. 6a). Rather, these observations were consistent with *LAMA4* promoting the proliferation and expansion of incipient cells that had extravasated into the lungs. Subsequent analysis revealed that *LAMA4*-depletion resulted in a significant decrease in the fraction of multi-cellular colonies relative to single cells present in the lungs of mice several days after tail-vein injection (Fig. 6c), demonstrating that *LAMA4* promotes the formation of micro-metastases *in vivo*. *LAMA4*-depletion also reduced the percentage of solitary cells that stained positively for the proliferative marker Ki67 (Fig. 6d), demonstrating that a greater fraction of solitary *LAMA4*-knockdown cells were not actively proliferating. Furthermore, knockdown of *LAMA4* also decreased the average size of metastatic colonies formed (Fig. 6b, Supplementary Fig. 6b), consistent with the continued proliferation of control cells giving rise to larger colonies relative to knockdown cells. An independent experiment employing the highly metastatic LM2 cell line confirmed these observations (Fig. 6e–h, Supplementary Fig. 6c). Collectively, these findings demonstrate that *LAMA4* promotes the active proliferation of disseminated solitary cells and the formation and expansion of incipient multicellular micro-metastases *in vivo*.

FOXQ1 promotes the expression of *LAMA4*

Having demonstrated a functional role for *LAMA4* in promoting primary and metastatic tumor reinitiation, we next asked if its expression level is influenced by either *FOXQ1* or *NAP1L3*, two other genes expressed at greater levels in tumorigenic-enriched and highly metastatic derivatives that also promote metastatic efficiency. Depletion of *FOXQ1* in multiple cell lines decreased *LAMA4* levels (Fig. 7a,b), while depletion of *LAMA4* in multiple cell lines decreased *FOXQ1* levels (Fig. 7c,d), revealing a positive reciprocal

relationship between these genes. This relationship was further supported through interrogation of a large set of primary breast tumor samples (*n*=988; see Methods), which revealed that *LAMA4* expression positively correlates with *FOXQ1* expression (Fig. 7e). *NAP1L3* knockdown in multiple cell lines did not reduce *LAMA4* expression (Supplementary Fig. 7a,b), and *LAMA4* knockdown did not reduce *NAP1L3* expression (Fig. 7c,d). These results demonstrate that *LAMA4* expression co-varies with *FOXQ1* in primary human tumors and reveal a positive reciprocal relationship between *FOXQ1* and *LAMA4* expression.

Increased expression of *LAMA4* marks early breast cancer progression and is correlated with clinical relapse

Given the increased expression of *LAMA4* by tumorigenic-enriched derivatives and its sufficiency in promoting tumor re-initiation *in vivo*, we next investigated the role of *LAMA4* in human clinical samples. As breast cancers transition from the pre-malignant stage (ductal carcinoma *in situ*) to the malignant stage (frank carcinoma), cancer cells acquire the ability to proliferate and survive in the absence of proper engagement to the ECM as they pass through the basement-membrane and enter the surrounding stromal microenvironment[46]. Having demonstrated a role for *LAMA4* in driving the expansion of breast cancer cells in the settings of non-attachment *in vitro* and multiple microenvironments *in vivo*, we asked whether *LAMA4* expression was increased during the transition from pre-malignancy to malignancy. Examination of multiple transcriptomic datasets[47–50] of laser-capture micro-dissected breast cancer tissues from pre-malignant or malignant areas of disease from within the same patients (Fig. 8a, see Methods) revealed that *LAMA4* expression was significantly increased in malignant breast cancer cells relative to nearby pre-malignant cancer cells (Fig. 8b–e; $P<0.0001$ combined). These findings are consistent with the increased expression of *LAMA4* in populations of cells enriched for tumor-forming potential and support our functional studies demonstrating a role for *LAMA4* in promoting breast cancer tumor re-initiation.

We next investigated whether *LAMA4* expression in established tumors correlates with relapse of human breast cancer. Analysis of multiple independent datasets[51–54] revealed that when patients with ER-negative breast cancer were stratified into those whose tumors expressed high or low levels of *LAMA4*, patients whose tumors expressed high levels of *LAMA4* had significantly reduced relapse-free survival relative to patients whose tumors expressed low levels of *LAMA4* (Fig. 8f–i, $P<0.0005$ combined). Tumors from patients that relapsed also expressed higher levels of *LAMA4* (Fig. 8k) and increased expression of *LAMA4* was associated with shorter overall survival in an independent cohort of ER-negative patients (Fig. 8j). Collectively, our selection and characterization of cancer cell populations with enhanced tumorigenic capacity establishes a model to study molecular and cellular determinants of tumor-forming potential, and has led to the identification of *LAMA4* as a key promoter of tumor re-initiation in the primary and metastatic microenvironments, whose expression correlates with human breast tumorigenesis and clinical outcome (Fig. 8l).

DISCUSSION

A molecular and cellular understanding of the features that govern tumor-forming potential is of great interest to the scientific and biomedical communities[6]. We herein describe a strategy to elucidate genes and cellular features that govern tumor re-initiation in ER-negative breast cancer using an unbiased approach to enrich for populations of cells with enhanced tumorigenic capacity (Fig. 1). Tumorigenic-enriched (TE) populations derived through *in vivo* selection demonstrated enhanced tumor re-initiation in multiple micro-environments (Fig. 1,2) and expressed a set of genes (*LAMA4*, *FOXQ1*, and *NAP1L3*) at greater levels relative to their parental populations (Fig. 3). These genes were also expressed at greater levels by highly metastatic cells and each of these genes promoted metastatic efficiency *in vivo* (Fig. 3). Further characterization of *LAMA4* revealed it to promote tumor re-initiation in the primary and metastatic micro-environment (Fig. 4) and to enhance the proliferation and multicellular expansion of cancer cells in the absence of substratum-attachment *in vitro* (Fig. 5) and during colonization *in vivo* (Fig. 6). Mechanistically, *LAMA4*'s promotion of cancer cell proliferation and tumor re-initiation requires β1-integrin (Fig. 5). In patient samples, *LAMA4* expression is increased in malignant cancer cells relative to adjacent pre-malignant cancer cells (Fig. 8) and increased *LAMA4* expression is correlated with reduced relapse-free survival (Fig. 8). Our unbiased application of *in vivo* selection to derive sub-populations of cells with enhanced tumorigenic capacity has enabled the identification of cellular and molecular determinants of tumor re-initiation in breast cancer and revealed *LAMA4* as a gene that promotes tumor re-initiation in multiple microenvironments and whose increased expression is associated with human breast cancer initiation and progression (Fig. 8).

It is thought that cancer cell interactions with the microenvironment during multiple stages of cancer progression can enable cells to resume cell-cycle progression and expand into colonies when they lack important extracellular cues[34,44,55–56] that can result in proliferative suppression[57] or cellular quiescence[34]. Our findings that the cell-autonomous expression of *LAMA4* promotes the proliferation of cancer cells in the absence of substratum-attachment *in vitro* and during incipient metastatic outgrowth *in vivo* suggest that *LAMA4* antagonizes the suppressive effects these contexts can have on proliferative potential. *LAMA4*'s promotion of tumor re-initiation in orthotopic and ectopic settings is also consistent with this gene governing a general feature of tumor reinitiation that is independent of a particular microenvironment.

LAMA4 is a member of a family of laminin genes encoding proteins normally present in the extracellular matrix (ECM)[58]. The roles of various laminins is thought to be restricted to their deposition into the BM and ECM, which is thought to enable cancer progression in part by allowing cancer cells to anchor to these surfaces in order to survive and proliferate[59–61]. *LAMA4* is distinct from most laminin alpha-chains in that it lacks an N-terminal 'head' region that is thought to allow laminins to polymerize *de novo*[62]. Studies have revealed that *LAMA4* requires enzymatic cross-linking in order to be incorporated into ECM generated *in vitro*[63] and have proposed a de-adhesive role for *LAMA4* in the progression of other solid cancers[64]. The cell-autonomous expression by cancer cells of this specific laminin gene lacking an anchoring head region provides a potential mechanism for cells to decouple

proliferative capacity from the physiologically constrained requirement of attachment to an immobile extracellular matrix, allowing them to in effect ‘carry’ their own ECM. Future studies could determine the potential contributions of other laminin family members that cooperate with *LAMA4* in mediating its effects.

In women, breast cancer leads all other cancers in the numbers of annual diagnoses and deaths incurred globally[65]. While aggressive cancers can lead to rapid death, metastatic relapse can occur even after the surgical resection of the primary tumor and without any prior radiographic evidence of metastases[44]. Such relapse is thought to be caused by a small number of disseminated cells that have re-activated proliferative potential in an inhospitable microenvironment[11,34,57]. Our identification of *LAMA4* as a gene expressed by ER-negative breast cancer cells that promotes tumor re-initiation in the metastatic microenvironment reveals a candidate target for therapeutic intervention. Functional blocking of *LAMA4* protein could potentially prevent the re-activation and subsequent outgrowth of disseminated metastatic cells. Our identification of this gene as a potential therapeutic target in ER-negative breast cancer has important implications for patients diagnosed with ER-negative subtypes, given that they are not candidates for treatment with anti-estrogen therapies[66].

METHODS

Animal studies

All mouse experiments were conducted in accordance with a protocol approved by the Institutional Animal Care and Use Committee (IACUC) at The Rockefeller University. Six- to twelve-week-old age-matched female NOD scid or NOD scid gamma mice obtained from Jackson Labs were used for orthotopic mammary fat pad, experimental lung metastasis, ectopic direct-lung, and experimental liver metastasis assays. For bioluminescence tracking of cells *in vivo*, cells were labeled with a triple-fusion reporter protein construct[67] through retroviral transduction followed by FACS to isolate GFP positive transduced cells 48–72 hours later. Non-invasive bioluminescence imaging was performed by anesthetizing mice with Isoflurane (Butler Schein), retro-orbital injection of D-Luciferin Luciferase (PerkinElmer) and exposing mice in an IVIS ® Lumina II (Caliper Life Science). Quantification of signal was performed with Living Image (PerkinElmer) software. *In vivo* caspase activity was measured by retro-orbital injections of VivoGlo™ Caspase 3/7 Substrate (Z-DEVD-Aminoluciferine Sodium Salt) (Promega) and bioluminescence signal was normalized to cancer cell luciferase signal[68]. Tumor volume was determined by measuring the small (s) and large (l) diameter of tumors using an Electronic Digital Caliper (Fisher Scientific) and quantifying volume using the formula $\pi s^2 l/6$.

In vivo selection

To generate tumorigenic-enriched (TE) derivatives from the MDA-MB-231 and CN34 human breast cancer cell lines, moderate numbers of parental populations (1×10^4 MDA-231-parental or 2×10^4 CN34-parental cells) were mixed in a 1:1 ratio of PBS and growth-factor reduced matrigel (356231, BD Biosciences) and injected orthotopically and bilaterally into the 2nd and 4th mammary glands of NOD scid mice. Tumors that were generated were dissociated into single cells (see below) and propagated *in vitro* to yield 1st

generation tumorigenic-enriched (TE1) derivatives. TE1 cells were then subjected to another round of *in vivo* selection by injecting 10-fold less (1×10^3 MDA-TE1 or 2×10^3 CN34-TE1) cells into the mammary fat pads of NOD scid mice, giving rise to a second round of tumors that were dissociated into single cells and propagated *in vitro* to yield 2nd generation tumorigenic-enriched (TE2) derivatives. For the MDA-231 cell line, a third round of *in vivo* selection was performed by injecting MDA-TE2 cells at a dose 10-fold less (1×10^2 cells) into the mammary fat pads of NOD scid mice to give rise to a third round of tumors that were then dissociated into single cells and propagated *in vitro* to yield 3rd generation tumorigenic-enriched (TE3) derivatives. CN34-TE2 cells did not undergo a third round of *in vivo* selection.

