

2015

# Molecular Determinants of Tumor Re-Initiation in Breast Cancer

Jason Barzel Ross

Follow this and additional works at: [http://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations](http://digitalcommons.rockefeller.edu/student_theses_and_dissertations)



Part of the [Life Sciences Commons](#)

---

## Recommended Citation

Ross, Jason Barzel, "Molecular Determinants of Tumor Re-Initiation in Breast Cancer" (2015). *Student Theses and Dissertations*. Paper 289.



MOLECULAR DETERMINANTS OF  
TUMOR RE-INITIATION IN BREAST CANCER

A Thesis Presented to the Faculty of  
The Rockefeller University  
in Partial Fulfillment of the Requirements for  
the degree of Doctor of Philosophy

by

Jason Barzel Ross

June 2015



# MOLECULAR DETERMINANTS OF TUMOR RE-INITIATION IN BREAST CANCER

Jason Barzel Ross, Ph.D.

The Rockefeller University 2015

Cancer progression is characterized by the formation of tumors in primary organs and the subsequent re-formation of tumors in distant metastatic sites. The tumorigenic, or tumor-forming, capacity of cancer cells drives cancer along a continuum that includes primary tumor formation, metastatic re-initiation, and macroscopic relapse. As such, a greater understanding of the biological features and molecular determinants that govern tumor-forming capacity in the primary and metastatic site is of great importance to the scientific and biomedical communities. This thesis presents an unbiased approach to systematically select and characterize breast cancer cells with enhanced tumorigenic capacity in order to elucidate genes and cell biological features that are important to tumor-forming potential in primary and metastatic sites.

The first part of this thesis describes a method to select for populations of cells with enhanced tumor-forming capacity from heterogeneous breast cancer populations. Two human Estrogen Receptor-negative (ER-negative) breast cancer cell lines, MDA-231 and CN34, were subjected to *in vivo* selection to yield ‘tumorigenic-enriched’ (TE)

derivatives, which demonstrated enhanced tumor re-initiation capacity in multiple organ sites in a xenograft model. In the second part of this study, a systematic approach was employed to identify genes that promote tumor re-initiation. Transcriptomic profiling revealed a set of genes—*LAMA4*, *FOXQ1*, and *NAPIL3*—that were expressed at greater levels in tumorigenic-enriched derivatives relative to their parental populations. These genes were also expressed at greater levels in independently derived metastatic populations. All three of these genes were found to promote metastatic efficiency, and one of these genes, *LAMA4*, was subjected to further characterization. Functional studies revealed *LAMA4* to promote clonal expansion during substratum-detachment *in vitro*, tumor re-initiation in multiple organ microenvironments, the proliferation of disseminated metastatic cells, and the formation of incipient micro-metastatic colonies *in vivo*.

The final phase of this study revealed the association of *LAMA4* expression with human breast cancer progression and clinical outcome in multiple patient datasets. In support of the role of *LAMA4* in promoting early tumor formation, malignant breast cancer cells isolated from incipient breast cancer lesions were found to express higher levels of *LAMA4* relative to pre-malignant cells from matched patient lesions. In addition, consistent with functional studies demonstrating the role of *LAMA4* in driving primary and metastatic tumor re-initiation, higher levels of *LAMA4* expression in ER-negative primary tumors correlated with worse relapse-free survival in multiple patient clinical cohorts.

*To my parents and siblings, for their continuous support and encouragement.*

## ACKNOWLEDGMENTS

First and foremost, I would like to sincerely thank my mentor and advisor Dr. Sohail Tavazoie for his mentorship, encouragement, and support over the past 5 years. His passion, dedication, and remarkable talent for science has served as an inspiring example to me, and I feel privileged for the opportunity to learn from him and to be a member of his lab. Furthermore, his creative ideas, generous support, and innovative approaches to solving problems have laid the foundation and made possible this work and many other exciting projects in the lab.

I would also like to extend my gratitude to several people for their contributions to this project. In particular, I would like to acknowledge and thank Doowon Huh, an outstanding member of the laboratory who worked closely with me for over 2.5 years (first as a technician, and then as an MD-PhD student) and whose hard work contributed to several important findings of this project. In addition, I would like to thank Lisa Noble for assisting me with molecular biology and immunohistochemistry during her rotation in the lab and later as a 1<sup>st</sup> year medical student. I would also like to express my appreciation to other members of the Tavazoie Lab, including Nora Pencheva, Alexander Nguyen, Jia Min Loo, and Dr. Nils Halberg for their assistance with difficult mouse surgeries, to Paul Furlow for his assistance in cloning *LAMA4*, and to Dr. Hani Goodarzi for his help with statistical and computational analysis. I want to especially thank Dr. Claudio Alarcon, Nora Pencheva, Alexander Nguyen, Lisa Fish, and Caitlin Sengelaub for helpful discussion of this project and their scientific advice.

I am also extremely grateful to the members of my thesis committee, Dr. Jeffrey Friedman, Dr. Sanford Simon, and Dr. Carl Blobel for their time and consideration, invaluable guidance, and for their rigorous and honest critique of my work. I would also like to extend my thanks to Dr. Jose Silva for serving as my external examiner. I would like to acknowledge the Tri-Institutional MD-PhD program, which has supported me over the past seven years. In particular, I am deeply thankful to the MD-PhD Program Director Dr. Olaf Andersen for his leadership, direction, and encouragement throughout all stages of my time in the program.

Finally, I want to express my sincere gratitude to my friends and family for their unwavering support.



# TABLE OF CONTENTS

<b>Acknowledgments</b> .....	iv
<b>List of Illustrations</b> .....	ix
<b>List of Figures</b> .....	x
<b>List of Tables</b> .....	xii
<b>List of Abbreviations</b> .....	xiii
 <b>Chapter I: Introduction</b> .....	 1
<b>Breast cancer: prevalence and prognosis</b> .....	1
<b>Breast cancer: origins and progression</b> .....	2
Pathophysiology.....	2
Pre-malignant and malignant breast cancer .....	4
Primary tumor formation .....	7
The metastatic cascade.....	7
Primary and Metastatic Relapse.....	12
<b>Breast cancer: therapeutic intervention</b> .....	12
<b>Breast cancer: experimental models</b> .....	13
Tumor re-initiation.....	13
Experimental metastasis.....	15
<i>In vivo</i> selection .....	16
<b>Specific aims</b> .....	18
 <b>Chapter II: Selection and Characterization of Breast Cancer Cells with</b>	
<b>Enhanced Tumorigenic Capacity</b> .....	20
<i>In vivo</i> selection for tumor re-initiation yields tumorigenic-enriched (TE) derivatives	20
Tumorigenic-enriched (TE) derivatives demonstrate enhanced orthotopic tumor re-	
initiation capacity relative to their parental populations.....	24
Tumorigenic-enriched (TE) derivatives robustly colonize multiple organs relative to	
their parental populations.....	26

Characterization of tumorigenic-enriched derivatives.....	30
<b>Summary of findings from Chapter II.....</b>	<b>35</b>
<b>Chapter III: Identification of Molecular Determinants of Tumor Re-Initiation</b>	<b>37</b>
<i>LAMA4</i> , <i>FOXQ1</i> , and <i>NAPIL3</i> display increased expression levels in tumorigenic-enriched and metastatic cells relative to parental populations.....	37
<i>LAMA4</i> , <i>FOXQ1</i> , and <i>NAPIL3</i> promote metastatic efficiency .....	47
<i>LAMA4</i> promotes tumor re-initiation in multiple organ microenvironments .....	52
<i>LAMA4</i> promotes the clonal expansion of breast cancer cells in the absence of substratum-attachment <i>in vitro</i> .....	61
<i>LAMA4</i> promotes metastatic proliferation and incipient micro-metastasis formation .	67
<i>FOXQ1</i> promotes the expression of <i>LAMA4</i> .....	72
<b>Summary of findings in Chapter III .....</b>	<b>76</b>
<b>Chapter IV: The Clinical Relevance of These Findings to Breast Cancer .....</b>	<b>77</b>
Increased expression of <i>LAMA4</i> marks early breast cancer progression.....	77
Increased expression of <i>LAMA4</i> is associated with poor relapse-free survival.....	80
<b>Summary of findings from Chapter IV .....</b>	<b>82</b>
<b>Chapter V: Summary and Perspectives.....</b>	<b>83</b>
<b>Overall summary .....</b>	<b>83</b>
<b>Derivation and characterization of tumorigenic-enriched populations.....</b>	<b>86</b>
<b>Identification of molecular determinants of tumor re-initiation in breast cancer</b>	<b>88</b>
<i>LAMA4</i> , <i>FOXQ1</i> , <i>NAPIL3</i> are a co-varying set of genes expressed at greater levels by tumorigenic-enriched and metastatic derivatives.....	88
<i>FOXQ1</i> .....	88
<i>NAPIL3</i> .....	89
<i>LAMA4</i> .....	90
<b>Phenotypes regulated by <i>LAMA4</i> .....</b>	<b>92</b>
Organ-independent tumor re-initiation .....	92
Clonal Expansion .....	93
<b>Implications of this study to human breast cancer .....</b>	<b>93</b>
Therapeutic applications .....	93

Primary and Metastatic Relapse.....	94
<b>Conclusion .....</b>	<b>95</b>
<b>Materials and Methods .....</b>	<b>96</b>
Animal studies .....	96
<i>In vivo</i> selection .....	97
Tumor dissociation .....	98
Orthotopic and ectopic tumor re-initiation assays .....	99
Metastasis assays .....	99
Cell culture.....	100
Cell proliferation and colony formation assay.....	102
Endothelial recruitment assay .....	102
Cell attachment assay.....	103
Generation of retrovirus-mediated knockdown cells.....	103
Generation of retrovirus-mediated over-expressing cells.....	104
RNA extraction and quantification of mRNA expression.....	105
siRNA knockdown experiments .....	106
Immunohistochemistry and histology.....	107
Western blotting.....	108
Microarray hybridization and transcriptomic analysis of <i>in vivo</i> selected derivatives.....	108
Correlation coefficient analysis .....	109
mRNA analysis of micro-dissected patient samples.....	109
Clinical Correlation Analysis.....	110
Statistical Analysis.....	111
Accession Numbers .....	111
<b>References .....</b>	<b>112</b>

## LIST OF ILLUSTRATIONS

### Chapter I:

- 1.1 The mammary gland
- 1.2 Pre-malignant and malignant breast cancer
- 1.3 The metastatic cascade

### Chapter V:

- 5.1 *In vivo* selection for tumorigenic-enriched cells uncovers *LAMA4* as a promoter of tumor re-initiation in breast cancer

## LIST OF FIGURES

### Chapter II:

- 2.1: *In vivo* selection for tumor re-initiation
- 2.2: Selection for multiple tumorigenic-enriched (TE) derivative populations
- 2.3: Tumorigenic-enriched (TE) derivatives demonstrate enhanced tumor re-initiation capacity
- 2.4: Tumorigenic-enriched cells demonstrate enhanced ectopic tumor re-initiation capacity relative to their parental populations
- 2.5: Tumorigenic-enriched cells demonstrate enhanced liver metastasis relative to their parental populations
- 2.6: *In vitro* characterization of tumorigenic-enriched derivatives
- 2.7: Immuno-phenotypic characterization of tumorigenic-enriched derivatives

### Chapter III:

- 3.1: Systematic identification of candidate promoters of tumor re-initiation
- 3.2: TE cells express increased levels of *LAMA4*, *FOXQ1*, and *NAPIL3*—genes also expressed at greater levels by highly metastatic cells
- 3.3: *LAMA4*, *FOXQ1*, and *NAPIL3* promote metastatic efficiency
- 3.4: Independent shRNA knockdown of *LAMA4*, *FOXQ1*, and *NAPIL3* leads to suppression of metastasis *in vivo*
- 3.5: TE3 cells secrete greater levels of laminin- $\alpha$ 4 relative to their parental populations
- 3.6: Over-expression of *LAMA4* in parental populations is sufficient to promote orthotopic tumor re-initiation
- 3.7: Assessment of *LAMA4*'s regulation of tumor growth upon injection of high cell doses
- 3.8: Over-expression of *LAMA4* in parental populations is sufficient to promote metastasis
- 3.9: Endogenous expression of *LAMA4* by TE derivatives promotes orthotopic tumor re-initiation

- 3.10: Endogenous expression of *LAMA4* by TE derivatives promotes ectopic tumor re-initiation
- 3.11: *LAMA4* promotes the clonal expansion of cancer cells in the absence of substratum-attachment *in vitro*
- 3.12: Tumorigenic-enriched and highly metastatic derivatives display enhanced clonal expansion in the absence of substratum-attachment *in vitro*
- 3.13: *LAMA4* represses the fraction of cells in G0/G1 in the absence of substratum-attachment *in vitro*
- 3.14: *LAMA4* promotes the proliferation of disseminated metastatic cells and micro-metastasis formation *in vivo*
- 3.15: *LAMA4* does not significantly regulate apoptosis at early time points during metastatic colonization
- 3.16: Independent validation of *LAMA4*'s promotion of disseminated metastatic cell proliferation and colony formation *in vivo*
- 3.17: *LAMA4* expression co-varies with the expression of *FOXQ1* and *NAPIL3* in primary breast tumors
- 3.18: *FOXQ1* promotes the expression of *LAMA4*

#### **Chapter IV:**

- 4.1: Increased expression of *LAMA4* marks early breast cancer progression
- 4.2: Increased expression of *LAMA4* is correlated with clinical relapse

## **LIST OF TABLES**

### **Chapter III**

3.1: Genes with increased expression (1.5x higher) in tumorigenic-enriched (TE) derivatives relative to their parental populations.

3.2: Genes with increased expression (1.5x higher) in lung-metastatic (LM) derivatives relative to their parental populations.

3.3: Genes with increased expression in tumorigenic-enriched (TE) and lung-metastatic (LM) cells relative to their parental populations (1.5x higher in all comparisons).