Tumor dissociation

For tumor dissociation, tumors were excised, placed in a 6cm tissue-culture dish and minced into fine pieces using a scalpel. Cells were collected in a 50mL conical tube by washing with PBS using a 25mL pipette and spun down. The pellet was re-suspended in 5mL of ACK buffer (Cambrex) and incubated at room temperature for 10 min for lysis of red blood cells. The mixture was washed with PBS, spun down, and re-suspended in 10mL of dissociation media comprised of a 1:1 mixture of DMEM:F12 supplemented with penicillin, streptomycin and fungizone, 1.25mg/mL Collagenase Type I (Worthington Biochemical) and 1mg/mL Hyaluronidase (Worthington Biochemical). The mixture was placed in shaking incubator at 37°C. Every 30 min, the mixture was removed and pipetted up and down with a 25mL pipette for a maximum of 2 hr until the mixture had little or no large pieces remaining. After dissociation, PBS was added to the mixture and it was spun down for 10 min. The pellet was re-suspended in 7mL of 0.25% Trypsin-EDTA (Invitrogen) and placed in shaking incubator at 37°C. After 10 min, 30mL of 10% FBS-containing DMEM was added to trypsin-neutralize and the solution was spun down at 3,000 rpm for 10 min. The pellet, now mainly a fine suspension, was re-suspended in 10mL of a 1:1 mixture of DMEM:F12 supplemented with penicillin, streptomycin and fungizone, 1mg/mL BSA, 25mM HEPES, and 20,000 Units/L DNase I (Worthington Biochemical) and placed in shaking incubator at 37°C. After 10 min, the pellet was re-suspended in 10mL of PBS and spun down. The final suspension was filtered through a 70μm and then 40μm filter to enrich for a single cell suspension of cells. The resulting cells were plated onto adherent tissue-culture plates and grown in appropriate media.

Orthotopic and ectopic tumor re-initiation assays

For orthotopic tumor re-initiation assays, cancer cells were mixed in a 1:1 ratio of PBS and growth-factor reduced matrigel (356231, BD Biosciences) and injected bilaterally into the 2nd and 4th mammary glands of age-matched NOD scid or NOD scid gamma mice. After 10 weeks, mice were sacrificed and necropsy was performed. Absence of tumor re-initiation was concluded if there was no evidence of tumor growth through visual inspection and palpation upon detailed necropsy. Any visual or palpable evidence of tumor growth was considered a re-initiating event and counted as positive tumor formation. For quantification, *n* was the number of independent mammary fat pad injections sites (4 per mouse). For ectopic tumor re-initiation assays, cancer cells were mixed in a 1:1 ratio of PBS and growth-factor reduced matrigel (356231, BD Biosciences) and injected into the 3rd intercostal space,

3mm deep, directly into the lung parenchyma of age-matched NOD scid or NOD scid gamma mice. For the ectopic tumor re-initiation blocking antibody experiment, mouse IgG1 isotype control antibody (clone 107.3, BD Pharmingen) or mouse anti-human β1-integrin blocking antibody (clone 6S6, Millipore) was added to the cell/PBS/matrigel mixture at a concentration of 20ug/mL prior to injection. For ectopic tumor reinitiation assays, *n* was the number of independent mice that were injected or the number of independent lungs that were harvested and subjected to histological analysis.

Metastasis assays

For lung metastasis assays, triple-reporter labeled cells were resuspended in 0.1mL PBS and injected using a 27G½ needle (BD) into the lateral tail-vein of age-matched NOD scid mice. For liver metastasis assays, triple-reporter labelled cells were resuspended in PBS and injected using a 27G½ needle (BD) into the portal circulation of NOD scid gamma mice via splenic injection. For both lung and liver metastasis assays, non-invasive bioluminescence imaging was performed immediately after cell inoculations using an IVIS ® Lumina II (Caliper Life Science) to assess a baseline level of injected cells on day 0. For all metastasis assays, *n* was the number of independent mice that were injected or the number of independent lung/livers that were harvested and subjected to histological analysis.

Cell culture

MDA-MB-231 (MDA-231) and CN34 breast cancer cell lines were propagated *in vitro* on standard tissue-culture treated plates. MDA-MB-231 cells and their derivatives were maintained in DMEM supplemented with 10% FBS, glutamine, pyruvate, penicillin, streptomycin and fungizone. CN34 cells and their derivatives were maintained in M199 supplemented with 2.5% FBS, 10ug/mL insulin, 0.5ug/mL hydrocortisone, 20ng/mL EGF, 100ng/mL cholera toxin, glutamine, pyruvate, penicillin, streptomycin and fungizone. The MDA-MB-231 and CN34 parental populations and their lung-metastatic (LM) derivatives MDA-LM2 and CN34-LM1a have been described in previous publications[14,16]. The MDA-MB-231 cell line (ATCC) was derived from the pleural effusion of a patient suffering from metastatic breast cancer[19] and the derivation of the MDA-LM2 subpopulation through *in vivo* selection has been described previously[14]. The CN34 parental population was formerly isolated from the pleural effusion of a patient with metastatic breast cancer treated at MSKCC upon IRB consent as described previously[20] and the CN34-LM1a derivative was formerly generated through inoculation of NOD scid mice with CN34 parental populations and dissociating metastatic nodules that had formed in the lungs as described previously[16]. Both cell lines are Estrogen Receptor-negative (ER-negative)[20]. The independent ER-negative HCC1806 cell line[69] was cultured in RPMI-based media supplemented with 10% FBS, glutamine, pyruvate, penicillin, streptomycin and fungizone. Cells in culture were routinely tested for mycoplasma contamination. For proliferation assays in the absence of substratum-attachment, cells were sorted using a FACSAria II (Becton Dickinson) at a clonal density of one cell per well into Ultra-Low Attachment Surface 96-well plates (Corning) containing EGM-2 media (Lonza). The presence of single cells was confirmed after sorting using bright-field microscopy and the number of cells contained in each well was counted on subsequent days. For quantification of cell proliferation, *n* was the number of individual wells seeded with single cancer cells. For the *in vitro* blocking antibody

experiments, cells were pre-incubated with mouse IgG1 isotype control antibody (clone 107.3, BD Pharmingen), mouse anti-human αVβ3-integrin blocking antibody (clone LM609, Millipore), or mouse anti-human β1-integrin blocking antibody (clone 6S6, Millipore) at 20ug/mL in HBSS for 20 min on ice prior to cell seeding. For the recombinant laminin-411 rescue experiments, cells were re-suspended in HBSS or pre-incubated with mIgG or anti-β1-integrin antibodies at 20ug/mL in HBSS for 20 min on ice, and then an equal volume of solution containing 160nM Bovine Serum Albumin (BSA) or 160nM laminin-411 (BioLamina) was added to the cells and they were incubated for 30 min at 37C prior to cell seeding. For cell cycle and Ki67 analysis in the absence of substratum-attachment, to enable quantitative analysis, a greater number of cells were seeded at a low density of 5,000 cells/well into Ultra-Low Attachment Surface 6-well plates (Corning) in media containing a 1:1 mixture of DMEM supplemented with 10%FBS, glutamine, pyruvate, penicillin, streptomycin and fungizone and 2x DMEM mixed 1:1 with 3% methylcellulose stock solution (HSC001, R&D). Cells were harvested on day 3 and analyzed using an LSR II (Becton Dickinson). Cell cycle analysis was based on DAPI staining. For Ki67 analysis, cells were fixed and permeabilized with BD Cytofix/Cytoperm Fixation/Permeabilization kit (BD Biosciences) per manufacturer's instructions and incubated with mouse anti-human Ki67 antibody (Alexa Fluor® 700-conjugated clone B56: BD Pharmingen) for 20 min on ice. Cell cycle phase and Ki67-positive fractions were determined using FlowJo (TreeStar) software. For flow cytometry immunophenotypic marker analysis, cells were incubated in PBS containing mouse anti-human CD44 antibody (APC-conjugated clone G44-26: BD Pharmingen) and mouse anti-human CD24 antibody (PE-conjugated clone ML5: BD Pharmingen) for 30 minutes on ice. Cells were washed, stained with propidium iodide for live-dead exclusion and analyzed using an LSR II (Becton Dickinson) and FlowJo (TreeStar) software.

Cell proliferation and colony formation assay

For cell proliferation assays, 2.5×10^4 cells were seeded into tissue-culture treated adherent 6-well plates (Falcon). Cells were collected through trypsin digestion and counted on days 1 and 5. Cell counts were normalized to day 1. For colony formation assays, 1×10^2 cells were seeded in triplicate into tissue-culture treated adherent 10cm plates (Falcon). On day 14, plates were washed with PBS, fixed in 6% glutaraldehyde (Sigma-Aldrich), stained with 0.5% crystal violet (Sigma-Aldrich), and washed with H20 to better enable the visualization of colonies.

Endothelial recruitment assay

For endothelial recruitment assays, 5×10^4 cancer cells were seeded overnight onto tissue-culture treated adherent 24-well plates (Falcon) while human umbilical vein endothelial cells (HUVECs; Lonza) at 70–80% confluence were serum-starved overnight. The next day, HUVECs were labelled with CellTracker Red™ CMTPX (Molecular Probes) according to manufacturer's protocol and 5×10^4 HUVECs were seeded onto the top layer of 3um HTS Fluoroblok transwell inserts (BD Falcon) fitted to the top of each well of 24-well plate. 0.5mL of 0.2% FBS EGM-2 (Lonza) was added to the top and bottom of each well. Plates were incubated at 37°C for 16 hours. Upon incubation, the transwell inserts were removed, washed twice with PBS and fixed with 4% paraformaldehyde. The inserts were then cut out

using a scalpel and mounted onto microscopy slides using VECTASHIELD® Mounting Media with DAPI (Vector Laboratories). For imaging and quantification, the basal side of each insert was imaged via fluorescence using an inverted microscope (Zeiss Axiovert 40 CFL) at 5x and 6–9 images per field per insert were collected. Quantification of migrated HUVECs was performed using ImageJ (NIH) software http://imagej.nih.gov/ij/.

Cell attachment assay

The cell attachment assay was performed by labelling cancer cells with CellTracker Red™ CMTPX (Molecular Probes) according to manufacturer's protocol and seeding 1×10^5 cells per well of a standard tissue culture six-well plate. After 18 hours, each well was carefully washed with PBS and the numbers of cells that had attached cells to the tissue culture plates was quantified. Quantification was performed by obtaining fluorescent images from nine random fields per well, subtracting background signal, and measuring the Area Fraction (the fractional area of fluorescence signal that covered each field) using ImageJ (NIH) software.

Generation of lentivirus-mediated knockdown cells

For generation of knockdown cell lines, virus was generated using 293T cells that had been grown to 70–80% confluence. Cells were transfected with Lipofectamine ® 2000 in antibiotic-free media and 6ug vector A, 12ug vector K, and 12ug of pLKO (with blasticidin or puromycin selection marker) shRNA vector. Virus was collected after 48–72hr, spun down at 2,000 rpm, and filtered using 0.45um nylon mesh. For cell transduction, virus was added to cancer cells along with polybrene (Millipore) and cells were incubated for 5–6 hrs. After at least 48hr, cells underwent antibiotic selection using either 2ug/mL puromycin or 1ug/mL blasticidin. Cells were removed from antibiotic selection once there were no viable cells remaining on a plate that was mock transduced. shRNA sequences are listed in Supplementary Table 4.

Generation of retrovirus-mediated over-expressing cells

For generation of over-expression cell lines, virus was generated using 293T cells that had been grown to 70–80% confluence. Cells were transfected with Lipofectamine ® 2000 in antibiotic-free media and 12ug vector gag/pol, 6ug vector VSVG, and 12ug of pBABE-puro empty vector or vector containing cloned full-length *LAMA4*. Virus was collected after 48–72hr, spun down at 2,000 rpm, and filtered using 0.45um nylon mesh. For cell transduction, virus was added to cancer cells along with polybrene (Millipore) for 24 hrs. After at least 48hr, cells underwent antibiotic selection using 2ug/mL puromycin. Cells were removed from antibiotic selection once there were no viable cells left on a kill plate that was mock transduced. Primer sequences used to clone full-length *LAMA4* are listed in Supplementary Table 4.