## LIST OF ABBREVIATIONS

BM	Basement Membrane
DCIS	Ductal Carcinoma <i>In Situ</i>
ECM	Extracellular Matrix
ER	Estrogen Receptor
FOXQ1	Forkhead Box Q1
IDC	Invasive Ductal Carcinoma
IGFBP4	Insulin-Like Growth Factor Binding Protein 4
LAMA4	Laminin, Alpha 4
NAP1L3	Nucleosome Assembly Protein 1-Like 3
MERTK	C-Mer Proto-Oncogene Tyrosine Kinase
SOX4	(Sex Determining Region Y)-Box 4
TE	Tumorigenic-Enriched
TNC	Tenascin C



## **CHAPTER I: INTRODUCTION**

### **Breast cancer: prevalence and prognosis**

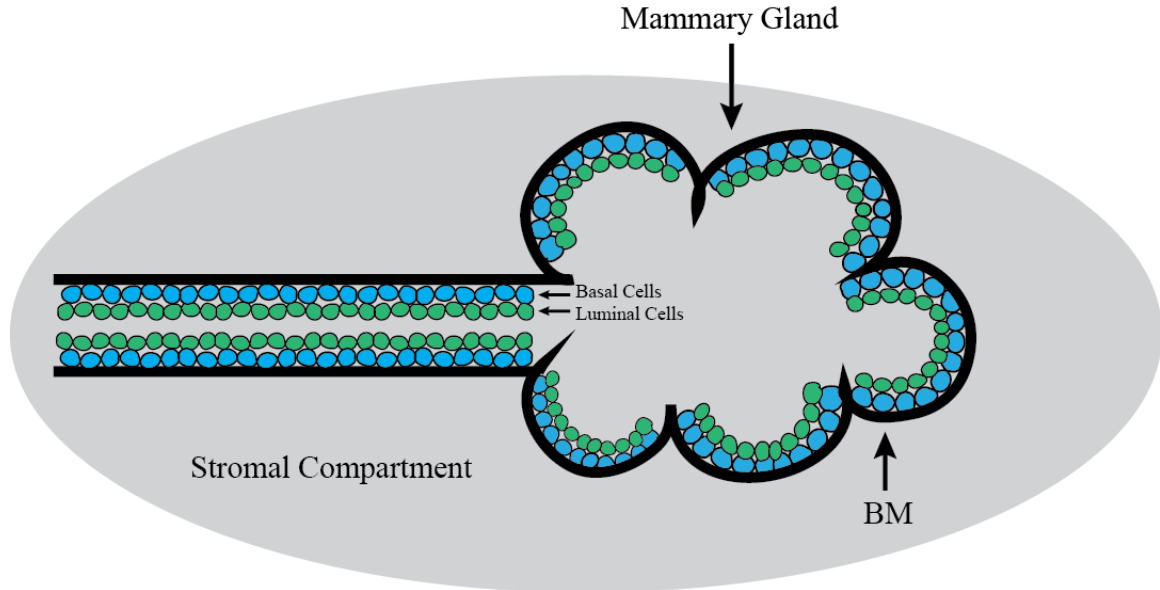
Breast cancer leads all other female cancers in the numbers of annual diagnoses and deaths incurred globally (Jemal et al., 2011). Without intervention, almost all breast cancers will follow a common progression path through pre-malignant, malignant, and metastatic stages (Espina and Liotta, 2011). Transition through each of these stages is thought to be intricately linked to the capacity of cancer cells to initiate colonies and give rise to tumors (Magee et al., 2012). Although pre-malignant breast cancers have a low risk of mortality, they are often excised because a portion of them will eventually transition to malignant cancers, which have the ability to spread into the surrounding breast tissue (Leonard and Swain, 2004). Localized pre-malignant or malignant breast cancers have a median 5-year survival time upwards of 95% (Jemal et al., 2010), but malignant breast cancers are treated more aggressively due to their risk of metastatic dissemination. The 5-year survival for patients with metastatic breast cancer, disease that has spread from the primary site to a distal organ, is less than 25% (Jemal et al., 2010). Unfortunately, patients that have undergone treatment at any stage of breast cancer are at risk for relapse of either primary or metastatic disease (Hedley and Chambers, 2009). Such relapse is often the result of the re-formation of tumors from small residual numbers of breast cancer cells that have escaped treatment (Chaffer and Weinberg, 2011). An understanding of the features that enable cancer cells to form tumors and to re-initiate

tumors in distal organs could lead to more effective therapeutic intervention at multiple stages of cancer progression.

### **Breast cancer: origins and progression**

#### *Pathophysiology*

Breast cancer arises from the mammary gland, an organ residing in the breast that is surrounded by fat and connective tissue (Thomas, 2011). The mammary gland is comprised of a branching network of ducts and lobules lined with a bi-layered epithelium (Illustration 1.1) (Gusterson, 2009). The inner layer of the mammary ducts are lined by luminal cells, while the outer layer is lined by basal cells that are in direct contact with the basement membrane (BM) (Vargo-Gogola and Rosen, 2007). The basement membrane, composed of extracellular matrix (ECM) proteins, functions to provide cells with anchorage and important instructive proliferative and survival cues (Ghajar and Bissell, 2008). Interaction of normal and neoplastic mammary epithelial cells with the BM and other ECM components plays important roles in mammary gland homeostasis and in the progression of breast cancer (Polyak and Kalluri, 2010).



**Illustration 1.1: The mammary gland**

The mammary gland is a branching organ comprised of ducts and lobules lined with a bi-layered epithelium. The inner layer of the mammary ducts are lined by luminal cells, while the outer layer is lined by basal cells that are in direct contact with the basement membrane (BM). BM is basement membrane.

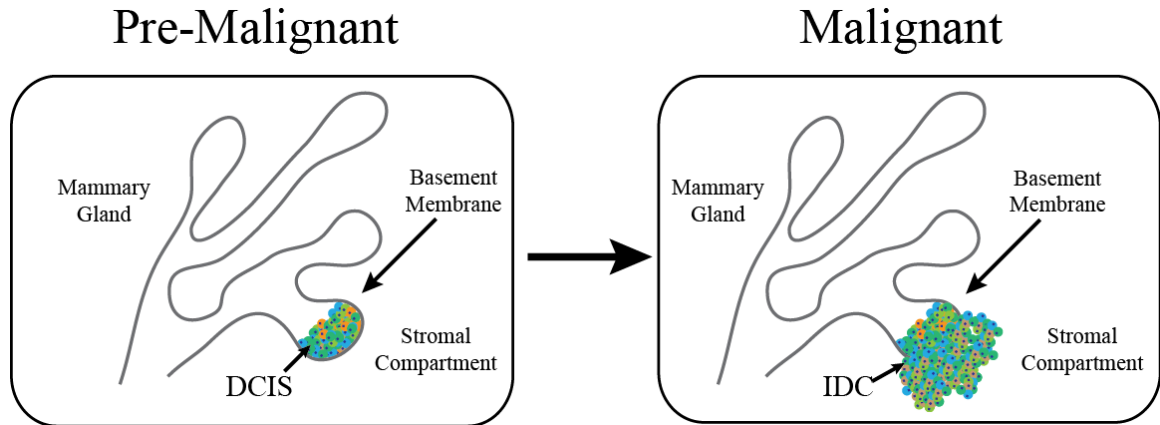
Mammary gland homeostasis is characterized by the expansion and involution of the mammary branches during development and the female reproductive life cycle (Ghajar and Bissell, 2008; Llobet-Navas et al., 2014). Deregulated homeostasis within the mammary gland can result in breast cancer, which is the unconstrained proliferation of oncogenically transformed mammary epithelial cells (Ercan et al., 2011). The oncogenic transformation of epithelial cells is the result of genetic and epigenetic alterations that lead to the acquisition of aberrant cell biological features that conflict with normal homeostatic control (Hanahan and Weinberg, 2000). These alterations enable cells to proliferate uncontrollably independent of extrinsic growth signals, evade normal safeguards that would lead to programmed cell death, and survive or become dormant in hostile external conditions, in addition to many other features (Hanahan and Weinberg, 2000; Hedley and Chambers, 2009). Further genetic or epigenetic changes in neoplastic cells drives cancer progression along a continuum, broadly demarcated into the pre-malignant, malignant, and metastatic stages.

#### *Pre-malignant and malignant breast cancer*

Upon oncogenic transformation of mammary epithelial cells, early breast cancer progression follows two distinct stages: the pre-malignant stage and the malignant stage (Illustration 1.2). Pre-malignant breast cancer is characterized by the proliferation and expansion of neoplastic cells into the mammary ducts without evidence of their penetration through the encapsulating basement membrane (Schatten, 2013). Based on their expansion but confinement to within the mammary ducts, pre-malignant breast cancers are also referred to as ductal carcinoma *in situ* (DCIS) (Espina and Liotta, 2011).

While pre-malignant cancer cells demonstrate many hallmarks of transformed cells, the benign nature of pre-malignant lesions is due to their confinement within the mammary gland and lack of escape into the surrounding stromal compartment (Chaffer and Weinberg, 2011). The confinement of pre-malignant lesions by the BM prevents them from growing to a large size, infiltrating into the surrounding tissue, and spreading to sites outside of the primary site (Talmadge and Fidler, 2010). The penetration of breast cancer cells through the BM and out of the mammary epithelium defines the transition from the pre-malignant stage to the malignant stage, and the biological factors that regulate progression to this stage remain largely unknown (Polyak, 2010).

Malignant breast cancers are defined by evidence of cancer cell penetration through the BM of the mammary epithelium and into the surrounding stromal tissue. Due to their invasion out of the ducts and into the stromal compartment, malignant breast cancers are also referred to as invasive ductal carcinomas (IDC) (Ali and Coombes, 2002). Exit of cancer cells through the BM and into the stromal compartment reflects both their capacity to break down and invade through this barrier and to proliferate and survive without proper engagement to extracellular matrix proteins that are present outside of the mammary gland (Pavlova et al., 2013; Wan et al., 2013). An understanding of the molecular differences between pre-malignant and malignant cancer cells could lead to the identification of factors that enable the transition to the malignant phase.



**Illustration 1.2: Pre-malignant and malignant breast cancer**

Pre-malignant breast cancer, also known as ductal carcinoma *in situ* (DCIS) is the precursor lesion to malignant breast cancer, also known as invasive ductal carcinoma (IDC). Pre-malignant breast cancer cells are confined within the basement membrane (BM) surrounding the mammary epithelium. The transition from pre-malignant to malignant breast cancer occurs once cancer cells have broken down and invaded through the BM and into the surrounding stromal micro-environment.

### *Primary tumor formation*

The penetration of malignant breast cancer cells through the basement membrane enables them to expand further and to develop into tumors of a substantial size (Talmadge and Fidler, 2010). The hallmarks of primary tumors include their large size, infiltration into the surrounding tissue, and capacity to enable vascularization to supply them with adequate nutrients and oxygen through the recruitment of vessel-forming endothelial cells (Folkman, 2002; Talmadge and Fidler, 2010). Cancer cells within primary tumors can also co-opt cells from the surrounding microenvironment, including fibroblasts, endothelial cells, and hematopoietic cells to induce them to produce important growth signals that further promote tumor maintenance and progression (Hanahan and Weinberg, 2000). The capacity of cancer cells originating from malignant tumors to expand and invade into the surrounding tissue is one feature that predisposes them to spread to further distant sites and to re-initiate tumors in a metastatic organ (Talmadge and Fidler, 2010).

### *The metastatic cascade*

The major cause of death in patients with breast cancer is metastasis, which involves the systemic spread of cancer cells and subsequent re-formation of tumors at distant organ sites (Hedley and Chambers, 2009). This process involves a cascade of several rate-limiting steps, all of which must be completed in order for a successful metastatic event to occur (Illustration 1.3). The execution of this complex series of events in concert makes metastasis highly inefficient and a significant bottleneck in the progression of cancer (Langley and Fidler, 2011).

The first step of the metastatic cascade involves the detachment of cancer cells from the primary tumor and their entrance into the circulatory or lymphatic system. The entrance of cancer cells into the blood or lymphatics, known as intravasation, requires cancer cells to invade through the endothelial or smooth muscle layers that surround the lumen of vessels (Chaffer and Weinberg, 2011). Because intravasation can only occur if cancer cells have first invaded through the basement membrane and into the stromal compartment, only malignant breast cancers are thought to be able to metastasize (Ali and Coombes, 2002). Once inside the lymphatics or blood vessels, cancer cells gain access to a conduit from which they can access every organ in the body.

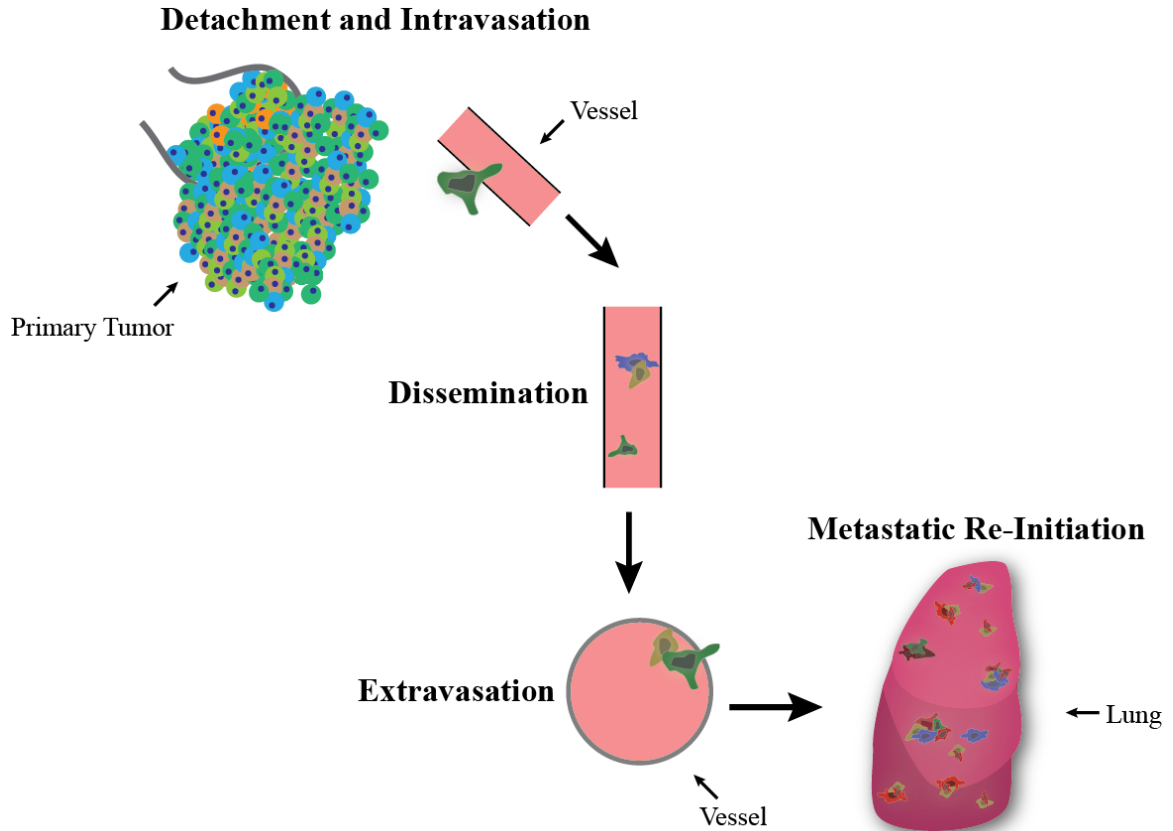
The hematogenous or lymphatic dissemination of cancer cells exposes them to multiple barriers that must be overcome in order for them to survive (Mehlen and Puisieux, 2006). While in the circulation or lymphatics, cancer cells lack their previous attachment to important ECM components and nearby cells (Simpson et al., 2008), are targets of engulfment by phagocytic macrophages (Wan et al., 2013), and can become mechanically destroyed due to tortuous or small vessels (Wirtz et al., 2011). All of these features represent potential activators of programmed cell death that must be overcome by cancer cells. Although it is unknown how long cancer cells can survive before progressing to the next step along the metastatic cascade, circulating cancer cells have been reproducibly detected in the blood of patients with breast cancer (Yu et al., 2011).



Metastatic progression of disseminated cancer cells continues through the re-entrance of cancer cells into distant organs and out of the circulation, a process known as extravasation. Extravasation occurs after the lodging, or embolization, of cancer cells in the endothelial linings of ectopic organs (Wirtz et al., 2011). In addition to physical lodging, cancer cell binding to endothelial cells through receptor-ligand interactions is thought to be required for their extravasation out of the circulatory system and back into the parenchyma of distant organs (Wirtz et al., 2011). However, different organs may have distinct requirements for extravasation that vary as a result of organ-specific features of the local vasculature. For example, the lungs are lined with an endothelial barrier that is highly restrictive to extravasation (Padua et al., 2008; Reymond et al., 2013), while the sinusoidal endothelium of the liver is more permissive to cell seeding (Nguyen et al., 2009a). It is thought that these different requirements for extravasation may contribute to organ-selectivity by metastatic cells (Wirtz et al., 2011).