RNA extraction and quantification of mRNA expression

Total RNA from cells was extracted and purified using the MiRvana (Applied Biosystems) or Total RNA Purification Kit (Norgen Biotek) according to manufacturer's protocol. For mRNA quantification, 0.2–2ug of total RNA was subject to reverse transcription using the cDNA First-Strand Synthesis Kit (Invitrogen). The resulting cDNA was diluted 1:2 or 1:5

and mixed with SYBR ® green PCR Master Mix (Applied Biosystems) with primers specific to each gene. Each independent reaction was split into quadruplicate wells of a 384-well plate and loaded onto an ABI Prism 7900HT Real-Time PCR System (Applied Biosystems) to conduct quantitative real-time PCR (qRT-PCR). Normalization for relative expression of target genes was performed using *HPRT*, *GAPDH*, or *SMAD4* as endogenous controls. Primer sequences are listed in Supplementary Table 4

siRNA knockdown experiments

For siRNA knockdown experiments, cells were seeded into tissue culture plates overnight. The following day, cells were washed with PBS, and incubated in Opti-MEM ® (Invitrogen) reduced-serum media. Lipofectamine ® 2000 (1:40, Invitrogen) and 50nM of each specific siRNA were pre-mixed and incubated for 20 min before being added to the cells. After addition of the Lipofectamine/siRNA mixture, the cells were incubated for 5 hours, washed with PBS, and normal cell culture media was added. For mRNA analysis, cells were collected and RNA was harvested 24–30hr post-transfection. For the HCC1806 *in vivo* metastasis experiment, cells were injected 48hr post-transfection. Control and target gene siRNAs were obtained from Integrated DNA Technologies (IDT) or Dharmacon (Thermo Fisher Scientific). siRNA sequences are listed in Supplementary Table 4

Immunohistochemistry and histology

Mammary tumors, whole lungs, or livers were excised and fixed through immersion in 4% para-formaldehyde overnight, and subsequently washed with PBS, 50% Ethanol, 70% Ethanol prior to being embedded in paraffin. Paraffin blocks were cut into 5μm thick sections. For macroscopic tumor quantification, paraffin sections were stained with mouse primary antibodies against human Vimentin (1:100, clone V9, Vector Laboratories) and visualized with Vectastain ABC kit (PK-6012, Vector Laboratories) and DAB chromogen (SK-4105, Vector Laboratories); macroscopic nodules were counted on the basis of staining for Vimentin. For immunofluorescence, paraffin sections underwent antigen retrieval and were stained with primary antibodies against human Vimentin (1:100, Vector Laboratories) and Ki-67 (1:200, ab15580, abcam), or for detection of endogenous laminin-α4, the mouse anti-human laminin-α4 antibody (1:50, clone 839084, R&D) was used and the isotype-matched mouse IgG1 antibody (1:50, clone 11711, R&D) was used as a control at an equivalent concentration. For detection of Collagen-IV, the rabbit anti-human Collagen-IV antibody (1:50, ab21295, Abcam) was used. Primary antibodies were detected using Alexa Flour dye-conjugated secondary antibodies (1:200, Invitrogen). Fluorescence was obtained using a Leica laser scanning confocal microscope (TCS SP5). Mouse on Mouse (M.O.M.) Blocking Reagent (Vector Laboratories) was use with all mouse antibodies. For solitary cell/micro-met comparisons, unprocessed images were inspected with ImageJ (NIH) software, and solitary cancer cells and multi-cellular micro-metastatic colonies were randomly counted in 10 fields per section or whole sections on the basis of positive staining for vimentin. Counts were quantified as percent of the total number of metastatic events (solitary cells + micro-metastatic colonies) to normalize for any differences in the initial seeding of metastatic cells. To compare colony size, the average colony size was quantified using ImageJ (NIH) software. To assess the proliferation of solitary cancer cells, the fraction of solitary cancer cells that stained positively for Ki-67 was quantified. This was achieved

by merging unprocessed green (vimentin) and red (Ki-67) channels using ImageJ (NIH) software; the fraction of vimentin-positive solitary cancer cells that were Ki-67-positive was quantified. *n* was the number of independent lungs that were extracted from mice.

Western blotting

For detection of laminin-α4, conditioned media was prepared by concentrating serum-free supernatant obtained from cultured cancer cells incubated for 24hr using 100k cutoff Amicon Ultra (Millipore) centrifuge tubes. For detection of p27, p21, and β-tubulin, cell lysates were collected from cells that had been cultured overnight in Ultra-Low Attachment Surface 6-well plates (Corning). Protein from conditioned media or cell lysates were separated using SDS–polyacrylamide gel electrophoresis, transferred to Immobilon-P Transfer Membrane (Millipore), and probed with antibodies against human laminin-α4 (1:200; clone 6C3, sc-130541, Santa Cruz Biotechnology) or p27 (1:1,000; clone D69C12, Cell Signaling Technology), p21 (1:1,000; clone 12D1, Cell Signaling Technology), or β-tubulin (1:5,000; clone 9F3, Cell Signaling Technology). Primary antibody was chemiluminescently detected using horseradish peroxidase–conjugated secondary antibody (1:10,000), ECL 2 Western Blotting Substrate (Pierce) and the SRX-101A (Konica Minolta) developer. Quantification of western blots was performed using ImageJ (NIH) software. p27 signal intensity was normalized to β-tubulin signal intensity.

Microarray hybridization and transcriptomic analysis of *in vivo* selected derivatives

To identify mRNAs whose levels were increased across tumorigenic-enriched (TE) and lung-metastatic (LM) *in vivo* selected derivatives as compared to their parental populations, total RNA derived from MDA-MB-231 (Parental, TE3 LM2) and CN34 (Parental, TE2, LM1a) populations was extracted, labeled and hybridized onto Illumina HT-12 v3 Expression BeadChip arrays by The Rockefeller University Genomics Resource Center. The raw signal intensities for each probe were median-normalized and replicates were averaged. Candidate promoters of tumor re-initiation were identified by over-lapping the set of genes whose expression was at least 1.5x fold-increased in the following conditions: i) MDA-TE3 vs. MDA-Parental, ii) MDA-LM2 vs. MDA-Parental, iii) CN34-TE2 vs. CN34-Parental, iv) CN34-LM1a vs. CN34-Parental. Eight candidate genes passed these criteria, and three of these genes demonstrated statistically significant differences upon independent validation using quantitative real-time PCR (qRT-PCR).

Correlation coefficient analysis

Ranked expression values for *LAMA4* and *FOXQ1* in primary breast tumors from 988 patients from The Cancer Genome Atlas (TCGA) Breast Cancer project provisional dataset were downloaded from http://www.cBioPortal.org/ and subjected to a linear regression analysis. *P* values were based on Spearman's coefficient test.

mRNA analysis of micro-dissected patient samples

For comparison of pre-malignant and malignant breast cancer tissues, the following datasets were used: Schuetz 2006 (GSE3893)[50], Lee 2012 (GSE41197)[47], Ma 2003: (Table 4)[48], Ma 2009 (GSE14548)[49]. Each dataset included patient-matched laser capture micro-dissected

epithelial cancer tissue (to restrict the analysis only to cancer cells) from regions of pre-malignant (ductal carcinoma *in situ*) or malignant (invasive ductal carcinoma) disease. Probe intensity values for *LAMA4* were used to quantify mRNA expression levels (multiple probe values were averaged) in each dataset, intra-sample normalized, and subject to a paired t-test to determine statistical significance and the average fold-change across individual or combined datasets.

Clinical Correlation Analysis

The expression of *LAMA4* in ER-negative tumors was assessed in the following datasets: Wang (GSE2034)[54], Desmedt (GSE7390)[51], Hatzis (GSE25066)[52], NKI[53], and The Cancer Genome Atlas (TCGA) Breast Cancer Project provisional dataset (http://www.cBioPortal.org/). To generate Kaplan-Meier curves, for each dataset patients were stratified according to those patients whose expression of *LAMA4* was higher (*LAMA4* High) or lower (*LAMA4* Low) than the median value for all ER-negative tumor samples within that dataset. ER-negativity was assessed from the clinical annotations provided with each dataset. Relapse-free or overall survival outcome was censored at 5 years, plotted and subjected to a one-tailed Mantel-Cox test. For Affymetrix chips with multiple probes, only the JetSet[70] probe (202202_s_at) was used. For comparison of *LAMA4* expression in tumors derived from patients that did, or did not relapse in the combined Wang/Hatzis/Desmedt/NKI dataset, *LAMA4* expression values were converted into z-scores within each dataset before being separated into Relapse or Relapse-Free groups and the data was uncensored.

Statistical Analysis

Throughout all figures: center values represent mean, error bars represent + or ± SEM; $*P<0.05$, $**P<0.01$, and $***P<0.001$ unless otherwise noted. Statistical significance was concluded at $P<0.05$. For animal studies, the investigator was not blinded during group allocation during the experiment or when assessing outcome, no statistical method was used to predetermine sample size, and the experiments were not randomized. The statistical tests used to determine *P* values are listed in each Figure Legend. Samples whose values were greater than two standard deviations from the mean were considered outliers and excluded from the analysis. Data analyzed using an unpaired Student's t-test were under the assumption of normality and equal variance unless otherwise specified. For non-normal data, analysis was performed using a Mann-Whitney test; normality was assessed using a Shapiro-Wilk test. If there was a significant difference in variance, a Mann-Whitney test or Welch's correction was applied; variance was determined using an F-test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to members of the Tavazoie laboratory for insightful discussion and C. Alarcon, N. Halberg, N. Pencheva, and A. Nguyen for providing comments on previous versions of this manuscript. We thank C. Blobel, S. Simon, and J. Friedman for intellectual input and helpful suggestions. We thank A. Nguyen and J. Min Loo for assistance with splenic injections. We thank N. Halberg for assistance with tail-vein injections. We thank H.

Goodarzi for assistance with statistical analysis. We thank P. Furlow for cloning assistance. We thank C. Zhao of the Rockefeller Genomics Resource Center for assistance with transcriptomic profiling. We thank S. Mazel and members of the Rockefeller Flow Cytometry Resource Center for FACS sorting. J.B.R., D.H., and L.B.N. are members of the Weill Cornell/Rockefeller/Sloan-Kettering Tri-Institutional MD-PhD Program supported by NIH MSTP grant GM07739. S.F.T. is a DOD Era of Hope Scholar and Collaborative Scholars and Innovators Award recipient and Head of the Elizabeth and Vincent Meyer Laboratory of Systems Cancer Biology.