The final rate-limiting step of the metastatic cascade, tumor re-initiation, occurs once cancer cells have extravasated and arrived in the parenchyma of a distal organ. During this stage, cancer cells arrive in an ectopic organ microenvironment that often lacks important proliferative and survival signals that would enable the formation of colonies (Fidler, 2003, 2011; Giancotti, 2013). The absence of cues from the ECM can lead to cell cycle arrest or a protracted period of cellular dormancy, whose length can vary from only a short period of time to up to several years (Wan et al., 2013). In order to re-initiate a tumor, a small number of cancer cells (or potentially only one cancer cell) must regain proliferative capacity and re-initiate a new colony in this non-permissive context. It is

thought that remodeling of the local microenvironment through the secretion or modulation of extracellular proteins contributes to the re-activation of disseminated cancer cells (Wan et al., 2013). The capacity of limiting numbers of cancer cells to overcome this final barrier through largely unknown mechanisms is the gateway to metastatic tumor re-initiation.



**Illustration 1.3: The metastatic cascade**

The metastatic cascade is a series of complex rate-limiting steps that all must be completed by cancer cells in order for a successful metastatic event to occur. Some of these steps include detachment from the primary tumor site, intravasation into the circulation or lymphatics, dissemination, extravasation of cancer cells out of the circulation and into the tissue of an ectopic organ, and the re-initiation of metastatic tumors.

### *Primary and Metastatic Relapse*

While many treatments for breast cancer can reduce tumor burden to undetectable levels, even the most effective treatments are limited in their ability to completely eliminate every cancer cell (Chaffer and Weinberg, 2011). As a result, it is thought that the capacity of cancer cells to re-initiate tumors from small numbers of residual cells exposes almost all breast cancer therapies to an inherent risk of relapse (Hu et al., 2013). Breast cancer relapse can occur at either at the primary or a metastatic site regardless of the stage of the patient at the time of therapy (Hedley and Chambers, 2009). For example, a patient treated for localized primary breast cancer can succumb to metastatic relapse in the lung years after her initial treatment due to the presence of a small numbers of cells escaping therapy and re-initiating tumors at a later time (Hedley and Chambers, 2009). The factors contributing to the latency observed in relapse are generally unknown, but they are thought to reflect the potential of cancer cells to remain dormant and resistant to therapy for a protracted period of time (Wan et al., 2013). Xenograft models that assess the capacity of cancer cells to propagate, or re-initiate, tumor growth in primary or metastatic sites model this process by mimicking the physiological formation of tumors from small numbers of cells (Magee et al., 2012).

### **Breast cancer: therapeutic intervention**

Breast cancer therapy is diverse in part due to the biological heterogeneity of this disease (Byler et al., 2014). Many subtypes of breast cancer exist with distinct clinical behavior, and these subtypes can be distinguished based on morphology or molecular analysis (Gusterson, 2009; Schatten, 2013). Breast cancer subtypes can broadly be distinguished

into two groups based on their expression of the estrogen hormone receptor (Putti et al., 2005). Estrogen Receptor-positive (ER-positive) breast cancers are often dependent on estrogen for their growth, maintenance, and progression (Kim et al., 2012). Estrogen Receptor-negative (ER-negative) breast cancers are generally more aggressive and metastasize more frequently than ER-positive cancers (Putti et al., 2005; Thomas and Gustafsson, 2011). In addition to being less aggressive, the hormone-dependence of ER-positive tumors can be exploited therapeutically through targeting with anti-estrogen therapy (Andre and Pusztai, 2006). Anti-hormone therapies such as tamoxifen have led to a significant decrease in the number of deaths incurred by ER-positive breast cancer and greatly improved the management of this disease (Briskin, 2013; Powles, 2002). In contrast, targeted therapies for ER-negative disease (which represent approximately 30% of all breast cancers (Zhang et al., 2011b)) are limited (Putti et al., 2005). An understanding of the biological mediators that drive ER-negative breast cancer has the potential to reveal novel targets for therapeutic intervention and is of great interest to the biomedical research community.

## **Breast cancer: experimental models**

### *Tumor re-initiation*

The tumor-forming capacity of cancer cells can be experimentally modeled with xenograft ‘tumor re-initiation assays,’ which assess the potential of human cancer cells to re-initiate macroscopic tumors in a secondary host upon implantation into a particular organ (such as the mammary gland of a mouse) (Williams et al., 2013). Tumor re-initiation assays physiologically mimic the formation of tumors from small numbers of

cells, which occurs during primary tumor formation, metastatic re-initiation, and relapse. It is thought that the capacity of cancer cells to give rise to tumors in a xenograft model is reflective of features important to each of these processes (O'Brien et al., 2010).

The identification and characterization of cancer populations with tumorigenic potential in xenograft models has been subject to numerous scientific investigations (Nguyen et al., 2012; O'Brien et al., 2010). The tumor-forming potential of human cancer cells can vary tremendously in such models, with studies reporting tumors being formed from just one cell to upwards of hundreds of thousands of cells (Ishizawa et al., 2010; Quintana et al., 2008). This variability has been attributed to the type of cancer studied, the xenograft model used, or whether the source was a primary tumor, metastatic lesion, or cell line (Buchstaller et al., 2012; Quintana et al., 2008). Even different sub-populations of cancer cells derived from the same tumor source may have substantial variability in their tumorigenic capacity (Al-Hajj et al., 2003; Prince et al., 2007). For example, several investigators have employed cell surface markers to prospectively isolate different sub-populations of cells that are then subject to phenotypic assays to determine their tumorigenic potential (Chaffer et al., 2013; Cordenonsi et al., 2011; De Vito et al., 2012; He et al., 2013; Shimono et al., 2009; Vermeulen et al., 2010; Visvader and Lindeman, 2012). While many of these studies have been greatly informative, most rely on the use of one or multiple markers to sort different populations of cells and then separately test their tumorigenic potential (Owens and Naylor, 2013). The discovery of important biological mediators of tumor-forming capacity could be aided through the development

of an alternative systematic approach to select and characterize breast cancer populations with enhanced tumorigenicity.

While characterization of mediators of tumorigenic capacity has important consequences for the understanding of primary disease, 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 also a question of considerable interest (Magee et al., 2012). While tumor-forming potential in the primary and metastatic site appear to share some similarities, the molecular features common to highly tumorigenic cells and metastatic cells has not been systematically studied (Magee et al., 2012; Vanharanta and Massague, 2013), and further characterization of both of these processes remains an area of active investigation (Chakrabarti et al., 2012; Korpala et al., 2011; Piao et al., 2014).

### *Experimental metastasis*

Multiple steps of the metastatic cascade can be examined experimentally using mouse xenograft models. The most complete xenograft models of breast cancer metastasis allow metastasis to occur in a mouse following the orthotopic implantation of primary tumors into the mammary gland (Hedley and Chambers, 2009). Cancer cells that metastasize to distant organs from this context must complete all steps of the metastatic cascade, including intravasation from the primary site, dissemination into the circulation, extravasation into distal organs, and the eventual re-initiation of tumors (Khanna and Hunter, 2005). While they have the advantage of encompassing all steps of the metastatic cascade, orthotopic models of metastasis are limited in their ability to draw conclusions

about specific steps (Khanna and Hunter, 2005). Specific stages of the metastatic cascade can be studied by introducing cancer cells into mice along particular stages of interest. For example, intravasation can be quantified by measuring the numbers of cancer cells that have disseminated into the circulation from a primary tumor implant (Gupta et al., 2007). Metastatic extravasation and colonization can be studied by injecting cells into the venous circulation of mice and evaluating the numbers of cells that have entered a metastatic organ and the numbers of colonies formed, respectively (Khanna and Hunter, 2005). While many of these experimental assays have been conducted and have yielded important insights into the underlying biology of multiple steps of the metastatic cascade, sufficient understanding of metastatic re-initiation, the final step of the metastatic cascade, is still lacking.

#### *In vivo selection*

*In vivo* selection is an experimental technique that can be used to select for populations enriched for a particular phenotype of interest. Populations with enhanced fitness for malignant phenotypes have been derived from heterogeneous cancer populations using this method (Fidler, 2003). Several investigators have employed *in vivo* selection of human cancer populations to enrich for cancer cells that demonstrate enhanced metastatic capacity relative to their parental populations in xenograft models of experimental metastasis (Clark et al., 2000; Fidler, 1973; Gumireddy et al., 2009; Gumireddy et al., 2007; Kang et al., 2003; Minn et al., 2005; Pencheva et al., 2012; Tavazoie et al., 2008; Xu et al., 2008). Comparison of such highly metastatic derivatives to their parental populations has led to the identification of functional molecular regulators of metastasis



with prognostic and therapeutic significance in multiple human cancers (Clark et al., 2000; Gumireddy et al., 2009; Gupta et al., 2007; Kang et al., 2003; Nguyen et al., 2009b; Pencheva et al., 2012; Png et al., 2012; Tavazoie et al., 2008; Zhang et al., 2013a).

In the study of breast cancer metastasis, two breast cancer cell lines, MDA-MB-231 (Cailleau et al., 1978) and CN34 (Gomis et al., 2006), were previously subjected to *in vivo* selection to generate the lung-metastatic (LM) derivative populations MDA-LM2 (Minn et al., 2005) and CN34-LM1a (Tavazoie et al., 2008), respectively. These populations were selected for their capacity to colonize the lungs upon intravenous inoculation into immunodeficient mice, and both the MDA-LM2 and CN34-LM1a demonstrate robust capacity to metastasize to the lungs in xenograft models (Png et al., 2012; Tavazoie et al., 2008). Systematic molecular characterization of LM populations has led to the identification of molecular regulators of specific features of the metastatic process, including cell invasion (Tavazoie et al., 2008), endothelial recruitment (Png et al., 2012), and trans-endothelial migration (Padua et al., 2008). Molecular regulators identified using this approach include the microRNAs miR-335 and miR-126 (Tavazoie et al., 2008), which suppress metastasis by repressing the expression of the effector genes *SOX4*, and *TNC* (in the case of miR-335) (Tavazoie et al., 2008) and *IGFBP2*, *MERTK* and *PITPNC1* in the case of (miR-126) (Png et al., 2012).

While important regulators of sequential steps of the metastatic cascade have been identified employing *in vivo* selection for metastasis, due to selection for this multi-step phenotype, characterizing the molecular basis of a specific step *a priori*, such as tumor

re-initiation, remains challenging. The identification of the molecular determinants that mediate tumor re-initiation in both the primary and metastatic settings could further advance the understanding of both of these processes.

### **Specific aims**

The goal of this thesis is to contribute to the understanding of the cellular and molecular properties that enable small number of cancer cells to form tumors in the primary or metastatic microenvironment. The first phase of this study describes the development of a system to select and characterize cells with enhanced tumorigenic capacity in a xenograft model. Populations of breast cancer cells with enhanced tumorigenic capacity were derived through the application of selective pressure for tumor formation from low cell numbers (Chapter II). These tumorigenic-enriched populations were extensively characterized and found to give rise to xenograft tumors at greater frequency in multiple organ microenvironments.

The second phase of this study applied a systematic approach to identify genes that promote tumor re-initiation and to uncover the cell biology governed by these genes (Chapter III). Three genes were identified– *LAMA4*, *FOXQ1*, and *NAPIL3*–to be expressed at greater levels by highly tumorigenic and highly metastatic cells and one of these genes *LAMA4*, was selected for further study. Functional characterization of *LAMA4* revealed it to promote tumor re-initiation in multiple microenvironments, clonal expansion of cancer cells in the absence of substratum attachment *in vitro*, and disseminated metastatic proliferation and colony formation.

The third phase of this study presents clinical evidence to support the significance of these findings to human breast cancer progression (Chapter IV). Multiple datasets revealed that increased expression of *LAMA4* marked the transition from pre-malignant to malignant breast cancer, and revealed that higher expression of *LAMA4* in ER-negative tumors was associated with poor relapse-free survival. Further discussion of the major findings of this thesis and future directions are presented in the final chapter (Chapter V).

## **CHAPTER II:**

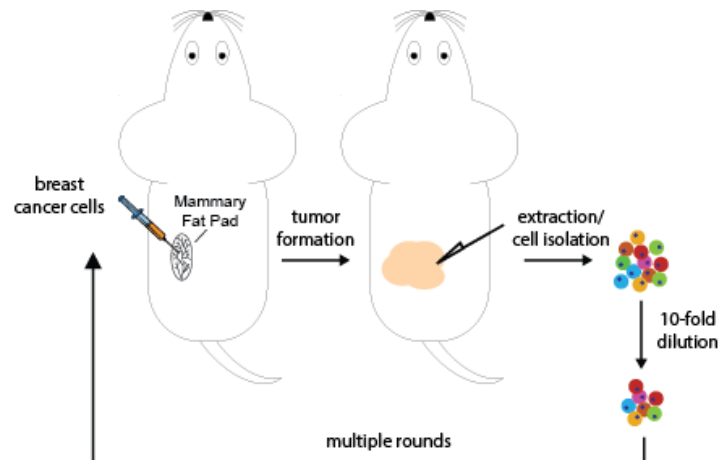
### **SELECTION AND CHARACTERIZATION OF BREAST CANCER CELLS WITH ENHANCED TUMORIGENIC CAPACITY**

The primary goal of this study is to better understand the biology through which small numbers of cancer cells give rise to tumors. As the first step to identify cellular and molecular determinants associated with this process, this chapter describes the development of an experimental system to select and characterize cancer cells with enhanced tumorigenic capacity.

#### ***In vivo* selection for tumor re-initiation yields tumorigenic-enriched (TE) derivatives**

In order to study the biology that governs breast cancer tumor re-initiation, *in vivo* selection was used to derive populations of human breast cancer cells with enhanced tumor-forming capacity. Independent human breast cancer cell lines, the established MDA-MB-231 (MDA-231) line (Cailleau et al., 1978; Minn et al., 2005) and the minimally passaged CN34 line (Tavazoie et al., 2008), were subjected to *in vivo* selection. These cell lines were selected for study in part due to their Estrogen Receptor-negative (ER-negative) status (Bos et al., 2009). ER-negative breast cancers are generally more aggressive and have fewer therapeutic options relative to ER-positive breast cancers (Putti et al., 2005). Selective pressure for tumor re-initiation was applied to the MDA-231 and CN34 cell lines by injecting increasingly limiting numbers of these populations

orthotopically into the mammary fat pads of immunodeficient mice in order to generate xenograft tumors over successive rounds of serial dilution (Figure 2.1).

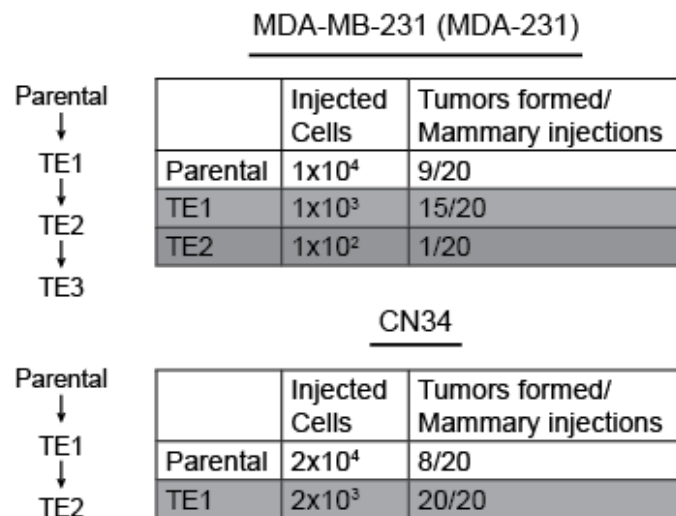


**Figure 2.1: *In vivo* selection for tumor re-initiation**

Schematic of the strategy of *in vivo* selection used to derive populations with enhanced tumorigenic potential. 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.

During the first round of *in vivo* selection for tumor re-initiation, both the MDA-231 and CN34 cell lines gave rise to tumors at non-saturating (less than 100-percent) frequency at the initial cell doses used (10,000 or 20,000 cells, for the MDA-231 or CN34 cell lines, respectively) upon injection into the mammary fat pads of immunodeficient mice (Figure 2.2). The non-saturating frequency of tumor re-initiation suggested that selective pressure for tumor formation existed at this dose.

Tumors that were formed during this first round of *in vivo* selection were dissociated to yield 1<sup>st</sup>-generation MDA-231 and CN34 tumorigenic-enriched (TE) derivatives MDA-TE1 and CN34-TE1, which were expanded and propagated *in vitro*. A second round of *in vivo* selection was conducted by re-injecting these 1<sup>st</sup>-generation derivatives at a 10-fold lower dose (1,000 or 2,000 cells, for MDA-TE1 and CN34-TE1, respectively) and dissociating the resulting tumors to yield 2<sup>nd</sup>-generation TE derivatives MDA-TE2 and CN34-TE2 (Figure 2.2). Notably, during the second round of *in vivo* selection, the frequency of tumors formed by the TE1 derivatives was greater than the initial frequency demonstrated by the parental populations (15/20 vs. 9/20 and 20/20 vs. 8/20 for the MDA-231 or CN34 cell lines, respectively) even though 10-fold fewer cells were injected (Figure 2.2). These initial results suggested that enrichment for tumor-forming capacity had occurred after only one round of *in vivo* selection. For the MDA-231 cell line, a third round of *in vivo* selection was performed to generate the MDA-TE3 derivative population, while the CN34 cell line only underwent two rounds of *in vivo* selection to generate the CN34-TE2 derivative population. Subsequent experiments focused on MDA-TE3 and CN34-TE2.



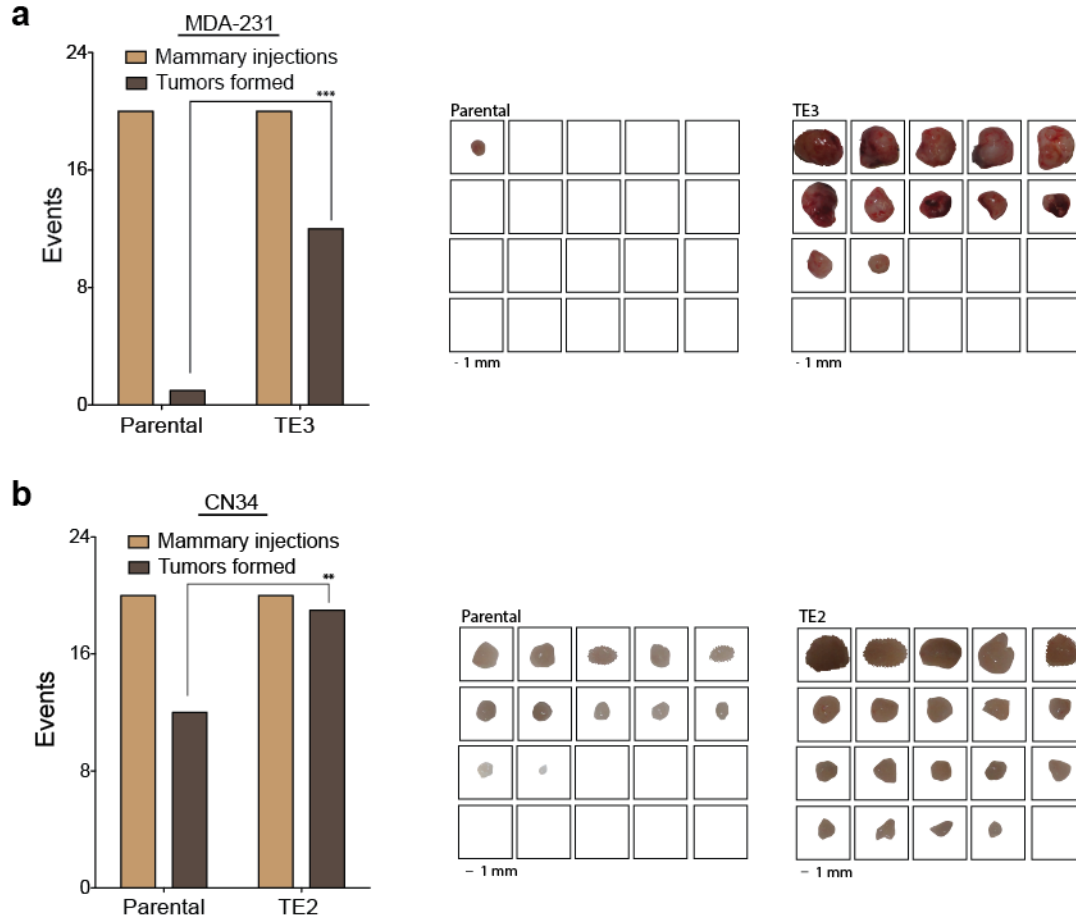
**Figure 2.2: Selection for multiple tumorigenic-enriched (TE) derivative populations**

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).

**Tumorigenic-enriched (TE) derivatives demonstrate enhanced orthotopic tumor re-initiation capacity relative to their parental populations**

Functional experiments revealed that tumorigenic-enriched derivatives MDA-TE3 and CN34-TE2 gave rise to xenograft tumors at significantly greater frequency (up to 12-fold) 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 (Figure 2.3a-b). These results demonstrate that populations of cells with significantly enhanced tumor re-initiation capacity can be derived from human breast cancer populations through *in vivo* selection.



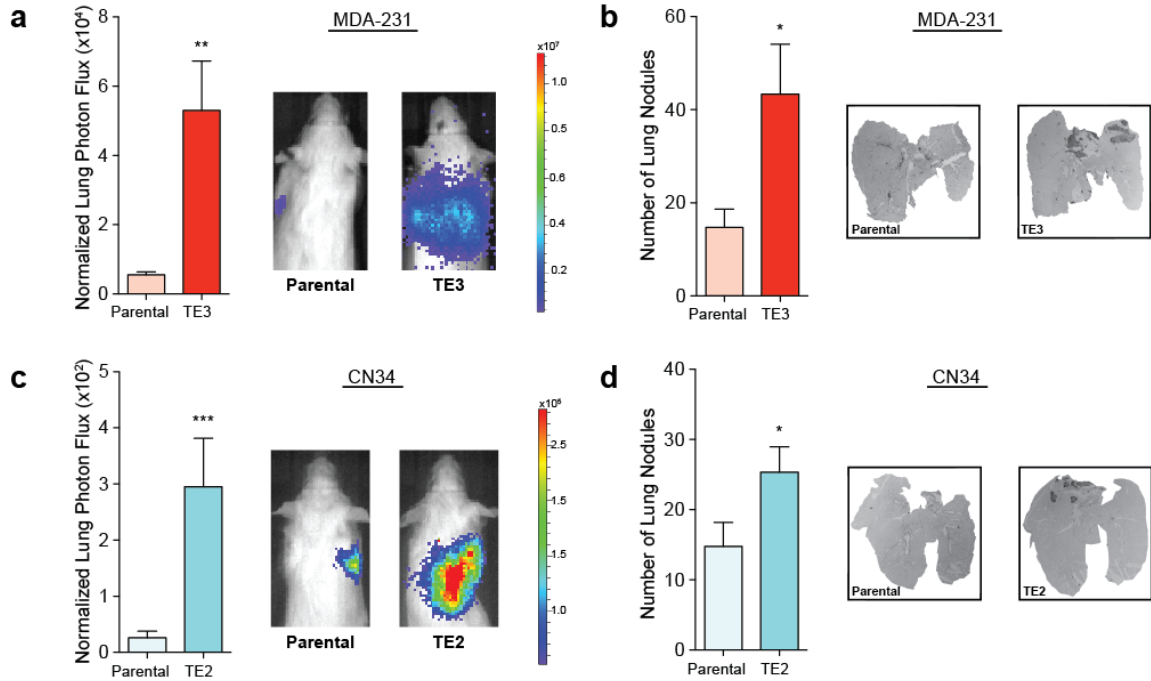


**Figure 2.3: Tumorigenic-enriched (TE) derivatives demonstrate enhanced tumor re-initiation capacity**

(a) MDA-TE3 cells exhibited enrichment for tumor re-initiation capacity as compared to MDA-parental cells following orthotopic injection of  $1 \times 10^2$  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 (represented as open squares, right). (b) CN34-TE2 cells exhibited enrichment for tumor re-initiation capacity as compared to CN34-parental cells following orthotopic injection of  $2 \times 10^3$  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 (represented as open squares, right).  $**P < 0.01$ ,  $***P < 0.001$  were obtained using one-sided Fisher's exact test (a-b).

### **Tumorigenic-enriched (TE) derivatives robustly colonize multiple organs relative to their parental populations**

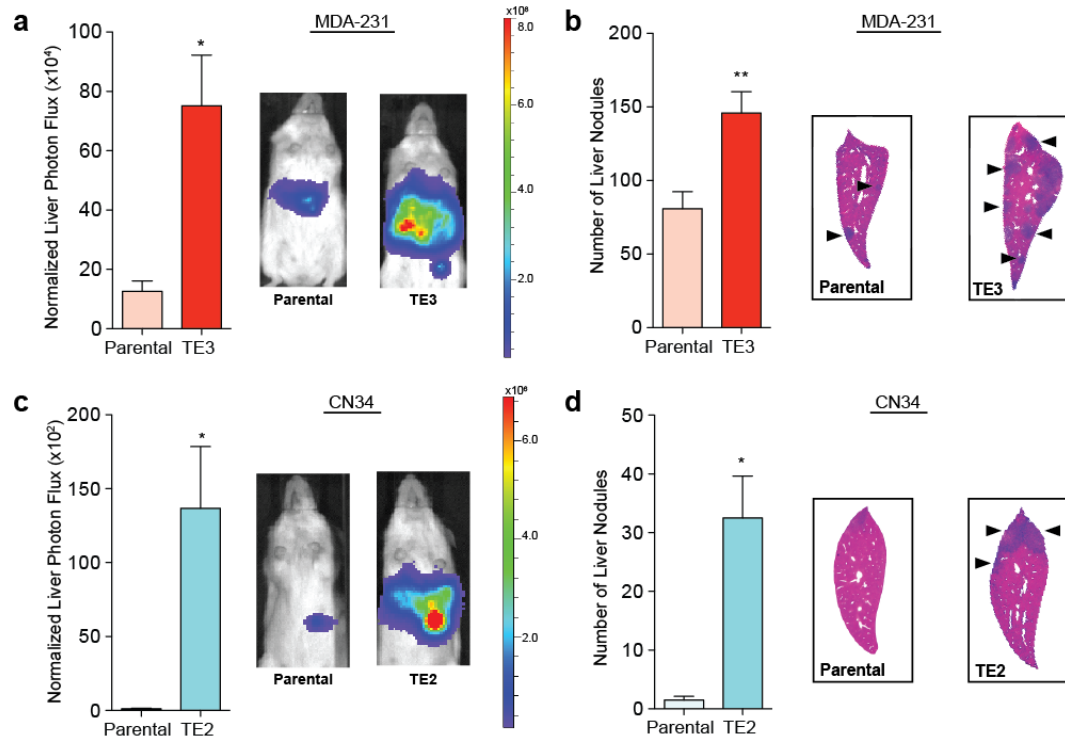
A fundamental question of great interest in cancer biology is the relationship between primary tumor-forming potential and metastatic activity, (Diehn and Majeti, 2010; Magee et al., 2012; Rhim et al., 2012; Song et al., 2013). Highly tumorigenic cells have been posited to be more able to metastasize (Brabletz et al., 2005), while other studies have suggested these processes to be independent (Brabletz, 2012; Visvader and Lindeman, 2012). As an important first step to address this issue, the capacity of tumorigenic-enriched derivatives to re-initiate tumors in an ectopic metastatic site was assessed. When injected directly into the lung parenchyma, TE derivatives gave rise to a greater number of macroscopic tumors in the lungs relative to their respective parental populations (Figure 2.4a-d), demonstrating that they were enriched in ectopic tumor re-initiation capacity in addition to orthotopic tumor re-initiation capacity.



**Figure 2.4: Tumorigenic-enriched cells demonstrate enhanced ectopic tumor re-initiation capacity relative to their parental populations**

**(a-b)**  $1 \times 10^3$  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)**.  $n = 3$  independent lungs. **(c-d)**  $1 \times 10^5$  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)**. On day 56 lungs were sectioned, vimentin stained, and the number of macroscopic nodules per lung was counted **(d)**.  $n = 8$  (Parental),  $n = 9$  (TE2) independent lungs.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  were obtained using a one-sided Mann-Whitney test **(a, c)** or a one-sided Student's t-test **(b, d)**. All data are represented as mean + S.E.M.

The enhanced ectopic tumor re-initiation capacity displayed by the TE derivatives suggested that these populations might also demonstrate enhanced colonization of organs during metastasis. Because organ-specific features such as the endothelial tight-junctions of the lungs represent significant barriers to cell extravasation (Padua et al., 2008; Reymond et al., 2013), TE derivatives would be expected to display 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 TE cells were assessed for their capacity to metastasize to the liver—an organ whose fenestrated sinusoidal capillaries more readily permit cell seeding (Nguyen et al., 2009a)—TE cells gave rise to a significantly greater number of macroscopic tumors in the liver relative to their respective parental populations (Figure 2.5a-d).



**Figure 2.5: Tumorigenic-enriched cells demonstrate enhanced liver metastasis relative to their parental populations**

(a-b)  $1 \times 10^5$  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 (a).  $n = 5$  independent mice. On day 35 livers were sectioned, H&E stained, and number of macroscopic nodules was counted (b). Arrows, representative nodules.  $n = 5$  independent livers. (c-d)  $2 \times 10^5$  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 (c).  $n = 4$  independent mice. On day 56 livers were sectioned, H&E stained, and number of macroscopic nodules was counted (d). Arrows, representative nodules.  $n = 4$  independent livers. \* $P < 0.05$ , \*\* $P < 0.01$  were obtained using a one-sided Mann-Whitney test (a, c) or a one-sided Student's t-test (b, d). All data are represented as mean + S.E.M.

Taken together, these results demonstrate that *in vivo* selection for orthotopic tumor re-initiation is sufficient to co-select for enhanced colonization of multiple organs. The generation of multiple tumorigenic-enriched derivatives with enhanced tumorigenic capacity supports the use of this system to interrogate the cellular and molecular features important to this process.