References

1. Klein CA. Parallel progression of primary tumours and metastases. Nature reviews. Cancer. 2009; 9:302–312.
2. Maheswaran S, Haber DA. Circulating tumor cells: a window into cancer biology and metastasis. Current opinion in genetics & development. 2010; 20:96–99. [PubMed: 20071161]
3. Chiang AC, Massague J. Molecular basis of metastasis. The New England journal of medicine. 2008; 359:2814–2823. [PubMed: 19109576]
4. Fidler IJ. The pathogenesis of cancer metastasis: the ‘seed and soil’ hypothesis revisited. Nature reviews. Cancer. 2003; 3:453–458.
5. Pavlova NN, et al. A role for PVRL4-driven cell-cell interactions in tumorigenesis. eLife. 2013; 2:e00358. [PubMed: 23682311]
6. Williams SA, Anderson WC, Santaguida MT, Dylla SJ. Patient-derived xenografts, the cancer stem cell paradigm, and cancer pathobiology in the 21st century. Laboratory investigation; a journal of technical methods and pathology. 2013; 93:970–982.
7. Shimono Y, et al. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell. 2009; 138:592–603. [PubMed: 19665978]
8. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. Cell stem cell. 2012; 10:717–728. [PubMed: 22704512]
9. Chaffer CL, et al. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell. 2013; 154:61–74. [PubMed: 23827675]
10. Magee JA, Piskounova E, Morrison SJ. Cancer stem cells: impact, heterogeneity, and uncertainty. Cancer cell. 2012; 21:283–296. [PubMed: 22439924]
11. Vanharanta S, Massague J. Origins of metastatic traits. Cancer cell. 2013; 24:410–421. [PubMed: 24135279]
12. Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. Nature. 2000; 406:532–535. [PubMed: 10952316]
13. Kang Y, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer cell. 2003; 3:537–549. [PubMed: 12842083]
14. Minn AJ, et al. Genes that mediate breast cancer metastasis to lung. Nature. 2005; 436:518–524. [PubMed: 16049480]
15. Pencheva N, et al. Convergent multi-miRNA targeting of ApoE drives LRP1/LRP8-dependent melanoma metastasis and angiogenesis. Cell. 2012; 151:1068–1082. [PubMed: 23142051]
16. Tavazoie SF, et al. Endogenous human microRNAs that suppress breast cancer metastasis. Nature. 2008; 451:147–152. [PubMed: 18185580]
17. Png KJ, Halberg N, Yoshida M, Tavazoie SF. A microRNA regulon that mediates endothelial recruitment and metastasis by cancer cells. Nature. 2012; 481:190–194. [PubMed: 22170610]
18. Putti TC, et al. Estrogen receptor-negative breast carcinomas: a review of morphology and immunophenotypical analysis. Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc. 2005; 18:26–35.
19. Cailleau R, Olive M, Cruciger QV. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In vitro. 1978; 14:911–915. [PubMed: 730202]
20. Bos PD, et al. Genes that mediate breast cancer metastasis to the brain. Nature. 2009; 459:1005–1009. [PubMed: 19421193]
21. Rhim AD, et al. EMT and dissemination precede pancreatic tumor formation. Cell. 2012; 148:349–361. [PubMed: 22265420]

22. Diehn M, Majeti R. Metastatic cancer stem cells: an opportunity for improving cancer treatment? Cell stem cell. 2010; 6:502–503. [PubMed: 20569685]
23. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. Nature reviews. Cancer. 2005; 5:744–749. [PubMed: 16148886]
24. Brabletz T. EMT and MET in metastasis: where are the cancer stem cells? Cancer cell. 2012; 22:699–701. [PubMed: 23238008]
25. Padua D, et al. TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. Cell. 2008; 133:66–77. [PubMed: 18394990]
26. Reymond N, d'Agua BB, Ridley AJ. Crossing the endothelial barrier during metastasis. Nature reviews. Cancer. 2013; 13:858–870. [PubMed: 24263189]
27. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. Nature reviews. Cancer. 2009; 9:274–284. [PubMed: 19308067]
28. Stenzel D, et al. Endothelial basement membrane limits tip cell formation by inducing Dll4/Notch signalling in vivo. EMBO reports. 2011; 12:1135–1143. [PubMed: 21979816]
29. Goering W, et al. Impairment of gastric acid secretion and increase of embryonic lethality in Foxq1-deficient mice. Cytogenetic and genome research. 2008; 121:88–95. [PubMed: 18544931]
30. Zhang H, et al. Forkhead transcription factor foxq1 promotes epithelial-mesenchymal transition and breast cancer metastasis. Cancer research. 2011; 71:1292–1301. [PubMed: 21285253]
31. Attia M, et al. Interaction between nucleosome assembly protein 1-like family members. Journal of molecular biology. 2011; 407:647–660. [PubMed: 21333655]
32. Buchstaller J, McKeever PE, Morrison SJ. Tumorigenic cells are common in mouse MPNSTs but their frequency depends upon tumor genotype and assay conditions. Cancer cell. 2012; 21:240–252. [PubMed: 22340596]
33. Grassian AR, Coloff JL, Brugge JS. Extracellular matrix regulation of metabolism and implications for tumorigenesis. Cold Spring Harbor symposia on quantitative biology. 2011; 76:313–324. [PubMed: 22105806]
34. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nature reviews. Cancer. 2007; 7:834–846. [PubMed: 17957189]
35. Polyak K, et al. p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. Genes & development. 1994; 8:9–22. [PubMed: 8288131]
36. Zhu X, Ohtsubo M, Bohmer RM, Roberts JM, Assoian RK. Adhesion-dependent cell cycle progression linked to the expression of cyclin D1, activation of cyclin E-cdk2, and phosphorylation of the retinoblastoma protein. The Journal of cell biology. 1996; 133:391–403. [PubMed: 8609171]
37. Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell. 1993; 75:805–816. [PubMed: 8242751]
38. Geberhiwot T, et al. Laminin-8 (alpha4beta1gamma1) is synthesized by lymphoid cells, promotes lymphocyte migration and costimulates T cell proliferation. Journal of cell science. 2001; 114:423–433. [PubMed: 11148143]
39. Fujiwara H, Kikkawa Y, Sanzen N, Sekiguchi K. Purification and characterization of human laminin-8. Laminin-8 stimulates cell adhesion and migration through alpha3beta1 and alpha6beta1 integrins. The Journal of biological chemistry. 2001; 276:17550–17558. [PubMed: 11278628]
40. Kortesmaa J, Yurchenco P, Tryggvason K. Recombinant laminin-8 (alpha(4)beta(1)gamma(1)). Production, purification,and interactions with integrins. The Journal of biological chemistry. 2000; 275:14853–14859. [PubMed: 10809728]
41. Hohenester E, Yurchenco PD. Laminins in basement membrane assembly. Cell adhesion & migration. 2013; 7:56–63. [PubMed: 23076216]
42. Moreno-Layseca P, Streuli CH. Signalling pathways linking integrins with cell cycle progression. Matrix biology: journal of the International Society for Matrix Biology. 2014; 34:144–153. [PubMed: 24184828]
43. Miranti CK, Brugge JS. Sensing the environment: a historical perspective on integrin signal transduction. Nature cell biology. 2002; 4:E83–90. [PubMed: 11944041]

44. Giancotti FG. Mechanisms governing metastatic dormancy and reactivation. Cell. 2013; 155:750–764. [PubMed: 24209616]
45. Shibue T, Weinberg RA. Integrin beta1-focal adhesion kinase signaling directs the proliferation of metastatic cancer cells disseminated in the lungs. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:10290–10295. [PubMed: 19502425]
46. Espina V, Liotta LA. What is the malignant nature of human ductal carcinoma in situ? Nature reviews. Cancer. 2011; 11:68–75.
47. Lee S, et al. Differentially expressed genes regulating the progression of ductal carcinoma in situ to invasive breast cancer. Cancer research. 2012; 72:4574–4586. [PubMed: 22751464]
48. Ma XJ, et al. Gene expression profiles of human breast cancer progression. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:5974–5979. [PubMed: 12714683]
49. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. Breast cancer research: BCR. 2009; 11:R7. [PubMed: 19187537]
50. Schuetz CS, et al. Progression-specific genes identified by expression profiling of matched ductal carcinomas in situ and invasive breast tumors, combining laser capture microdissection and oligonucleotide microarray analysis. Cancer research. 2006; 66:5278–5286. [PubMed: 16707453]
51. Desmedt C, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clinical cancer research: an official journal of the American Association for Cancer Research. 2007; 13:3207–3214. [PubMed: 17545524]
52. Hatzis C, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA: the journal of the American Medical Association. 2011; 305:1873–1881. [PubMed: 21558518]
53. van de Vijver MJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002; 347:1999–2009. [PubMed: 12490681]
54. Wang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet. 2005; 365:671–679. [PubMed: 15721472]
55. Peduto L, et al. ADAM12 is highly expressed in carcinoma-associated stroma and is required for mouse prostate tumor progression. Oncogene. 2006; 25:5462–5466. [PubMed: 16607276]
56. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. Cell. 2010; 141:52–67. [PubMed: 20371345]
57. Goss PE, Chambers AF. Does tumour dormancy offer a therapeutic target? Nature reviews. Cancer. 2010; 10:871–877. [PubMed: 21048784]
58. Nelson CM, Bissell MJ. Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. Annual review of cell and developmental biology. 2006; 22:287–309.
59. Kusuma N, et al. Integrin-dependent response to laminin-511 regulates breast tumor cell invasion and metastasis. International journal of cancer. Journal international du cancer. 2012; 130:555–566. [PubMed: 21387294]
60. Yurchenco PD. Basement membranes: cell scaffoldings and signaling platforms. Cold Spring Harbor perspectives in biology. 2011; 3
61. Fujita M, et al. Overexpression of beta1-chain-containing laminins in capillary basement membranes of human breast cancer and its metastases. Breast cancer research: BCR. 2005; 7:R411–421. [PubMed: 15987446]
62. Hamill KJ, Kligys K, Hopkinson SB, Jones JC. Laminin deposition in the extracellular matrix: a complex picture emerges. Journal of cell science. 2009; 122:4409–4417. [PubMed: 19955338]
63. Ilani T, et al. A secreted disulfide catalyst controls extracellular matrix composition and function. Science. 2013; 341:74–76. [PubMed: 23704371]
64. Vainionpaa N, Lehto VP, Tryggvason K, Virtanen I. Alpha4 chain laminins are widely expressed in renal cell carcinomas and have a de-adhesive function. Laboratory investigation; a journal of technical methods and pathology. 2007; 87:780–791.

65. Jemal A, et al. Global cancer statistics. CA: a cancer journal for clinicians. 2011; 61:69–90. [PubMed: 21296855]

66. Ali S, Coombes RC. Endocrine-responsive breast cancer and strategies for combating resistance. Nature reviews. Cancer. 2002; 2:101–112. [PubMed: 12635173]

67. Ponomarev V, et al. A novel triple-modality reporter gene for whole-body fluorescent, bioluminescent, and nuclear noninvasive imaging. European journal of nuclear medicine and molecular imaging. 2004; 31:740–751. [PubMed: 15014901]

68. Hickson J, et al. Noninvasive molecular imaging of apoptosis *in vivo* using a modified firefly luciferase substrate, Z-DEVD-aminoluciferin. Cell Death Differ. 2010; 17:1003–1010. [PubMed: 20057500]

69. Gazdar AF, et al. Characterization of paired tumor and non-tumor cell lines established from patients with breast cancer. International journal of cancer. Journal international du cancer. 1998; 78:766–774. [PubMed: 9833771]

70. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene. BMC bioinformatics. 2011; 12:474. [PubMed: 22172014]