### **Characterization of tumorigenic-enriched derivatives**

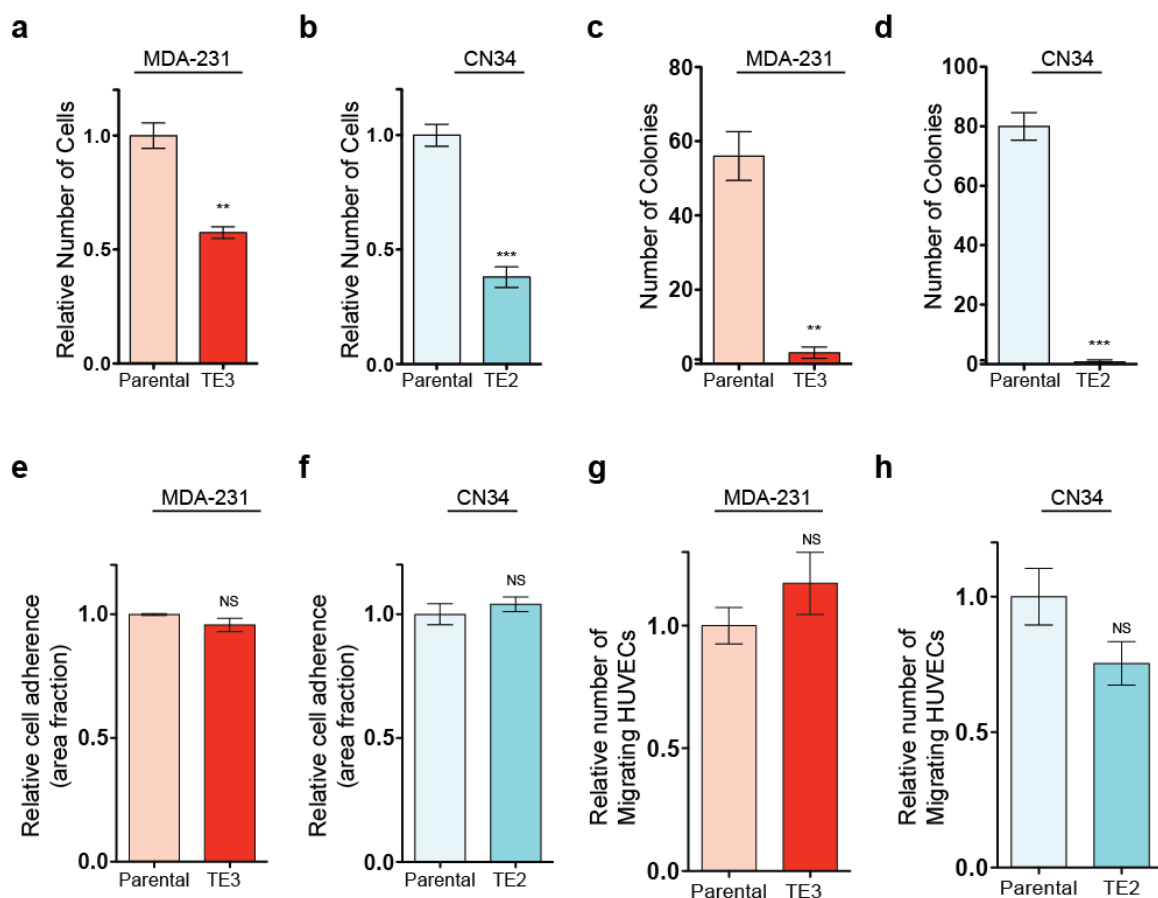
Phenotypic differences between tumorigenic-enriched cells and their parental populations would be expected to reveal biology important to the process of tumor re-initiation. Experiments were conducted to compare TE derivatives with their parental populations across multiple *in vitro* phenotypes typically considered to confer a pro-tumorigenic advantage. Surprisingly, TE derivatives proliferated and formed colonies *in vitro* to a lesser extent than their parental populations upon standard adherent cell culture conditions (Figure 2.6a-d), did not demonstrate significant differences in their capacity to attach to tissue culture plates *in vitro* (Figure 2.6e-f), and did not recruit a greater number of endothelial cells relative to their parental populations *in vitro* (Figure 2.6g-h). These findings suggested that the enhanced *in vivo* tumor-forming capacity demonstrated by TE derivatives was independent of traditional *in vitro* phenotypes thought to be associated with enhanced tumorigenic potential. Additionally, immuno-phenotypic characterization of the MDA-TE3 or CN34-TE2 derivatives did not reveal enrichment of CD44<sup>+</sup>/CD24<sup>-</sup> markers relative to their respective parental populations (Figure 2.7a,b). The absence of enhancement of these multiple phenotypes in tumorigenic-enriched derivatives relative to their parental populations suggested that a molecular approach to systematically identify

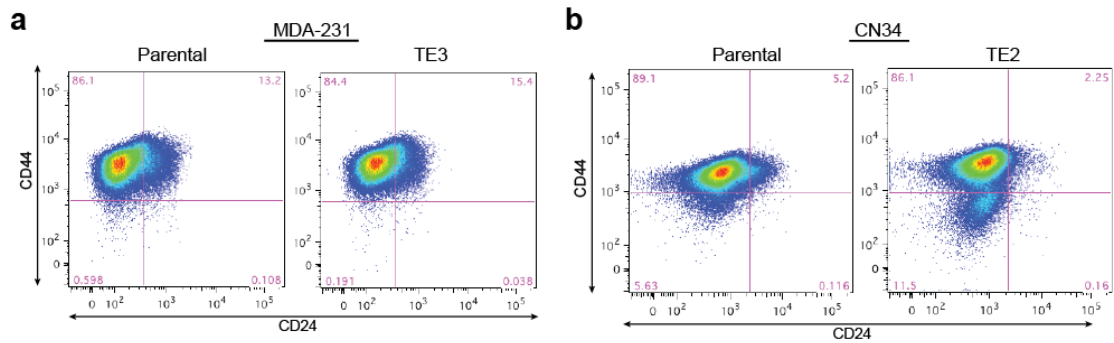
differences between these cell populations might be more revealing of important determinants of tumor re-initiation in breast cancer.

**Figure 2.6: *In vitro* characterization of tumorigenic-enriched derivatives**

(a)  $2.5 \times 10^4$  MDA-parental or MDA-TE3 cells were seeded into 6-well adherent tissue-culture plates. The number of cells per well on day 5 was counted and normalized to cell counts on day 1.  $n = 3$  independent wells. (b)  $2.5 \times 10^4$  CN34-parental or CN34-TE2 cells were seeded into 6-well adherent tissue-culture plates. The number of cells per well on day 5 was counted and normalized to cell counts on day 1.  $n = 3$  independent wells. (c)  $1 \times 10^2$  MDA-parental or MDA-TE3 cells were seeded into 10cm adherent tissue-culture plates. The number of colonies per well on day 14 was counted upon staining with crystal-violet.  $n = 3$  independent plates. (d)  $1 \times 10^2$  CN34-parental or CN34-TE2 cells were seeded into 10cm adherent tissue-culture plates. The number of colonies per well on day 14 was counted upon staining with crystal-violet.  $n = 3$  independent plates. (e) Assessment of cell attachment to adherent tissue culture plates of MDA-Parental cells compared to MDA-TE3 cells.  $n = 3$  independent wells. (f) Assessment of cell attachment to adherent tissue culture plates of CN34-Parental cells compared to CN34-TE3 cells.  $n = 6$  independent wells. (g) Endothelial recruitment assay comparing the relative capacity of MDA-Parental cells and MDA-TE3 cells recruit Human Vein Endothelial Cells (HUVECs).  $n = 6$  independent trans-well inserts. (h) Endothelial recruitment assay comparing the relative capacity of CN34-Parental cells and CN34-TE3 cells recruit Human Vein Endothelial Cells (HUVECs).  $n = 8$  independent trans-well inserts. NS is not significant.  $**P < 0.01$ ,  $***P < 0.001$  were obtained using a two-sided student's t-test (a-h). All data are represented as mean  $\pm$  S.E.M.







**Figure 2.7: Immuno-phenotypic characterization of tumorigenic-enriched derivatives**

(a) MDA-parental and MDA-TE3 populations were stained with antibodies specific to the cell surface markers CD44 and CD24 and analyzed using flow cytometry. For the MDA-parental population, 86.1% of the cells were CD44<sup>+</sup>CD24<sup>-</sup>. For the MDA-TE3 population, 84.4% of the cells were CD44<sup>+</sup>CD24<sup>-</sup>. (b) CN34-parental and CN34-TE2 populations were stained with antibodies specific to the cell surface markers CD44 and CD24 and analyzed using flow cytometry. For the CN34-parental population, 89.1% of the cells were CD44<sup>+</sup>CD24<sup>-</sup>. For the CN34-TE2 population, 86.1% of the cells were CD44<sup>+</sup>CD24<sup>-</sup>.

## Summary of findings from Chapter II

This chapter describes the selection and characterization of breast cancer populations with enhanced tumor-forming capacity. Two ER-negative breast cancer cell lines, MDA-MB-231 and CN34, were subjected to *in vivo* selection for tumor re-initiation, leading to the generation of independent tumorigenic-enriched (TE) derivatives. Tumorigenic-enriched derivatives were propagated and expanded *in vitro*, and gave rise to tumors at greater frequency relative to their parental populations when injected orthotopically into immunodeficient mice at low cell numbers. TE derivatives also demonstrated enhanced capacity to re-initiate tumors relative to their parental populations when introduced into multiple ectopic organs. These findings demonstrated that propagatable populations of cells with enhanced tumor re-initiation capacity can be derived from breast cancer cell lines using *in vivo* selection and that selection for tumor-forming potential in the primary site is sufficient to co-select for enhanced tumor re-initiation capacity in multiple ectopic organs.

Tumorigenic-enriched derivatives were further characterized by subjecting them to multiple *in vitro* experimental assays thought to be associated with enhanced tumor-forming capacity. Surprisingly, tumorigenic-enriched derivatives performed in these assays to the same or lesser extent relative to their parental populations, suggesting that the enhanced tumorigenic potential of TE populations was independent of features represented by these assays. The following chapter presents a systematic effort to identify molecular differences that could distinguish tumorigenic-enriched cells from their

parental populations in order to uncover potential molecular regulators of tumor re-initiation (Chapter III).

# CHAPTER III

## IDENTIFICATION OF MOLECULAR DETERMINANTS OF TUMOR RE-INITIATION

This chapter describes a systematic approach to identify putative general promoters of tumor re-initiation through the molecular characterization of tumorigenic-enriched derivatives. Whole-genome transcriptomic profiling of tumorigenic-enriched and parental breast cancer populations was employed in order to identify candidate promoters of tumor re-initiation that function independent of the local microenvironment. This approach uncovered a set of genes expressed at greater levels by tumorigenic-enriched and metastatic derivative populations relative to their parental populations. One of these genes, *LAMA4*, was then extensively characterized. Further examination of *LAMA4* revealed it to promote tumor re-initiation in multiple organ microenvironments, clonal expansion in the absence of substratum attachment *in vitro*, and disseminated metastatic proliferation and colony formation *in vivo*.

### ***LAMA4*, *FOXQ1*, and *NAPIL3* display increased expression levels in tumorigenic-enriched and metastatic cells relative to parental populations**

Upon generating tumorigenic-enriched (TE) derivatives with enhanced tumor-forming capacity relative to their parental populations, a systematic approach was used to identify candidate genes that might govern this process. Promoters of tumor re-initiation would be

expected to have greater expression levels in TE cells relative to their parental populations. Whole-genome transcriptomic profiling of TE derivatives from both the MDA-231 and CN34 cell lines revealed a large set of genes (169, Table 3.1) that displayed at least 1.5-fold greater expression in these cells relative to their respective parental populations (Figure 3.1b). This set of genes represented potential drivers of tumor re-initiation, promoters of selective growth in the mammary gland, as well as passenger genes lacking functional significance (Figure 3.1a).

**Table 3.1: Genes with increased expression (1.5x higher) in tumorigenic-enriched (TE) derivatives relative to their parental populations.**

Gene ID	MDA-TE3 FC*	CN34-TE2 FC*	#	Gene ID	MDA-TE3 FC*	CN34-TE2 FC*	#
BCHE	7.17	11.14	1	TGFBR2	1.96	1.90	86
SLCO1B3	2.74	10.81	2	MAFF	1.76	2.08	87
TIMP3	1.89	10.17	3	IL7	2.16	1.65	88
BCHE	3.89	7.84	4	CD82	1.68	2.11	89
ALOX5AP	1.85	8.27	5	TGFBR2	1.79	2.00	90
PKIA	2.51	7.23	6	NAV3	1.52	2.25	91
C13orf15	2.16	7.38	7	GPR37	1.73	2.04	92
IL1A	2.67	6.75	8	STAT3	1.55	2.21	93
IL1B	5.66	3.11	9	NAV2	1.72	2.02	94
IGFBP1	1.97	6.52	10	CLDN7	1.96	1.77	95
HOXB5	1.99	6.48	11	TMTC2	1.76	1.95	96
TMEM163	2.18	5.86	12	HNMT	2.05	1.64	97
LOC653506	1.93	5.93	13	<b>FOXQ1</b>	<b>1.60</b>	<b>2.07</b>	<b>98</b>
PKIA	2.37	5.37	14	PUAEN2	1.67	1.99	99
SLIT2	1.91	5.76	15	EMP1	1.94	1.71	100
TMEM71	1.82	5.24	16	NFE2L3	1.79	1.86	101
ADAMTS1	2.49	4.13	17	SH3BGRL	1.53	2.12	102
IGFBP1	1.76	4.76	18	FRMD3	1.76	1.88	103
<b>NAP1L3</b>	<b>1.78</b>	<b>4.62</b>	<b>19</b>	COG6	1.77	1.87	104
CACNA2D4	3.60	2.72	20	CDKN1B	1.70	1.93	105
KLF2	1.95	4.37	21	PCMTD1	1.98	1.65	106
HIST1H1C	1.75	4.47	22	IRAK3	2.11	1.51	107
HIST2H2AA3	1.85	4.34	23	PAN3	1.55	2.06	108
<b>LAMA4</b>	<b>1.81</b>	<b>4.39</b>	<b>24</b>	GOLT1B	1.65	1.97	109
KHDRBS3	1.75	4.34	25	COG3	1.82	1.78	110
APOLD1	2.04	4.00	26	SLITRK5	2.06	1.54	111
LOC100008589	3.52	2.51	27	RGS17	1.73	1.85	112
METRNL	1.69	4.30	28	C11orf54	1.86	1.71	113
INSIG1	2.78	3.17	29	RCN3	1.79	1.78	114
BTG1	1.59	4.24	30	CR738291	1.56	2.00	115
KYNU	3.93	1.83	31	NDFIP2	1.57	1.99	116
GPC6	1.56	4.19	32	LOC442597	1.93	1.62	117
RASD1	1.70	4.05	33	NAMPT	1.72	1.82	118
HIST2H2AA3	1.66	4.07	34	ZMIZ2	1.60	1.93	119
PTGS2	2.21	3.50	35	TM7SF3	1.59	1.93	120
LMTK3	1.98	3.57	36	C13orf23	1.85	1.67	121
CRADD	3.45	2.08	37	PITPNC1	1.55	1.96	122
NSBP1	1.82	3.69	38	SATB2	1.75	1.76	123
TMEM154	3.25	2.22	39	KIAA1024	1.58	1.92	124
KHDRBS3	1.95	3.44	40	ARMCX3	1.53	1.96	125
PRDM8	1.51	3.81	41	HNMT	1.95	1.54	126
C21orf7	2.76	2.54	42	ERCC5	1.74	1.75	127
TMEM154	3.08	2.21	43	FBXO33	1.56	1.93	128
ARHGDIB	1.65	3.59	44	DDIT3	1.57	1.91	129
PCDHB2	2.22	3.01	45	WDR51B	1.58	1.90	130

\*Fold-change (FC) reflective the expression of (tumorigenic-enriched derivatives)/(parental populations).

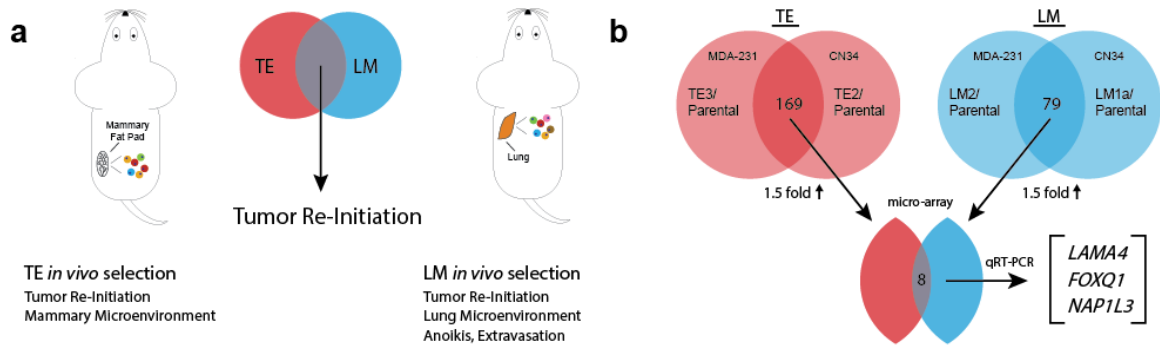
**Table 3.1 continued.**

Gene ID	MDA- TE3 FC*	CN34- TE2 FC*	#	Gene ID	MDA- TE3 FC*	CN34- TE2 FC*	#
ANTXR2	2.02	3.17	46	CCDC28A	1.76	1.72	131
SRGN	3.54	1.63	47	TAF15	1.65	1.79	132
SRGN	3.43	1.68	48	MBNL2	1.51	1.93	133
ANKRD10	1.78	3.22	49	AP1S1	1.67	1.73	134
SPRY1	2.45	2.49	50	MANSC1	1.73	1.66	135
LOC730525	2.52	2.38	51	ICAM3	1.67	1.72	136
SPRY1	2.44	2.43	52	MYCBP2	1.70	1.69	137
PTPRR	2.24	2.61	53	KIAA0355	1.74	1.63	138
C10orf90	1.66	3.12	54	LOC286467	1.79	1.56	139
ID3	1.95	2.80	55	C12orf35	1.80	1.54	140
ST6GAL1	2.54	2.19	56	GPR162	1.83	1.50	141
DLG4	2.39	2.30	57	RAB38	1.70	1.60	142
EREG	2.38	2.29	58	ESAM	1.57	1.73	143
STAT5A	1.52	3.10	59	MANSC1	1.56	1.72	144
DOCK10	2.16	2.43	60	C7orf46	1.67	1.58	145
HDAC9	1.76	2.70	61	CKAP2	1.50	1.75	146
ANKRD10	1.60	2.81	62	EAF1	1.61	1.64	147
DYNC111	1.84	2.50	63	TGDS	1.63	1.61	148
C4orf34	1.65	2.68	64	TM9SF2	1.59	1.65	149
DUSP6	1.57	2.74	65	MBNL2	1.54	1.69	150
MRIP	1.56	2.72	66	CBX5	1.71	1.52	151
INSIG1	2.01	2.26	67	CARKD	1.66	1.56	152
QPCT	2.47	1.79	68	PPM1F	1.62	1.60	153
UFM1	1.81	2.39	69	ATF4	1.55	1.65	154
ICAM2	1.75	2.44	70	GNG11	1.67	1.52	155
VGLL4	1.87	2.29	71	ASB1	1.68	1.51	156
KYNU	2.60	1.56	72	ANKDD1A	1.50	1.68	157
CDK6	1.55	2.57	73	LOC285074	1.57	1.61	158
TM4SF1	1.71	2.41	74	FLJ21986	1.53	1.65	159
SEMA3A	2.20	1.89	75	ERRFI1	1.61	1.55	160
MTSS1	1.78	2.29	76	MTIF3	1.55	1.60	161
CDKN1B	1.65	2.41	77	JMJD1C	1.56	1.57	162
UFM1	1.89	2.13	78	IER5L	1.56	1.56	163
DUSP6	1.50	2.49	79	SERINC1	1.51	1.60	164
SLCO4A1	2.03	1.93	80	SMARCC2	1.51	1.60	165
GEM	2.01	1.93	81	TTC14	1.53	1.52	166
PPFIBP1	1.73	2.15	82	IFRD1	1.51	1.54	167
CREBZF	2.04	1.85	83	DIAPH3	1.54	1.50	168
GEM	1.92	1.96	84	UBE2CBP	1.50	1.53	169
DAB2	2.34	1.52	85				

\*Fold-change (FC) reflective the expression of (tumorigenic-enriched derivatives)/(parental populations).



This list of genes was narrowed to focus on candidates that might represent general regulators of tumor re-initiation. Such general regulators would be expected to function independent of the local micro-environment, and not be restricted to regulating growth only in the organ used for selection or to regulating features either only specific to primary tumor formation or metastasis (Figure 3.1a). To identify such candidates, this list was combined with the set of genes expressed at greater levels in highly metastatic cells that had been previously derived from the same parental populations. Comparison of the set of genes that displayed greater expression in the TE derivatives (Table 3.1) and in lung-metastatic (LM) sub-lines (Table 3.2) that had been previously generated from the MDA-231 (Minn et al., 2005) and CN34 (Tavazoie et al., 2008) cell lines revealed eight candidate genes that displayed a greater than 1.5-fold increase in expression in all *in vivo* selected derivatives relative to their respective parental populations (Figure 3.1b, Table 3.3).



**Figure 3.1: Systematic identification of candidate promoters of tumor re-initiation**

(a) Genes whose expression was greater in tumorigenic-enriched (TE) and lung-metastatic (LM) *in vivo* selected derivatives were considered candidate promoters of general features of tumor re-initiation. The intersection of features common to TE and LM derivatives was hypothesized to exclude factors only specific to growth in the primary organ (mammary microenvironment) or only specific to the metastatic cascade (lung microenvironment, extravasation, anoikis). (b) Transcriptomic gene expression 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 through subsequent quantitative real-time PCR (qRT-PCR) assessment.

**Table 3.2: Genes with increased expression (1.5x higher) in lung-metastatic (LM) derivatives relative to their parental populations.**

Gene ID	MDA-LM2 FC*	CN34-LM1a FC*	#	Gene ID	MDA-LM2 FC*	CN34-LM1a FC*	#
MMP1	1.97	32.14	1	KCNK1	1.93	2.98	41
IL13RA2	27.33	3.69	2	DCLK1	2.05	2.78	42
WDR72	11.32	5.46	3	ZNF503	1.54	3.08	43
SPANXA1	6.23	10.46	4	SERPINE2	1.92	2.62	44
SPANXA1	3.09	11.46	5	ZIC2	2.65	1.88	45
DCLK1	5.38	9.01	6	KIAA0746	2.06	2.39	46
GYPC	2.42	8.41	7	FBLN1	2.47	1.97	47
FOXA1	1.88	8.82	8	CTHRC1	2.06	2.31	48
ALDH3A1	4.17	5.52	9	RAB37	1.71	2.49	49
GYPC	1.87	6.97	10	LOC401233	2.12	1.97	50
LOC731895	1.63	7.00	11	<b>LAMA4</b>	<b>2.56</b>	<b>1.51</b>	<b>51</b>
KISS1	2.21	6.14	12	CCND2	1.59	2.44	52
INHBB	2.98	5.19	13	LAMA5	2.50	1.52	53
PLCB1	3.96	3.69	14	COL6A2	2.42	1.51	54
CADM1	3.10	4.38	15	NQO1	1.55	2.28	55
SYK	1.84	5.56	16	NR2F1	1.65	2.17	56
ITGB4	2.63	4.70	17	QPCT	2.31	1.50	57
ITGB4	2.45	4.74	18	BICD1	2.10	1.71	58
PHGDH	3.97	3.18	19	ALDH3A2	1.66	2.07	59
FRMD3	1.65	5.35	20	SHISA3	1.78	1.84	60
ODZ3	1.86	5.09	21	SLC22A18	1.73	1.85	61
CTHRC1	2.64	3.85	22	ALDH3A2	1.63	1.88	62
LOC124220	1.77	4.65	23	NQO1	1.63	1.88	63
<b>NAP1L3</b>	<b>1.63</b>	<b>4.65</b>	<b>24</b>	ITGA6	1.71	1.76	64
POU3F2	1.93	4.14	25	FBLN1	1.78	1.69	65
TENM1	2.32	3.68	26	TGFB1	1.86	1.60	66
GARNL4	1.67	4.21	27	PPARG	1.55	1.90	67
TNFRSF19	1.52	4.29	28	TMEM154	1.90	1.51	68
FSCN1	2.00	3.80	29	<b>FOXQ1</b>	<b>1.63</b>	<b>1.73</b>	<b>69</b>
KCNK1	1.91	3.53	30	COL6A2	1.58	1.75	70
AMPH	2.28	3.12	31	COL13A1	1.66	1.66	71
CHRM3	2.22	3.10	32	LOC642236	1.62	1.70	72
SLC47A1	2.18	3.11	33	<b>LAMA4</b>	<b>1.75</b>	<b>1.53</b>	<b>73</b>
IL_1914933	2.91	2.34	34	SLIT2	1.73	1.54	74
LDOC1	1.91	3.30	35	SOX21	1.53	1.71	75
PCDH20	2.44	2.76	36	TRAK2	1.59	1.64	76
SRGN	3.57	1.62	37	RORA	1.64	1.57	77
BMP2	1.75	3.40	38	PLAG1	1.53	1.61	78
CHN2	1.63	3.48	39	FLRT2	1.50	1.63	79
MPP4	1.58	3.39	40				

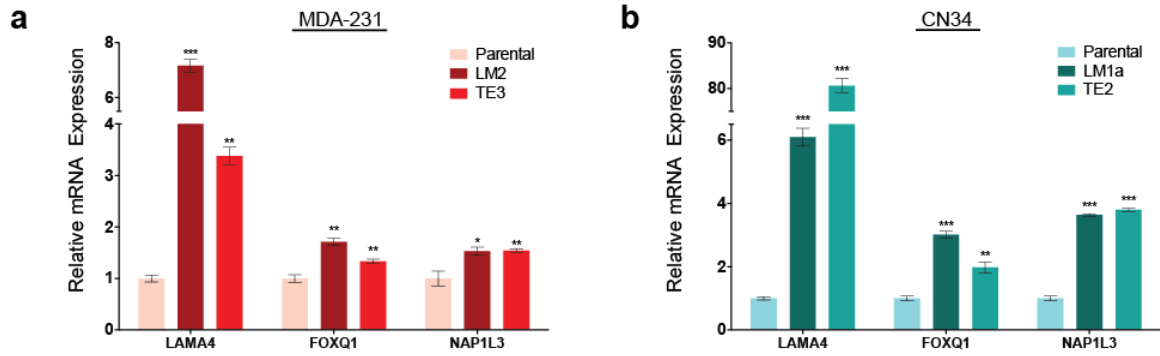
\*Fold-change (FC) reflective the expression of (lung-metastatic derivatives)/(parental populations).

**Table 3.3: Genes with increased expression in tumorigenic-enriched (TE) and lung-metastatic (LM) cells relative to their parental populations (1.5x higher in all comparisons).**

Gene ID	MDA-TE3 FC*	CN34-TE2 FC*	MDA-LM2 FC*	CN34-LM1a FC*	#
<b>NAP1L3</b>	<b>1.78</b>	<b>4.62</b>	<b>1.63</b>	<b>4.65</b>	<b>1</b>
SLIT2	1.91	5.76	1.73	1.54	2
FRMD3	1.76	1.88	1.65	5.35	3
SRGN	3.54	1.63	3.57	1.62	4
<b>LAMA4</b>	<b>1.81</b>	<b>4.39</b>	<b>2.56</b>	<b>1.51</b>	<b>5</b>
TMEM154	3.25	2.22	1.90	1.51	6
QPCT	2.47	1.79	2.31	1.50	7
<b>FOXQ1</b>	<b>1.60</b>	<b>2.07</b>	<b>1.63</b>	<b>1.73</b>	<b>8</b>

\*Fold-change (FC) reflective the expression of (*in vivo* selected derivatives)/(parental populations).

Quantitative real-time PCR validation of this set of eight yielded three genes, *LAMA4*, *FOXQ1*, and *NAPIL3*, that displayed significantly greater levels of expression in TE and LM derivatives from both cell lines relative to their respective parental populations (Figure 3.2a-b). *LAMA4* encodes the secreted alpha chain isoform protein laminin- $\alpha$ 4, an ECM component mainly present in basement membranes (Stenzel et al., 2011). *FOXQ1* is a member of the FOX family of transcription factors with roles in development (Goering et al., 2008) and cancer progression (Zhang et al., 2011a). *NAPIL3* encodes a poorly characterized member of the nucleosome assembly protein 1-like (NAP1L) family (Attia et al., 2011).



**Figure 3.2: TE cells express increased levels of *LAMA4*, *FOXQ1*, and *NAP1L3*–genes also expressed at greater levels by highly metastatic cells (a-b)** qRT-PCR assessment 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 (a) and CN34 (b) cell lines relative to their respective parental populations.  $n = 3$  samples per group. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  were obtained using one-sided Student's t-test (a-b). All data are represented as mean + or  $\pm$  S.E.M.

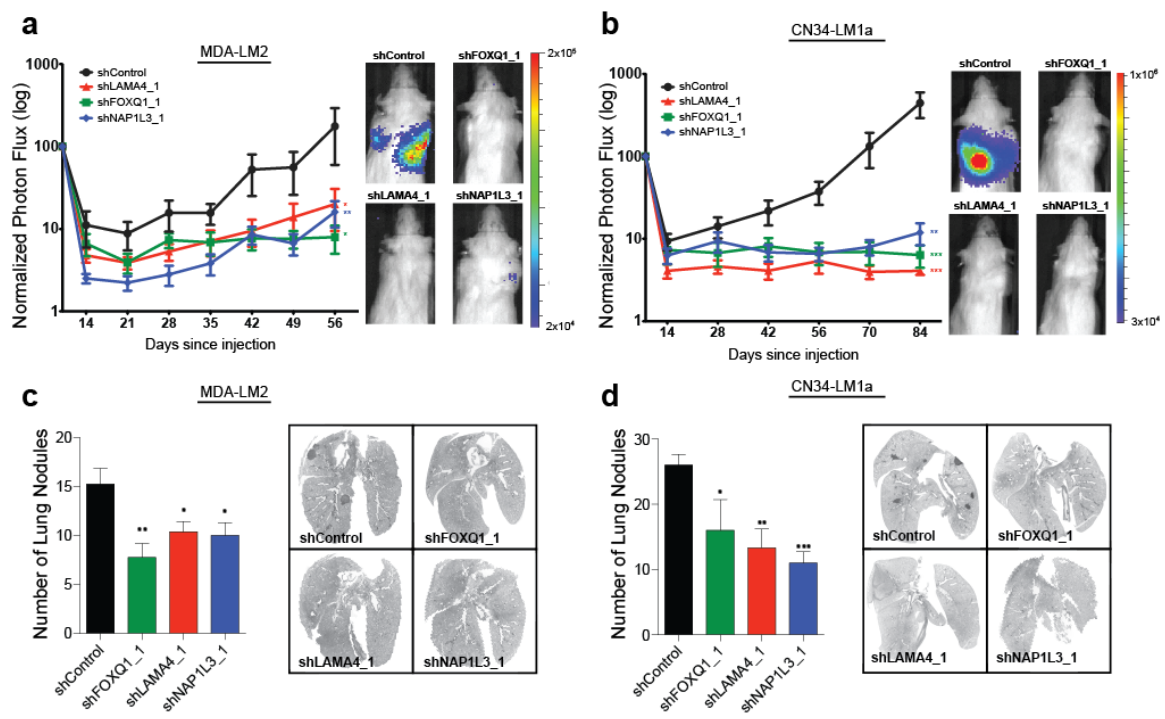
### ***LAMA4*, *FOXQ1*, and *NAPIL3* promote metastatic efficiency**

Given their increased expression in tumorigenic-enriched and metastatic derivatives, *LAMA4*, *FOXQ1*, and *NAPIL3* represented candidate promoters of primary and metastatic tumor re-initiation. As a first step to assess the function of these genes, *in vivo* studies were conducted to determine if these genes regulate the formation of macroscopic nodules by highly metastatic cells. Knockdown of *LAMA4*, *FOXQ1*, or *NAPIL3* in highly metastatic MDA-LM2 and CN34-LM1a derivatives significantly decreased metastatic colonization of the lungs based on bioluminescence tracking in both the MDA-LM2 and CN34-LM1a cell lines (Figure 3.3a-b, Figure 3.4a-c). Consistent with these genes promoting the re-initiation of metastatic colonies, histological analysis of lungs from these mice revealed significant decreases in the number of macro-metastases that were formed upon knockdown of each of these genes (Figure 3.3c-d).