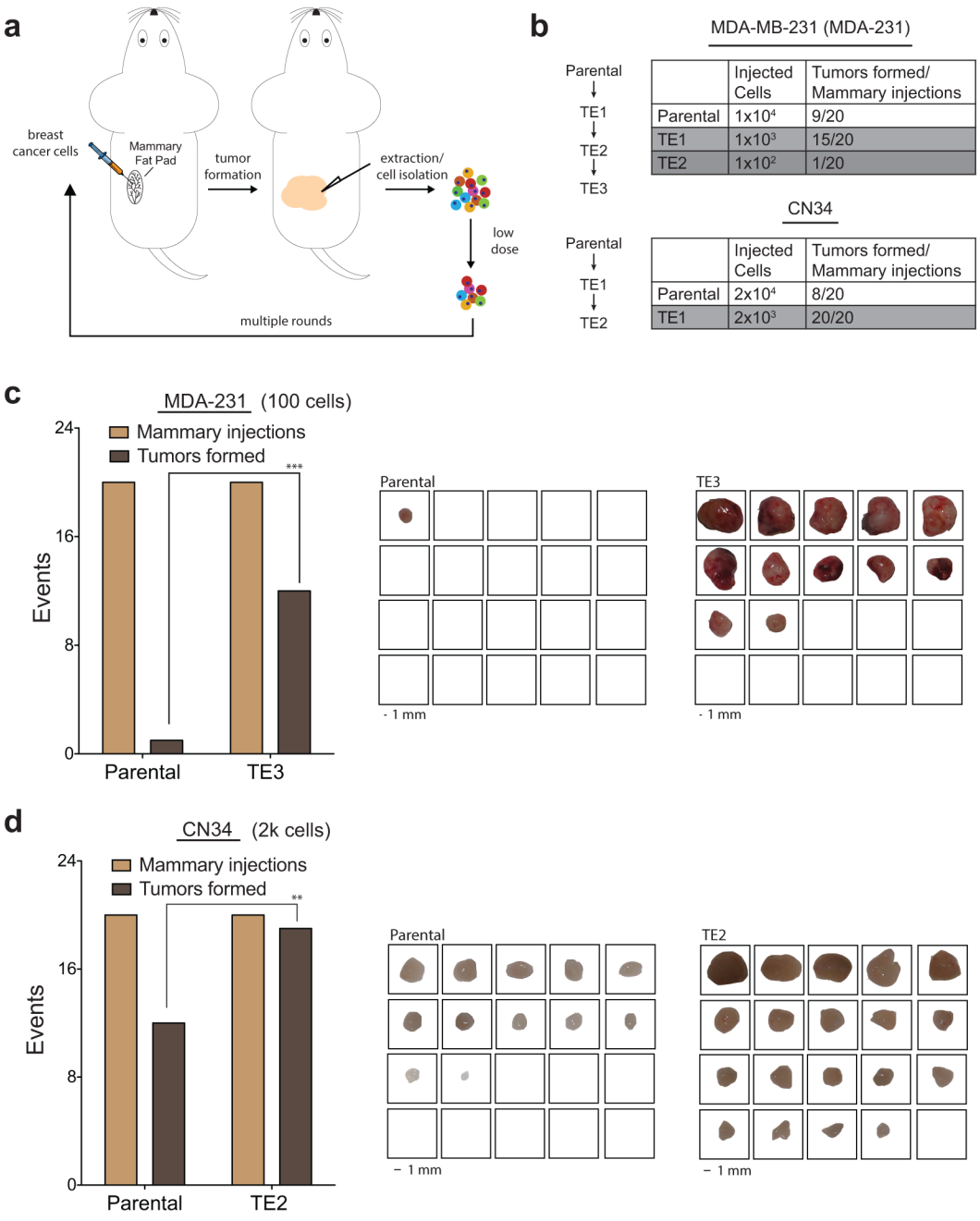

Figure 1. *In vivo* selection for tumor re-initiation

(**a**) Schematic of the *in vivo* selection strategy used to derive tumorigenic-enriched (TE) populations. Breast cancer cells were injected into the mammary fat pads of immunodeficient mice at low cell numbers and the resulting tumors formed were dissociated into single-cell suspensions and re-injected over multiple rounds of serial dilution. (**b**) Flow chart depicting the generation of tumorigenic-enriched (TE) derivatives from the MDA-MB-231 (MDA-231) and CN34 parental cell lines (left). Table depicting the number of injected cells and the numbers of tumors formed per mammary fat pad injection into immunodeficient NOD scid mice during the process of generating the *in vivo* selected

derivatives (right). (**c**) MDA-TE3 cells exhibited enrichment for tumor re-initiation capacity as compared to MDA-parental cells following orthotopic injection of 1×10^2 cancer cells into NOD scid mice. MDA-TE3 cells yielded tumors in 12/20 sites compared to 1/20 sites for the MDA-parental cells after 10 weeks (left). Gross tumor explants (right). $n = 20$ independent mammary fat pad injections (pooled from 5 mice with 4 injections each per condition and represented as open squares, right). (**d**) CN34-TE2 cells exhibited enrichment for tumor re-initiation capacity as compared to CN34-parental cells following orthotopic injection of 2×10^3 cancer cells into NOD scid gamma mice. CN34-TE2 cells yielded tumors in 19/20 sites compared to 12/20 sites for the CN34-parental cells after 10 weeks (left). Gross tumor explants (right). $n = 20$ independent mammary fat pad injections (pooled from 5 mice with 4 injections each per condition and represented as open squares, right). $^{**}P<0.01$, $^{***}P<0.001$ were obtained using one-sided Fisher's exact test (**c**–**d**).

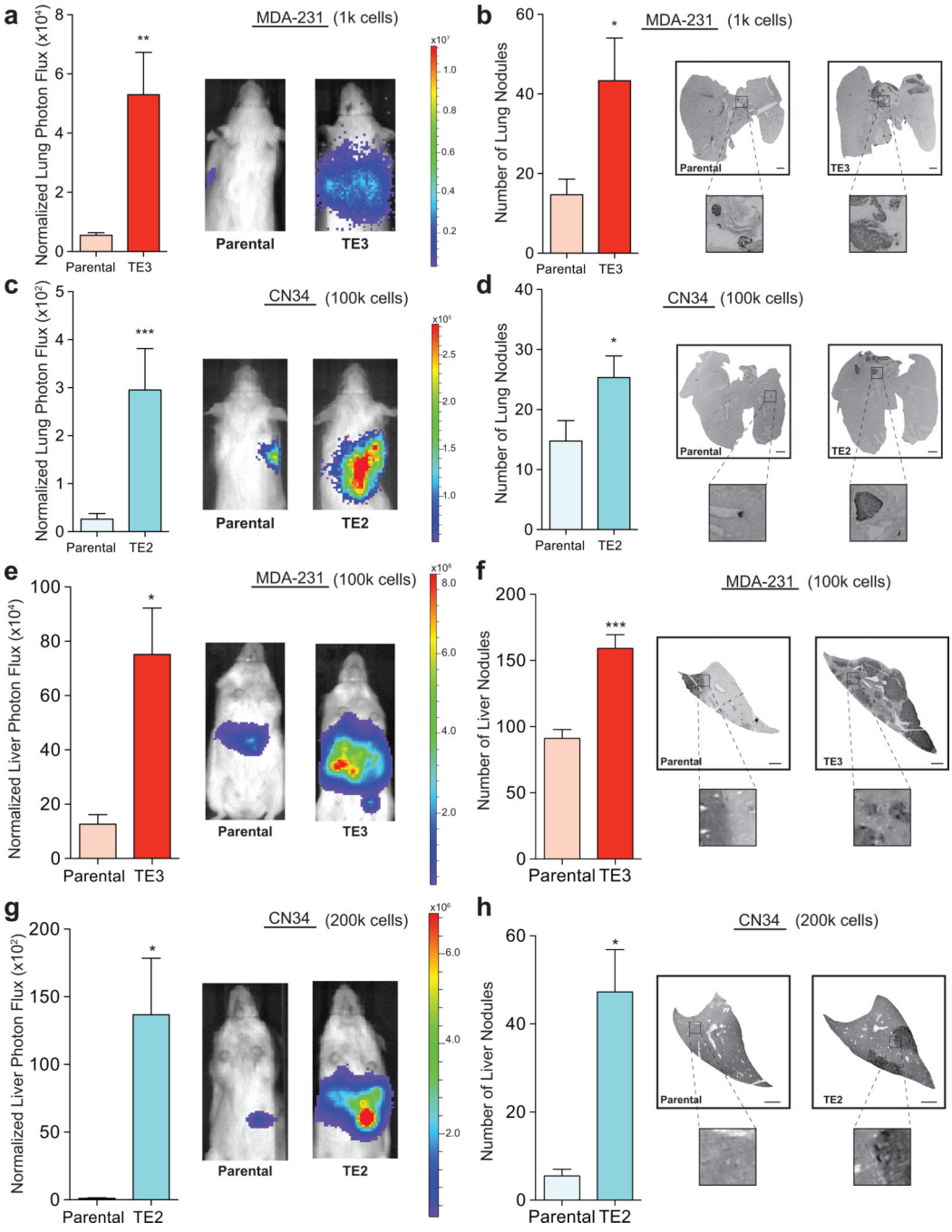

Figure 2. Tumorigenic-enriched cells robustly colonize multiple organs as compared to their parental populations

(**a**–**b**) 1×10^3 MDA-parental or MDA-TE3 cells were injected directly into the lung parenchyma to assess ectopic tumor re-initiation capacity. Lung bioluminescence was measured on day 53 and normalized to post-injection signal at day 0 (**a**). *n* = 7 independent mice. On day 53 lungs were sectioned, vimentin stained, and the number of macroscopic nodules per lung was counted (**b**). *n* = 3 lungs from 3 independent mice. Scale bars: 1mm. Insets are magnified 5x. (**c**–**d**) 1×10^5 CN34-parental or CN34-TE2 cells were injected

directly into the lung parenchyma to assess ectopic tumor re-initiation capacity. Lung bioluminescence was measured on day 56 and normalized to post-injection signal at day 0 (**c**). *n* = 8 (CN34-parental), *n* = 9 (CN34-TE2) independent mice. On day 56 lungs were sectioned, vimentin stained, and the number of macroscopic nodules per lung was counted (**d**). *n* = 8 (CN34-parental), *n* = 9 (CN34-TE2) lungs from independent mice. Scale bars: 1mm. Insets are magnified 5x. (**e**–**f**) 1×10^5 MDA-parental or MDA-TE3 cells were injected into the portal circulation via splenic injection in order to assess metastasis to the liver. Liver colonization was measured by bioluminescence imaging at day 35 normalized to post-injection signal at day 0 (**e**). *n* = 5 independent mice. On day 35 livers were sectioned, vimentin stained, and number of macroscopic nodules was counted (**f**). *n* = 5 livers from 5 independent mice. Scale bars: 1mm. Insets are magnified 5x. (**g**–**h**) 2×10^5 CN34-parental or CN34-TE2 cells were injected into the portal circulation via splenic injection in order to assess metastasis to the liver. Liver colonization was measured by bioluminescence imaging on day 56 and normalized to post-injection signal at day 0 (**g**). *n* = 4 independent mice. On day 56 livers were sectioned, vimentin stained, and number of macroscopic nodules was counted (**h**). *n* = 4 livers from 4 independent mice. Scale bars: 1mm. Insets are magnified 5x. **P*<0.05, ***P*<0.01, ****P*<0.001 were obtained using a one-sided Mann-Whitney test (**a**, **c**, **e**, **g**) or a one-sided Student's t-test (**b**, **d**, **f**, **h**). All data are represented as mean + S.E.M.

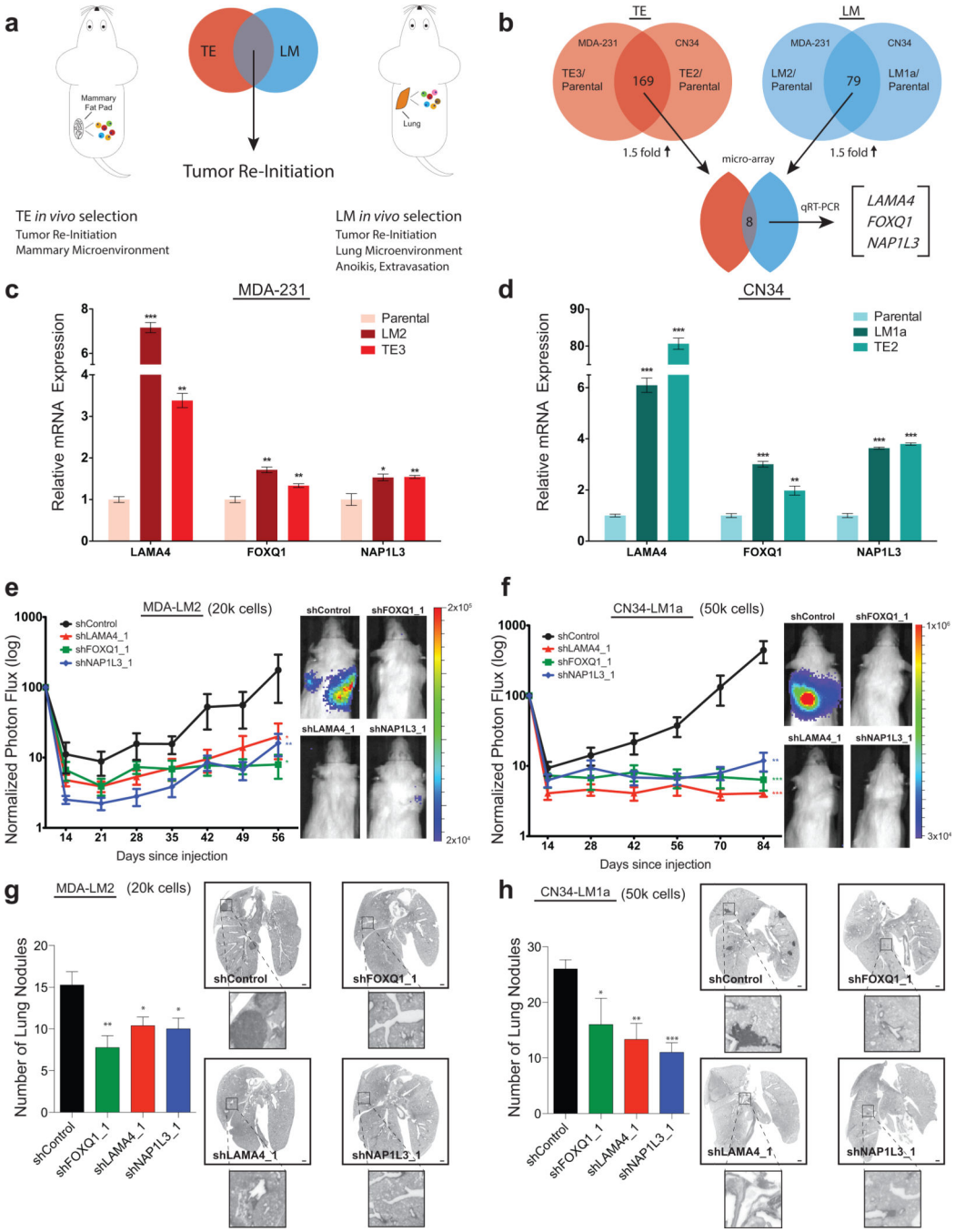

Figure 3. TE cells express increased levels of *LAMA4*, *FOXQ1*, and *NAP1L3*–genes also expressed at greater levels by highly metastatic cells that promote metastatic efficiency