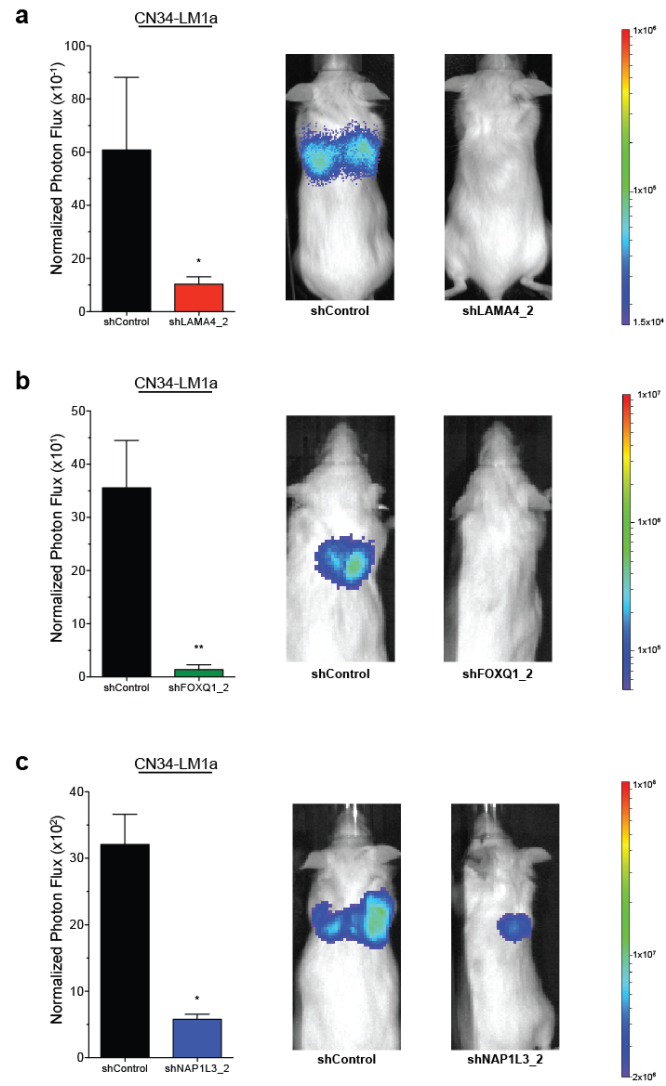
**Figure 3.3: *LAMA4*, *FOXQ1*, and *NAPIL3* promote metastatic efficiency**

(a and c)  $2 \times 10^4$  MDA-LM2 cells transduced with either a control shRNA or shRNAs targeting *LAMA4*, *FOXQ1*, or *NAPIL3* were inoculated intravenously into immunodeficient mice. shRNA-depletion of *LAMA4*, *FOXQ1*, and *NAPIL3* led to a significant reduction in metastasis as demonstrated by bioluminescence measurements over 56 days normalized to post-injection signal at day 0 (a).  $n = 8$  (shControl),  $n = 8$  (shLAMA4\_1),  $n = 7$  (shNAPIL3\_1),  $n = 7$  (shFOXQ1\_1) independent mice. Representative vimentin-stained lungs and quantification of the number of macroscopic lung nodules on day 56 (c).  $n = 4$  (shControl),  $n = 5$  (shLAMA4\_1),  $n = 4$  (shNAPIL3\_1),  $n = 5$  (shFOXQ1\_1) independent lungs. (b and d)  $5 \times 10^4$  CN34-LM1a cells transduced with either a control shRNA, or shRNAs targeting *LAMA4*, *FOXQ1*, or *NAPIL3* were injected intravenously into immunodeficient mice. shRNA-depletion of *LAMA4*, *FOXQ1*, and *NAPIL3* led to a significant reduction in metastasis as demonstrated by bioluminescence measurements over 84 days normalized to post-injection signal at day 0 (b).  $n = 26$  (shControl),  $n = 6$  (shLAMA4\_1),  $n = 5$  (shNAPIL3\_1),  $n = 5$  (shFOXQ1\_1) independent mice. Representative vimentin-stained lungs and quantification of the number of macroscopic lung nodules on day 84 (c).  $n = 5$  (shControl),  $n = 3$  (shLAMA4\_1),  $n = 3$  (shNAPIL3\_1),  $n = 3$  (shFOXQ1\_1) independent lungs. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  were obtained using one-sided Mann-Whitney test (a-b), or one-sided Student's t-test (c-d). All data are represented as mean + or  $\pm$  S.E.M.



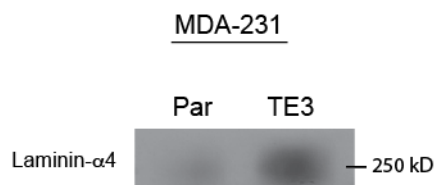


**Figure 3.4: Independent shRNA knockdown of *LAMA4*, *FOXQ1*, and *NAPIL3* leads to suppression of metastasis *in vivo*** (a)  $1.5 \times 10^5$  CN34-LM1a cells transduced with either a control shRNA hairpin or an independent shRNA hairpin targeting *LAMA4* were inoculated intravenously into immunodeficient mice. shRNA depletion of *LAMA4* led to a significant reduction in metastasis as measured by bioluminescence imaging on day 25 normalized to post-injection signal on day 0.  $n = 4$  (shControl),  $n = 3$  (shLAMA4\_2) independent mice. (b)  $5 \times 10^4$  CN34-LM1a cells transduced with either a control shRNA hairpin or an independent shRNA hairpin targeting *FOXQ1* were inoculated intravenously into immunodeficient mice. shRNA depletion of *FOXQ1* led to a significant reduction in metastasis as measured by bioluminescence imaging on day 84 normalized to post-injection signal on day 0.  $n = 4$  (shControl),  $n = 5$  (shFOXQ1\_2) independent mice. (c)  $5 \times 10^4$  CN34-LM1a cells transduced with either a control shRNA hairpin or an independent shRNA hairpin targeting *NAPIL3* were inoculated intravenously into immunodeficient mice. shRNA depletion of *NAPIL3* led to a significant reduction in metastasis as measured by bioluminescence imaging on day 42 normalized to post-injection signal on day 0.  $n = 4$  independent mice.  $*P < 0.05$ ,  $**P < 0.01$ , were obtained using one-sided Mann-Whitney test (a-c). All data are represented as mean + S.E.M.



### ***LAMA4* promotes tumor re-initiation in multiple organ microenvironments**

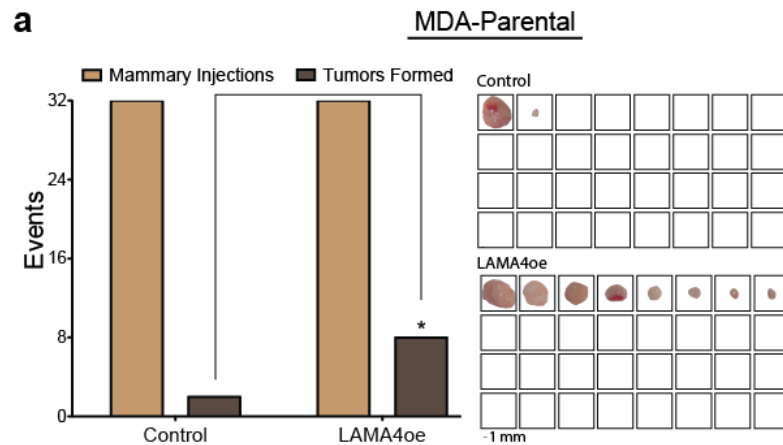
Upon identifying a set of genes (*LAMA4*, *FOXQ1*, and *NAPIL3*) that displayed increased expression in cells with enhanced tumorigenic and metastatic potential, one of these genes, *LAMA4*, was chosen for further study. *LAMA4* was selected for further characterization based on i) its greater fold expression increase observed in *in vivo* selected derivatives compared to *FOXQ1* and *NAPIL3*, and ii) its encoding of a secreted ECM protein that could be therapeutically targeted. Consistent with its extracellular role, greater levels of laminin- $\alpha$ 4 protein were detected in conditioned media obtained from MDA-TE3 cells relative to their parental population (Figure 3.5).



**Figure 3.5: TE3 cells secrete greater levels of laminin- $\alpha$ 4 relative to their parental populations**

Anti-laminin- $\alpha$ 4 antibody was used to detect endogenous protein in supernatant collected from *in vitro* culture of MDA-parental or MDA-TE3 cells.

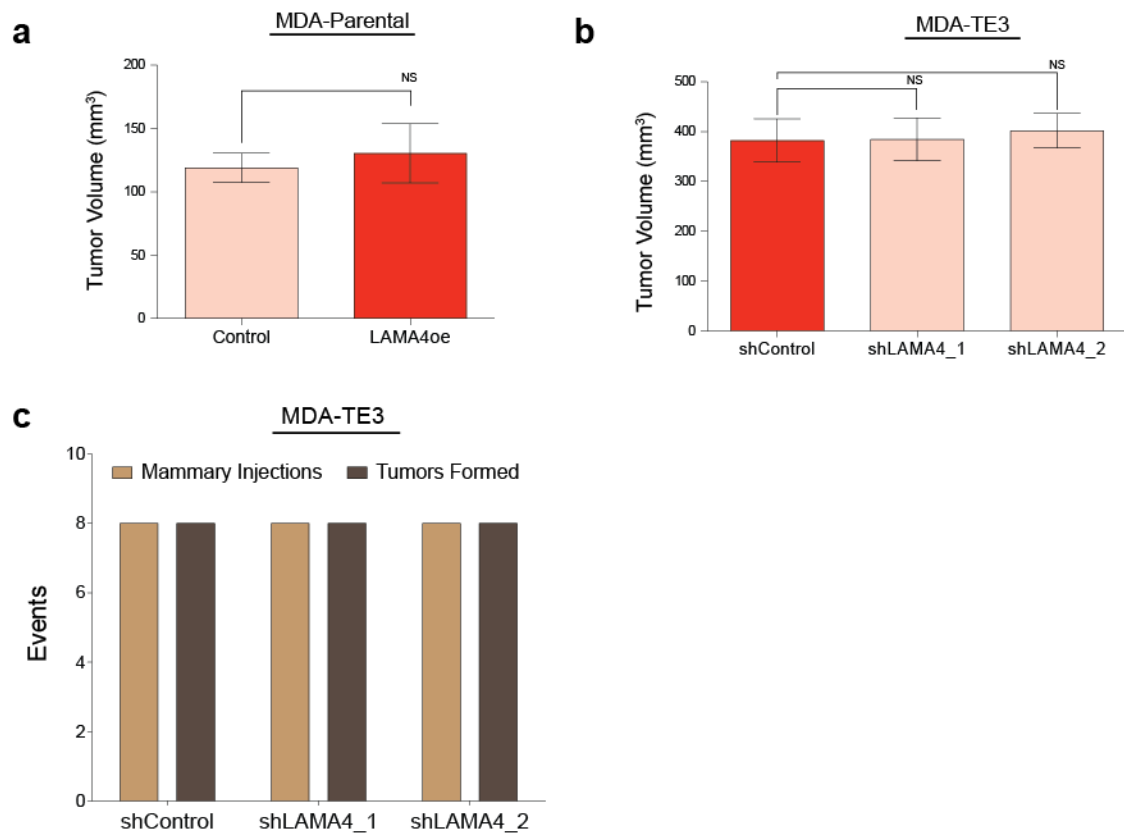
The increased expression of *LAMA4* in *in vivo* selected derivatives suggested that expression of this gene could enhance the tumor re-initiation capacity of parental breast cancer populations. Over-expression of *LAMA4* in parental populations led to a significant increase (4-fold) in the number of tumors formed upon injection of low numbers of cells (10 cells per injection) into the mammary fat pads of immunodeficient mice (Figure 3.6a).



**Figure 3.6: Over-expression of *LAMA4* in parental populations is sufficient to promote orthotopic tumor re-initiation**

(a)  $1 \times 10^1$  MDA-parental cells transduced with either an empty vector control or *LAMA4* over-expression vector were injected into the mammary fat pads of immunodeficient mice. MDA-parental *LAMA4*-over-expressing cells yielded tumors in 8/32 sites as compared to 2/32 sites for MDA-parental empty vector control cells after 10 weeks (left). Gross tumor explants (right).  $n = 32$  independent mammary fat pad injections (represented as open squares, right). \* $P < 0.05$  was obtained using one-sided Fisher's exact test (a).

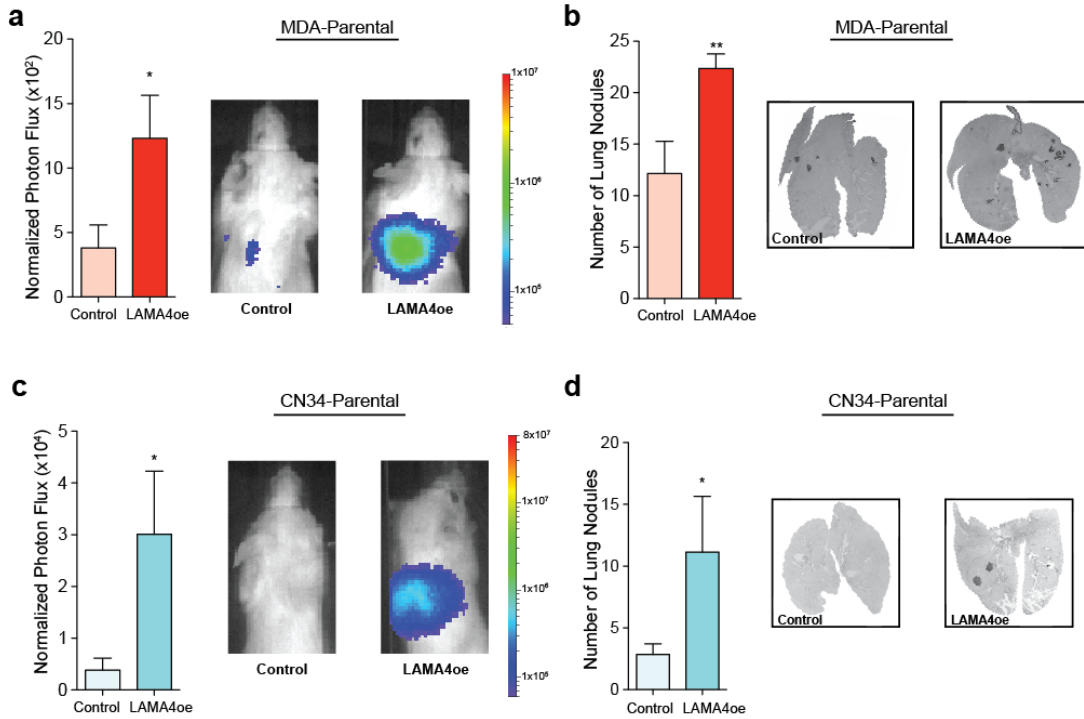
In the context of injecting a high number of cells ( $5 \times 10^5$  cells per injection), overexpression or depletion of *LAMA4* did not significantly alter tumor growth, as there were no significant differences in tumor size (Figure 3.7a-b), and depletion of *LAMA4* had no effect on the frequency of tumors formed (Figure 3.7c).