(**a**) To identify candidate genes that promote general features of tumor re-initiation, we searched for genes whose expression was greater in tumorigenic-enriched (TE) and lung-metastatic (LM) *in vivo* selected derivatives. We hypothesized that the intersection of features common to TE and LM derivatives might exclude factors only specific to growth in the primary organ (mammary microenvironment) or the metastatic cascade (lung microenvironment, extravasation, anoikis). (**b**) Transcriptomic analysis of TE and LM

derivatives from the MDA-231 (MDA-TE3, MDA-LM2) and CN34 (CN34-TE2, CN34-LM1a) cell lines were compared relative to their parental populations. Eight candidate promoter genes passed a relative 1.5-fold increase cutoff criteria. Three genes (*LAMA4*, *FOXQ1*, and *NAP1L3*) demonstrated statistical significance by quantitative real-time PCR (qRT-PCR). (**c**–**d**) qRT-PCR of the mRNA expression levels of *LAMA4*, *FOXQ1*, and *NAP1L3* by *in vivo* selected derivatives (MDA-TE3, MDA-LM2, CN34-TE2, CN34-LM1a) from the MDA-231 (**c**) and CN34 (**d**) cell lines relative to their respective parental populations. $n = 3$ samples per group. (**e** and **g**) 2×10^4 MDA-LM2 cells transduced with control shRNA or shRNAs targeting *LAMA4*, *FOXQ1*, or *NAP1L3* were inoculated intravenously into immunodeficient mice. shRNA-depletion of *LAMA4*, *FOXQ1*, and *NAP1L3* led to significant reduction in metastasis based on bioluminescence over 56 days normalized to day 0 post-injection signal (**e**). $n = 8$ (shControl), n = 8 (shLAMA4_1), $n = 7$ (shNAP1L3_1), $n = 7$ (shFOXQ1_1) independent mice. Representative vimentin-stained lungs and quantification of macroscopic lung nodules on day 56 (**g**). $n = 4$ (shControl), $n = 5$ (shLAMA4_1), $n = 4$ (shNAP1L3_1), n = 5 (shFOXQ1_1) lungs from independent mice. Scale bars: 1mm. Insets are magnified 5x. (**f** and **h**) 5×10^4 CN34-LM1a cells transduced with control shRNA, or shRNAs targeting *LAMA4*, *FOXQ1*, or *NAP1L3* were injected intravenously into immunodeficient mice. shRNA-depletion of *LAMA4*, *FOXQ1*, and *NAP1L3* led to significant reduction in metastasis based on bioluminescence over 84 days normalized to day 0 post-injection signal (**f**). $n = 26$ (shControl), $n = 6$ (shLAMA4_1), $n = 5$ (shNAP1L3_1), $n = 5$ (shFOXQ1_1) independent mice. Representative vimentin-stained lungs and quantification of macroscopic lung nodules on day 84 (**h**). $n = 5$ (shControl), $n = 3$ (shLAMA4_1), $n = 3$ (shNAP1L3_1), $n = 3$ (shFOXQ1_1) lungs from independent mice. Scale bars: 1mm. Insets are magnified 5x. $^*P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$ obtained using one-sided Student's t-test (**c**–**d**, **g**–**h**), or one-sided Mann-Whitney test (**e**–**f**). All data are represented as mean +/± S.E.M.

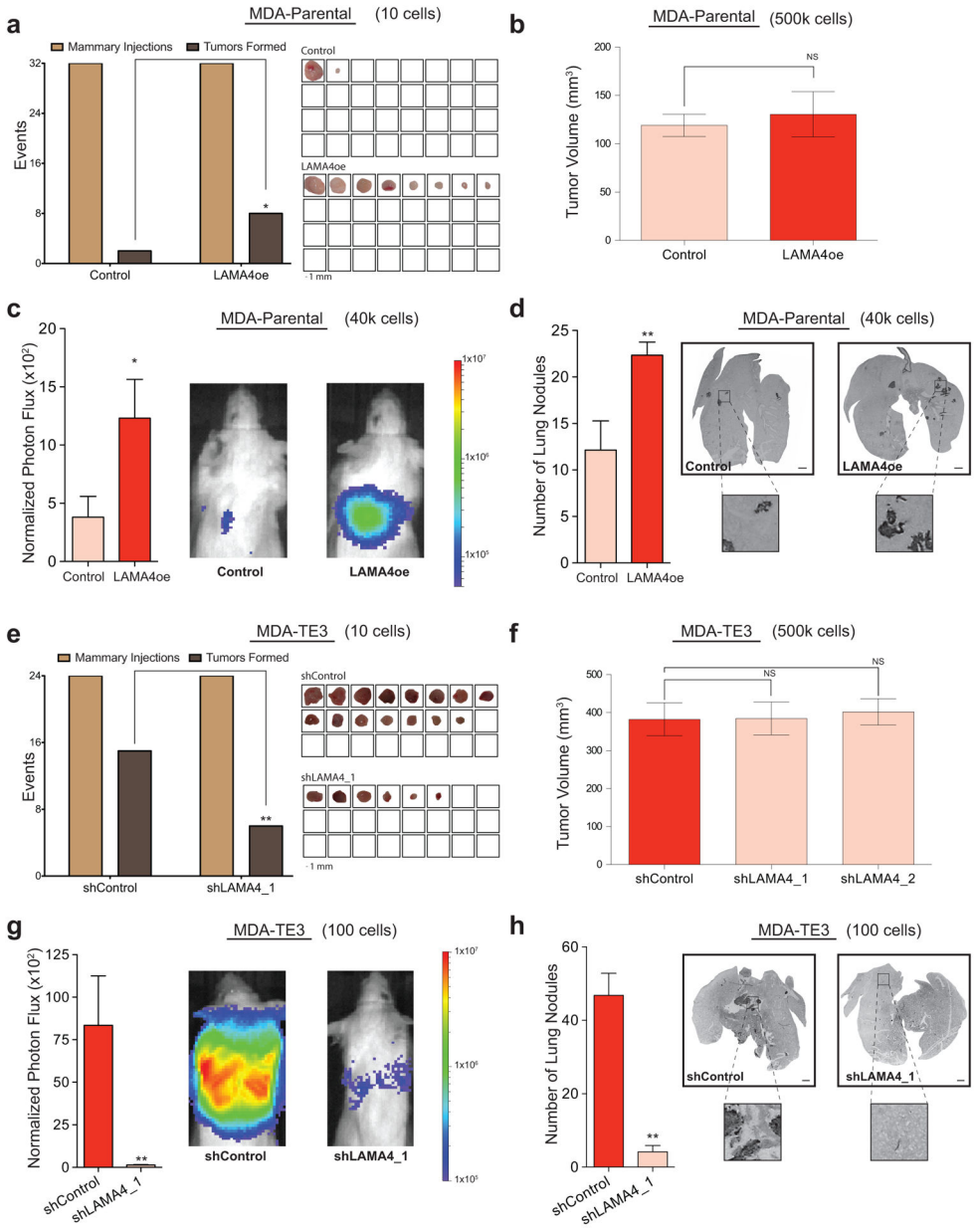

Figure 4. *LAMA4* promotes the re-initiation of tumors in multiple organ microenvironments

(**a**) 1×10^1 MDA-parental cells transduced with empty vector control or *LAMA4* over-expression vector (LAMA4oe) were injected into the mammary fat pads of immunodeficient mice. MDA-parental *LAMA4*-over-expressing cells yielded tumors in 8/32 sites versus 2/32 sites for MDA-parental control cells after 10 weeks (left). Tumor explants (right). n = 32 independent mammary fat pad injections (pooled from 8 mice with 4 injections each per condition and represented as open squares, right). (**b**) 5×10^5 MDA-parental cells transduced with either empty vector control or *LAMA4* over-expression vector were injected into the mammary fat pads of immunodeficient mice. There were no significant differences in the size of the tumors formed on day 29. n = 8 independent mammary fat pad injections. (**c**–**d**)

4×10^4 MDA-parental cells transduced with empty vector control or *LAMA4* over-expression vector (LAMA4oe) were inoculated intravenously into immunodeficient mice. Lung bioluminescence was measured on day 56 and normalized to day 0 post-injection signal (**c**). *n* = 7 independent mice. Lungs were harvested on day 56, vimentin-stained, and the number of macroscopic nodules per lung was counted. Representative vimentin-stained lungs on day 56 (**d**). *n* = 7 (control), n = 6 (LAMA4oe) lungs from independent mice. Scale bars: 1mm. Insets are magnified 5x. (**e**) 1×10^1 MDA-TE3 cells transduced with either control shRNA or shRNA targeting *LAMA4* were injected into the mammary fat pads of immunodeficient mice. MDA-TE3 shControl cells yielded tumors in 15/24 sites versus 6/24 sites for MDA-TE3 shLAMA4_1 cells after 10 weeks (left). Tumor explants (right). *n* = 24 independent mammary fat pad injections (pooled from 6 mice with 4 injections each per condition and represented as open squares, right). (**f**) 5×10^5 MDA-TE3 cells transduced with either control shRNA or two independent shRNAs targeting *LAMA4* were injected into the mammary fat pads of immunodeficient mice. There were no significant differences in the size of the tumors formed on day 39. *n* = 8 independent mammary fat pad injections (pooled from 2 mice with 4 injections each per condition). (**g**–**h**) 1×10^2 MDA-TE3 cells transduced with either control shRNA or an shRNA targeting *LAMA4* were injected directly into the lung parenchyma to assess ectopic tumor re-initiation capacity. Lung bioluminescence was measured on day 63 and normalized to day 0 post-injection signal (**g**). *n* = 5 independent mice. On day 63 lungs were sectioned, vimentin-stained, and the number of macroscopic nodules per lung was counted (**h**). *n* = 5 lungs from 5 independent mice. Scale bars: 1mm. Insets are magnified 5x. **P*<0.05, ***P*<0.01 obtained using one-sided Fisher's exact test (**a**, **e**), one-sided Mann-Whitney test (**b**, **c**, **f**, **g**) or one-sided Student's t-test (**d**, **h**). All data are represented as mean +/± S.E.M. NS, not significant.

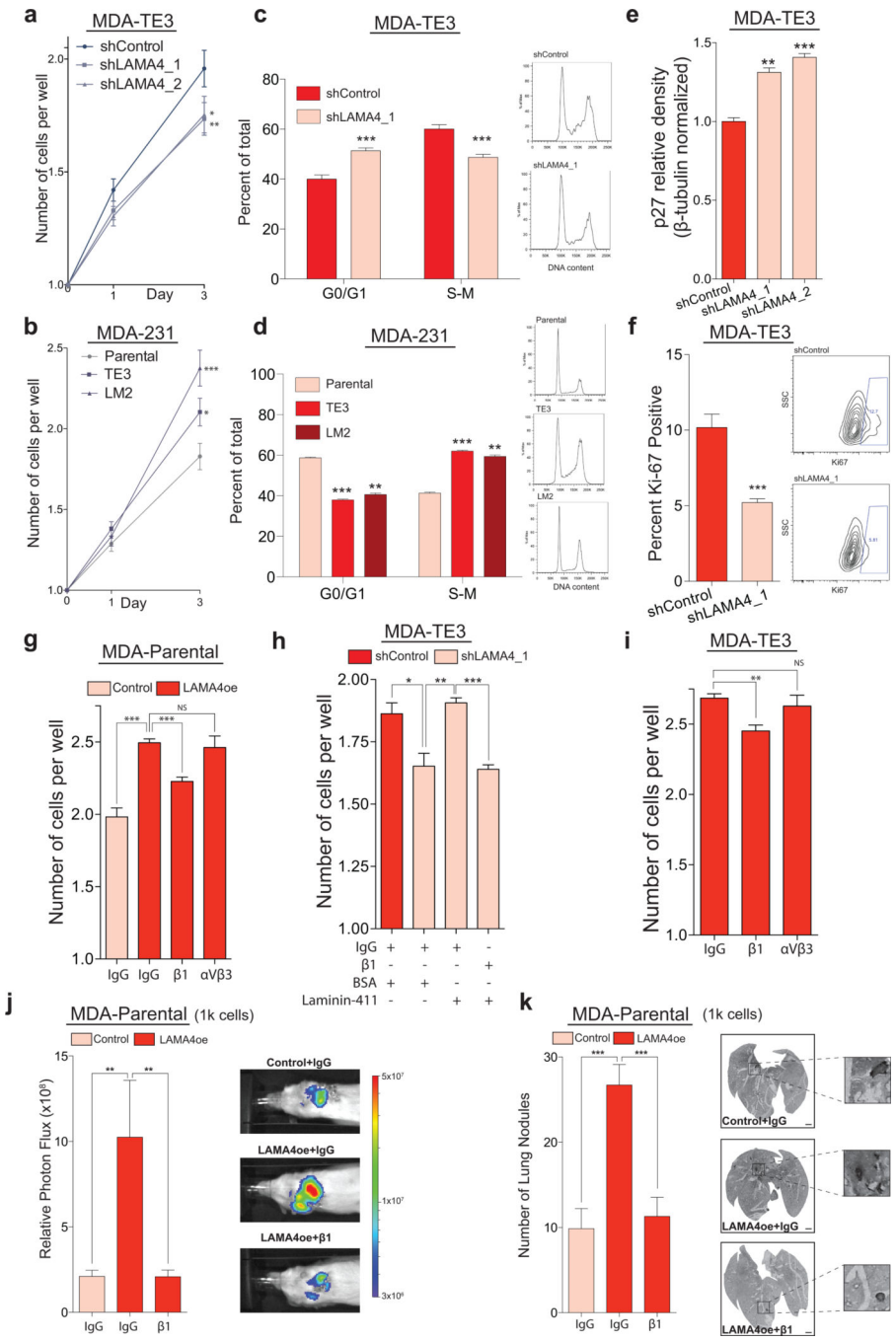

Figure 5. *LAMA4* promotes cancer cell proliferation in the absence of substratum-attachment *in vitro* in a β1-integrin dependent manner

(**a**) MDA-TE3 control cells or MDA-TE3 *LAMA4*-knockdown cells were sorted at clonal density (one cell per well) into low-attachment 96-well plates. *LAMA4*-depletion led to reduced proliferation. *n* = 212 (shControl), *n* = 219 (shLAMA4_1), *n* = 224 (shLAMA4_2) wells. (**b**) MDA-parental, MDA-TE3, or MDA-LM2 cells were sorted at clonal density into low-attachment plates. MDA-TE3 and MDA-LM2 cells proliferated more extensively than MDA-parental cells. *n* = 152 (parental), *n* = 163 (TE3), *n* = 141 (LM2) wells. (**c**) Cell-cycle

analysis of MDA-TE3 control cells or MDA-TE3 *LAMA4*-knockdown cells (left). Representative DNA content histograms (right). *n* = 6 independent samples. (**d**) Cell-cycle analysis of MDA-parental, MDA-TE3, or MDA-LM2 cells (left). Representative DNA content histograms (right). *n* = 4 independent samples. (**e**) Relative p27 protein levels in MDA-TE3 control cells or MDA-TE3 *LAMA4*-knockdown cells quantified by Western. p27 values normalized to β-tubulin. *n* = 3 independent samples. Representative blot in Supplementary Fig. 5l. (**f**) The fraction of Ki67-positive cells was assessed in MDA-TE3 control cells or MDA-TE3 *LAMA4*-knockdown cells using flow-cytometry (left). Representative plots (right). *n* = 6 independent samples. (**g**) Incubation with β1-integrin blocking antibody suppressed the proliferation of MDA-parental cells over-expressing *LAMA4* upon sorting at clonal density into low-attachment plates. Counts on day 3. *n* = 4 independent experiments. (**h**) The capacity of recombinant *LAMA4*-containing protein (laminin-411) to rescue the proliferation of MDA-TE3 *LAMA4*-knockdown cells was assessed in the presence/absence of β1-integrin blocking antibody upon sorting at clonal density into low-attachment plates. Counts on day 3. *n* = 3 independent experiments. (**i**) Incubation with β1-integrin blocking antibody suppressed the proliferation of MDA-TE3 cells upon sorting a clonal density into low-attachment plates. Counts on day 3. *n* = 3 independent experiments. (**j**–**k**) 1×10^3 MDA-Parental control cells or *LAMA4*-over-expressing (LAMA4oe) cells were injected directly into the lung parenchyma in the presence of IgG control or β1-integrin blocking antibody to assess ectopic tumor re-initiation capacity. Lung bioluminescence was measured on day 39 (**j**). *n* = 7 independent mice. On day 39 lungs were sectioned, vimentin-stained, and the number of macroscopic nodules per lung was counted (**k**). *n* = 7 lungs from 7 independent mice. Scale bars: 1mm. Insets are magnified 5x. **P*<0.05, ***P*<0.01, ****P*<0.001 obtained using one-sided Mann-Whitney test (**a**–**b**, **j**–**k**), or one sided Student's t-test (**c**–**i**). Experiments (**a**–**f**) are representative and were replicated at least 2 times. All data are represented as mean +/± S.E.M.

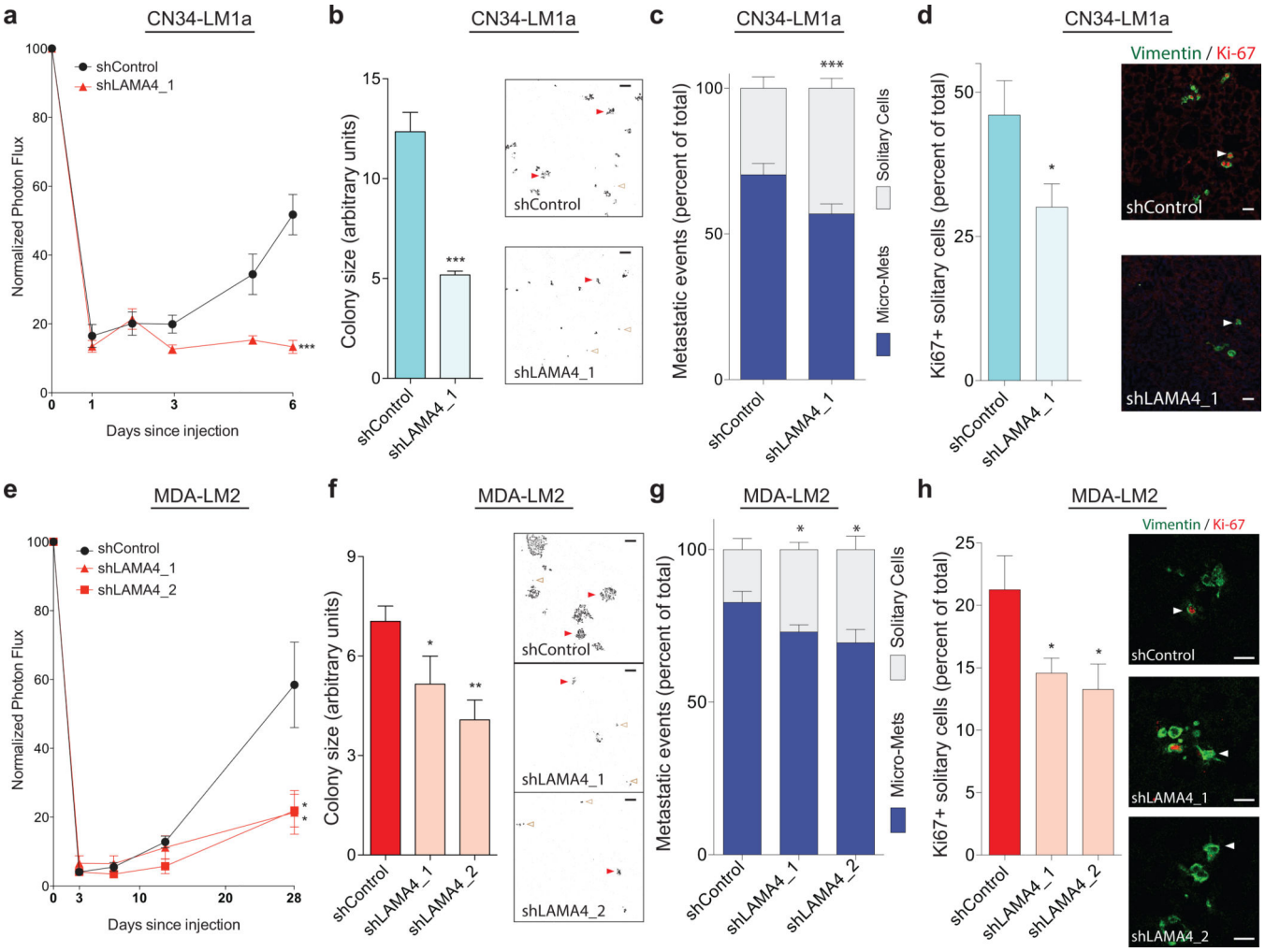

Figure 6. *LAMA4* promotes the proliferation of disseminated metastatic cells and micro-metastasis formation *in vivo*

(**a**–**d**) 3×10^5 CN34-LM1a cells transduced with either control shRNA or an shRNA targeting *LAMA4* were injected intravenously into immunodeficient mice and lung bioluminescence was measured over time (**a**). *n* = 5 independent mice. On day 6, lungs were harvested, sectioned, and stained with vimentin and Ki-67. Upon fluorescence image acquisition using confocal microscopy and software analysis, the average size of metastatic colonies was quantified (**b**; representative background-subtracted vimentin stained images, right. Scale bars: 100μm), the number of solitary single cells relative to multi-cellular colonies (micro-metastases) was quantified (**c**; see panel **b** for representative images. Red closed arrows, micro-metastases. Brown open arrows, solitary cells), and the percentage of solitary single cells that stained positively for Ki67 was quantified (**d**; vimentin staining in green. Ki67 staining in red. Arrows depict solitary cells. Scale bars: 25μm). *n* = 5 lungs from 5 independent mice. Micro-mets is micro-metastases. (**e**–**h**) 3×10^5 MDA-LM2 cells transduced with either control shRNA or two independent shRNAs targeting *LAMA4* were injected intravenously into immunodeficient mice and lung bioluminescence was measured over time (**e**). *n* = 6 (shControl), *n* = 4 (shLAMA4_1), *n* = 4 (shLAMA4_2) independent

mice. On day 28 lungs were harvested, sectioned, and stained with vimentin and Ki-67. Upon fluorescence image acquisition using confocal microscopy and software analysis, the average size of metastatic colonies was quantified (**f**; representative background-subtracted vimentin stained images, right. Scale bars: 100μm), the number of solitary single cells relative to multi-cellular colonies (micro-metastases) was quantified (**g**; see panel **f** for representative images. Red closed arrows, micro-metastases. Brown open arrows, solitary cells), and the percentage of solitary single cells that stained positively for Ki67 was quantified (**h**; vimentin staining in green. Ki67 staining in red. Arrows depict solitary cells. Scale bars: 25μm). n = 4 lungs from 4 independent mice. Micro-mets is micro-metastases. *P<0.05, **P<0.01, ***P<0.001 were obtained using a one-sided Student's t-test (**a**–**h**). All data are represented as mean +/± S.E.M.

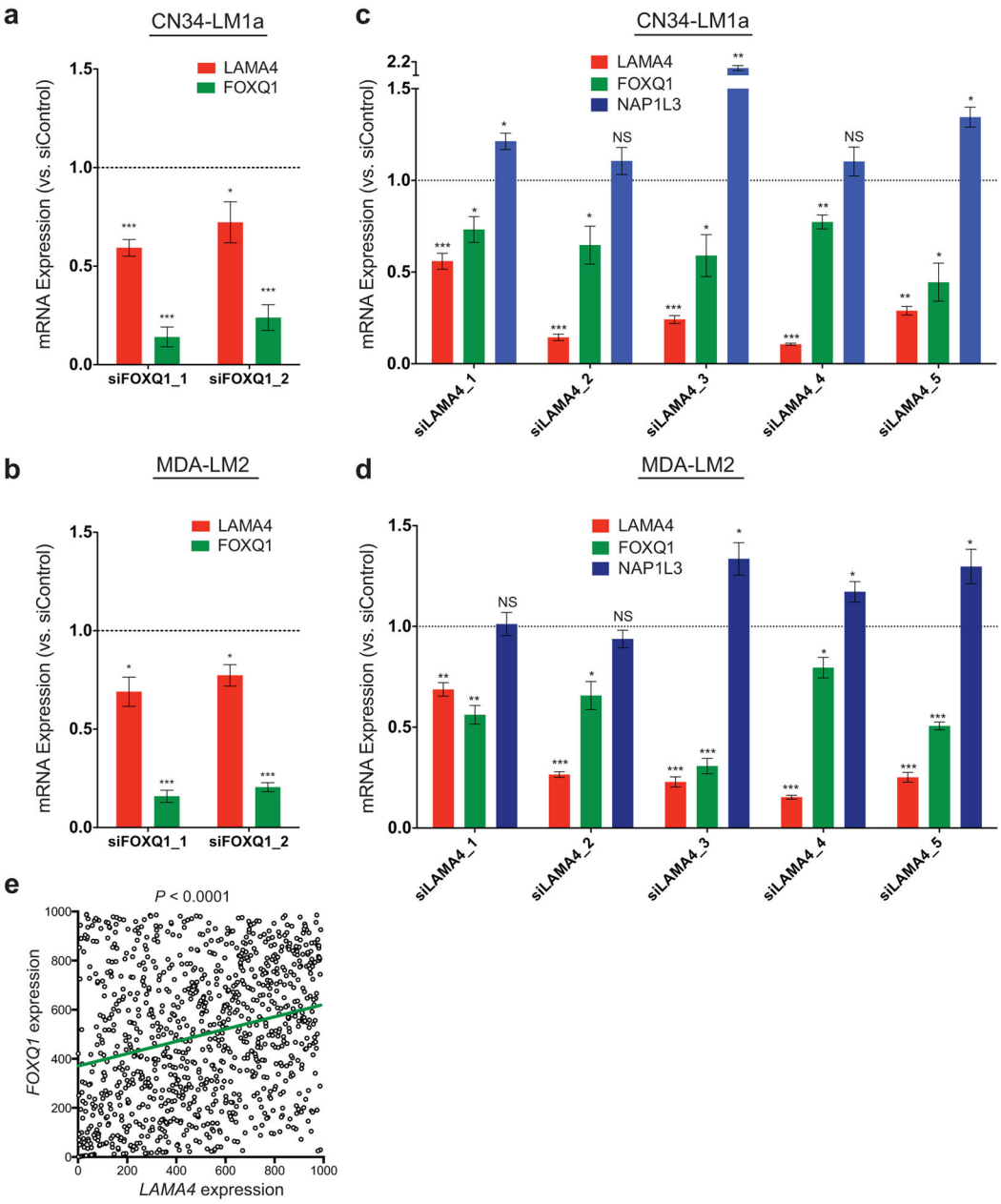

Figure 7. *FOXQ1* promotes the expression of *LAMA4*

(**a**–**b**) Relative expression of *LAMA4* and *FOXQ1* following siRNA knockdown of *FOXQ1* in LM1as (**a**; *n* = 9) or LM2s (**b**; *n* = 6) compared to control siRNA. *n* is the number of samples collected over 3 (**a**) or 2 (**b**) independent experiments. (**c**–**d**) Relative expression of *LAMA4*, *FOXQ1* and *NAP1L3* following siRNA knockdown of *LAMA4* in LM1as (**c**; *n* = 4 for siLAMA4_1,2,3,4, *n* = 6 for siLAMA4_5) or LM2s (**d**; *n* = 4 for siLAMA4_1,2,3,4, *n* = 6 for siLAMA4_5) compared to control siRNA. *n* is the number of samples collected over 2 independent experiments. (**e**) *LAMA4* expression was positively correlated with the expression of *FOXQ1* in a large set of primary breast tumor samples. Ranked *LAMA4* (x-axis) expression values were plotted against ranked *FOXQ1* (y-axis) expression values. *n* =

988 independent primary breast cancer patient samples. R = 0.25. **P*<0.05, ***P*<0.01, ****P*<0.001 were obtained using a two-sided one-sample t-test (**a**–**d**) or a linear regression (**e**). All data are represented as mean ± S.E.M.

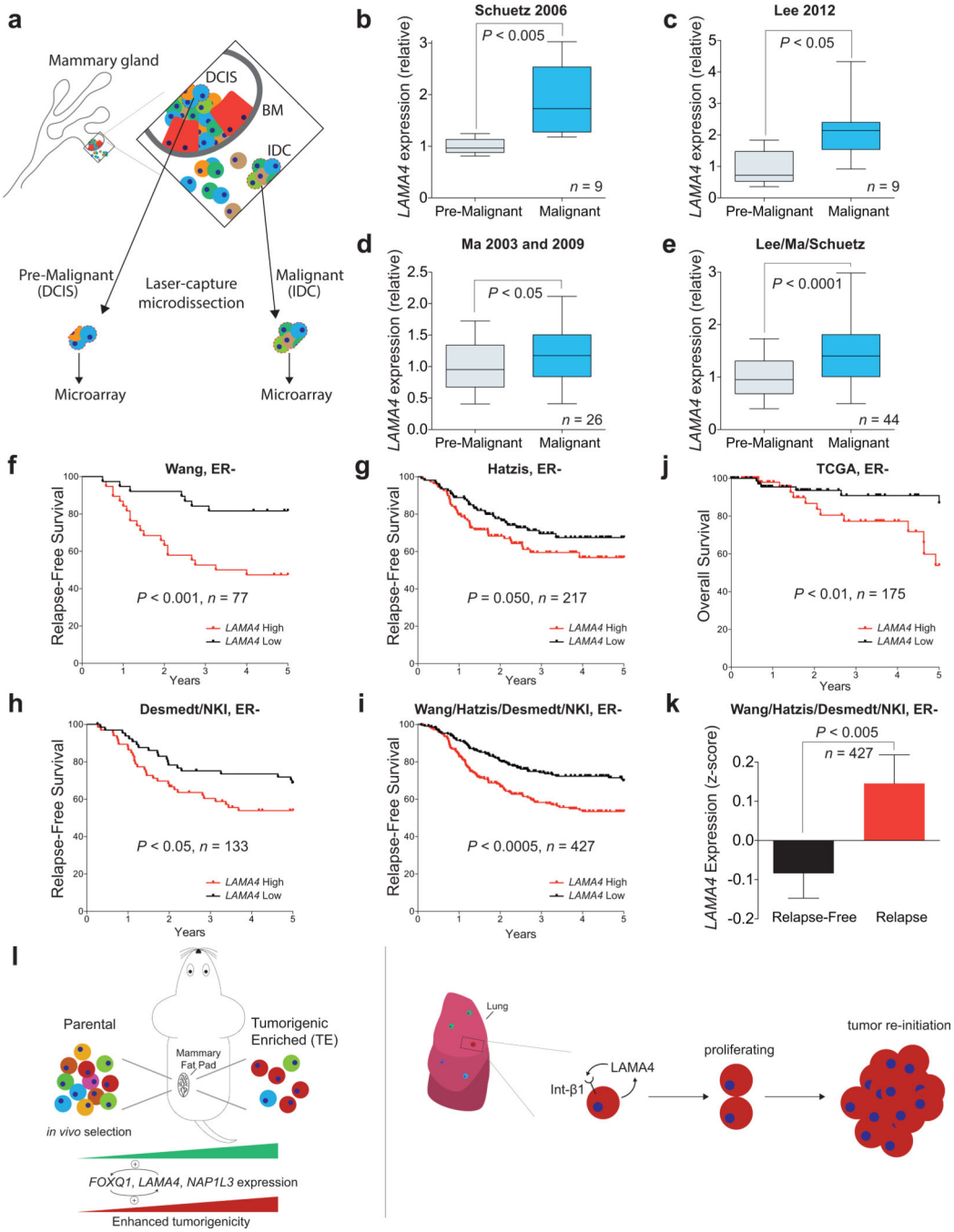

Figure 8. Increased expression of *LAMA4* marks early breast cancer progression and is correlated with clinical relapse

(**a**–**e**) *LAMA4* mRNA expression was assessed in multiple gene expression datasets (see Methods) from laser capture microdissected pre-malignant or malignant human breast cancer tissue isolated from individual patients (**a**). DCIS is ductal carcinoma *in situ*. IDC is intra-ductal carcinoma. BM is basement membrane. Schuetz 2006 (**b**; $n = 9$), Lee 2012 (**c**; $n = 9$), Ma 2003 and 2009 (**d**; $n = 26$), Lee/Ma/Schuetz (**e**; $n = 44$). n is the number of independent paired DCIS and IDC patient samples. (**f**–**i**) Kaplan-Meier curves depicting

relapse-free survival of patients with ER-negative breast cancer as a function of their primary tumor's expression of *LAMA4*. Patients whose tumor's mRNA expression of *LAMA4* was greater or lower than the median for the population were classified as either *LAMA4* High or *LAMA4* Low. Wang (**f**, *n* = 77), Hatzis (**g**, *n* = 217), Desmedt/NKI (**h**, *n* = 133), Wang/Hatzis/Desmedt/NKI (**i**; *n* = 427). *n* is the number of independent patient samples. ER- is ER-negative. (**j**) Kaplan-Meier curve depicting overall survival of patients with ER-negative breast cancer as a function of their primary tumor's expression of *LAMA4*. Patients whose tumor's mRNA expression of *LAMA4* was greater or lower than the median for the population were classified as either *LAMA4* High or *LAMA4* Low. *n* = 175 independent patient samples. ER− is ER-negative. (**k**) The expression of *LAMA4* in primary tumors from ER-negative patients that relapsed or did not relapse in the combined Wang/Hatzis/Desmedt/NKI dataset was compared. *n* = 427. ER− is ER-negative. (**l**) Schematic depicting *in vivo* selection for tumorigenic-enriched cells, which demonstrated enhanced tumor re-initiation capacity and increased expression of *LAMA4*, *FOXQ1*, and *NAP1L3* relative to their parental populations (left). Model depicting *LAMA4* promotion of proliferation and multi-cellular colony formation in the metastatic niche (right). *P* values were obtained using a one-sided paired t-test (**b**–**e**), a one-sided Mantel-Cox test (**f**–**j**), or a one-tailed Mann-Whitney test (**k**). For box and whiskers graphs, box extends from the 25th to the 75th percentile, line inside box is median, whiskers are 5th and 95th percentile.