**Figure 3.7: Assessment of *LAMA4*'s regulation of tumor growth upon injection of high cell doses** (a)  $5 \times 10^5$  MDA-parental cells transduced with either an 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. (b-c)  $5 \times 10^5$  MDA-TE3 cells transduced with either a 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 (b). Tumors were formed in all cases (c).  $n = 8$  independent mammary fat pad injections. NS is not significant based on one-sided Mann-Whitney test (a-b).

Over-expression of *LAMA4* in parental populations also led to a significant increase in metastatic colonization of the lungs upon tail-vein injection of cancer cells (Figure 3.8a,c), resulting in a significant increase in the number of macroscopic colonies formed (Figure 3.8b,d). Taken together, these results demonstrate that over-expression of *LAMA4* in parental populations is sufficient to enhance both their tumorigenic and metastatic potential.

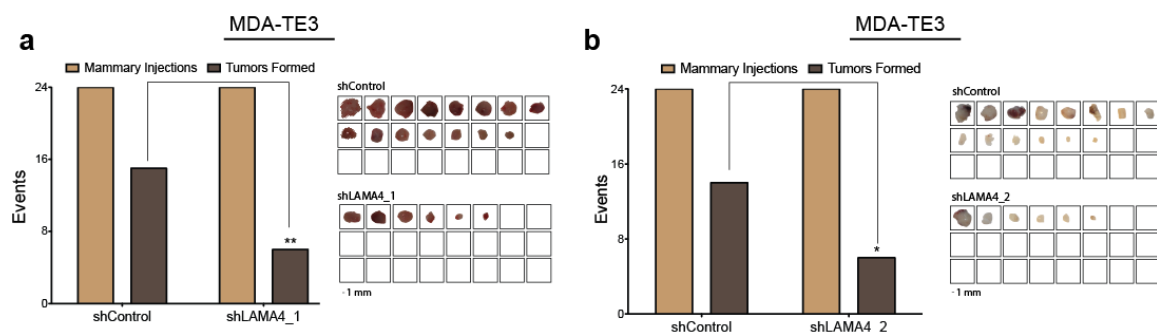




**Figure 3.8: Over-expression of *LAMA4* in parental populations is sufficient to promote metastasis**

(a-b)  $4 \times 10^4$  MDA-parental cells transduced with an empty vector control or *LAMA4* over-expression vector were inoculated intravenously into immunodeficient mice. Lung bioluminescence was measured on day 56 and normalized to post-injection signal at day 0 (a).  $n = 7$  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 (b).  $n = 7$  (control),  $n = 6$  (LAMA4oe) independent lungs. (c-d)  $5 \times 10^5$  CN34-parental cells transduced with either an empty vector control or *LAMA4* over-expression vector were inoculated intravenously into immunodeficient mice. Lung bioluminescence was measured on day 133 and normalized to post-injection signal at day 0 (a).  $n = 7$  independent mice. Lungs were harvested on day 133, vimentin-stained, and the number of macroscopic nodules per lung section was counted. Representative vimentin-stained lungs on day 133 (b).  $n = 7$  independent lungs. \* $P < 0.05$ , \*\* $P < 0.01$ , were obtained using one-sided Mann-Whitney test (a,c-d) or one-sided Student's t-test (b). All data are represented as mean  $\pm$  S.E.M. NS, not significant.

Next, the role of endogenous *LAMA4* in TE derivatives was assessed and revealed that depletion of *LAMA4* in TE3 cells significantly decreased (shRNA\_1, 2.5-fold; shRNA\_2, 2.33-fold) the numbers of tumors formed upon orthotopic injection of limiting numbers of cells (10 cells per injection) into the mammary fat pads of immunodeficient mice (Figure 3.9a-b). Depletion of *LAMA4* in TE3 cells also led to a significant (11-fold) decrease in the numbers of macroscopic colonies formed upon ectopic injection into the lungs (Figure 3.10a-b). 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.

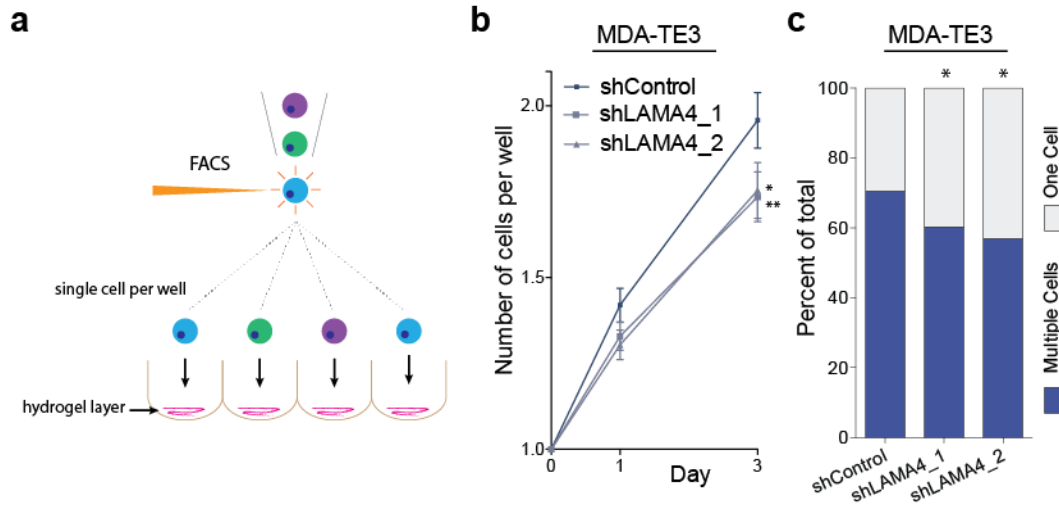


**Figure 3.9: Endogenous expression of *LAMA4* by TE derivatives promotes orthotopic tumor re-initiation** (a)  $1 \times 10^1$  MDA-TE3 cells transduced with either a control shRNA or an shRNA targeting *LAMA4* were injected into the mammary fat pads of immunodeficient mice. MDA-TE3 shControl cells yielded tumors in 15/24 sites as compared to 6/24 sites for MDA-TE3 shLAMA4\_1 cells after 10 weeks (left). Gross tumor explants (right).  $n = 24$  independent mammary fat pad injections (represented as open squares, right). (b)  $1 \times 10^1$  MDA-TE3 cells transduced with either a control shRNA or an independent shRNA targeting *LAMA4* were injected into the mammary fat pads of immunodeficient mice. MDA-TE3 shControl cells yielded tumors in 14/24 sites as compared to 6/24 sites for MDA-TE3 shLAMA4\_2 after 10 weeks (left).  $n = 24$  independent mammary fat pad injections (represented as open squares, right).  $*P < 0.05$  and  $**P < 0.01$  were obtained using a one-sided Fisher's exact test (a-b)

**Figure 3.10: Endogenous expression of *LAMA4* by TE derivatives promotes ectopic tumor re-initiation (a-b)**  $1 \times 10^2$  MDA-TE3 cells transduced with either a 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 post-injection signal at day 0 (**a**).  $n = 5$  independent mice. On day 63 lungs were sectioned, vimentin stained, and the number of macroscopic nodules per lung was counted (**b**).  $n = 5$  independent lungs.  $**P < 0.01$  was obtained using one-sided Mann-Whitney test (**a**) or one-sided Student's t-test (**b**). All data are represented as mean + or  $\pm$  S.E.M.

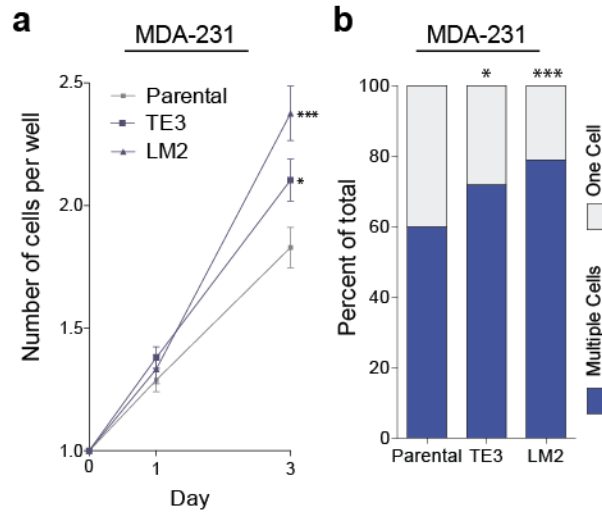
***LAMA4* promotes the clonal expansion of breast cancer cells in the absence of substratum-attachment *in vitro***

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 (Buchstaller et al., 2012; Grassian et al., 2011; Pavlova et al., 2013)—conditions that can lead to protracted cell cycle arrest or programmed cell death (Aguirre-Ghiso, 2007). The cell-autonomous expression of *LAMA4* by cancer cells suggested that this gene might enable them to proliferate independent of attachment to an underlying substratum. In order to test this hypothesis, cancer cells were seeded at clonal density *in vitro* into plates that had been treated with a hydrogel layer that prevents adherence to substratum. When TE3 cells were seeded at a clonal density of one cell per well onto low-attachment 96-well plates that prevent substratum-attachment (Figure 3.11a), wells containing *LAMA4*-depleted cells contained a smaller number cells after 3 days relative to wells seeded with cells transduced with a control hairpin (Figure 3.11b).



**Figure 3.11: *LAMA4* promotes the clonal expansion of cancer cells in the absence of substratum-attachment *in vitro*** (a) To assess the clonal expansion of cancer cells in the absence of substratum-attachment *in vitro*, single cells were sorted at a clonal density of one cell per well into low-attachment 96-well plates and monitored over several days. (b-c) MDA-TE3 cells transduced with either control shRNA or two independent shRNAs targeting *LAMA4* were sorted at a clonal density of one cell per well into low-attachment 96-well plates. The number of cells contained in each well was counted on days 1 and 3. MDA-TE3 cells transduced with control shRNA proliferated more extensively than MDA-TE3 cells transduced with shRNAs targeting *LAMA4* (b).  $n = 212$  (shControl\_1),  $n = 219$  (shLAMA4\_1),  $n = 224$  (shLAMA4\_2) independent wells. The number of wells containing only one cell compared to the number of wells containing multiple cells for each condition was assessed on day 3 (c).  $n = 189$  (shControl\_1),  $n = 196$  (shLAMA4\_1),  $n = 197$  (shLAMA4\_2) independent wells. Graph c depicts percent of total. \* $P < 0.05$ , \*\*\* $P < 0.001$  were obtained using one-sided Mann-Whitney test (b) or one-sided Fisher's exact test (c) based on counts on day 3.

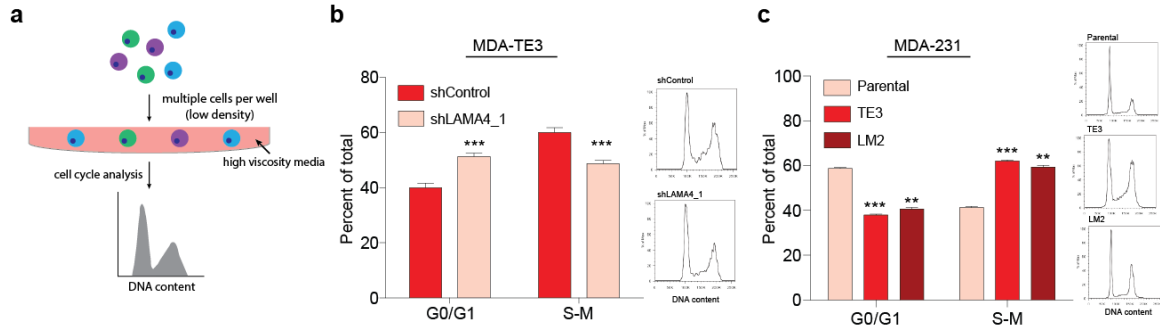
Based on these findings, tumorigenic-enriched and highly metastatic derivatives, which physiologically express elevated levels of *LAMA4* relative to their parental populations, would be expected to demonstrate greater clonal expansion in the absence of substratum-attachment *in vitro*. When TE3 or LM2 cells were seeded onto low-attachment plates at a clonal density of one cell per well, they gave rise to a greater number of cells than their parental population (Figure 3.12a). In addition, further examination of the wells seeded with single cells after several days revealed that a greater fraction of wells seeded with TE3 and LM2 cells contained multiple cells relative to wells that had been seeded with the parental population, which had a greater fraction of wells that still contained only one cell (Figure 3.12b). This effect was also observed in the *LAMA4*-dependent experiments, in which a smaller fraction of wells containing *LAMA4* knockdown cells contained multiple cells compared to single cells relative to wells containing control hairpin expressing cells (Figure 3.11b).



**Figure 3.12: Tumorigenic-enriched and metastatic derivatives display enhanced clonal expansion in the absence of substratum-attachment *in vitro* (a-b)** MDA-parental, MDA-TE3, or MDA-LM2 cells were sorted at a clonal density of one cell per well into low-attachment 96-well plates. The number of cells contained in each well was counted on days 1 and 3. MDA-TE3 and MDA-LM2 cells proliferated more extensively than MDA-parental cells (**a**).  $n = 152$  (parental),  $n = 163$  (TE3),  $n = 141$  (LM2) independent wells. The number of wells containing only one cell compared to the number of wells containing multiple cells for each condition was assessed on day 3 (**b**).  $n = 145$  (parental),  $n = 156$  (TE3),  $n = 133$  (LM2) independent wells. Graph **b** depicts percent of total.  $*P < 0.05$ ,  $***P < 0.001$  were obtained using one-sided Mann-Whitney test (**a**) or one-sided Fisher's exact test (**b**) based on counts on day 3.



Taken together, these observations suggested that *LAMA4* might regulate cell cycle dynamics during conditions when cells are detached from substratum. This assay was modified to enable quantitative analysis of a larger population of cells maintained at a low density (Figure 3.13a) and revealed that *LAMA4* depletion in the absence of substratum-attachment *in vitro* increased the fraction of cells in G0/G1 (Figure 3.13b). In addition, TE3 and LM2 populations cultured under the same conditions contained a smaller fraction of cells in G0/G1 relative to their parental populations (Figure 3.13c). 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 clonal expansion of cancer cells.

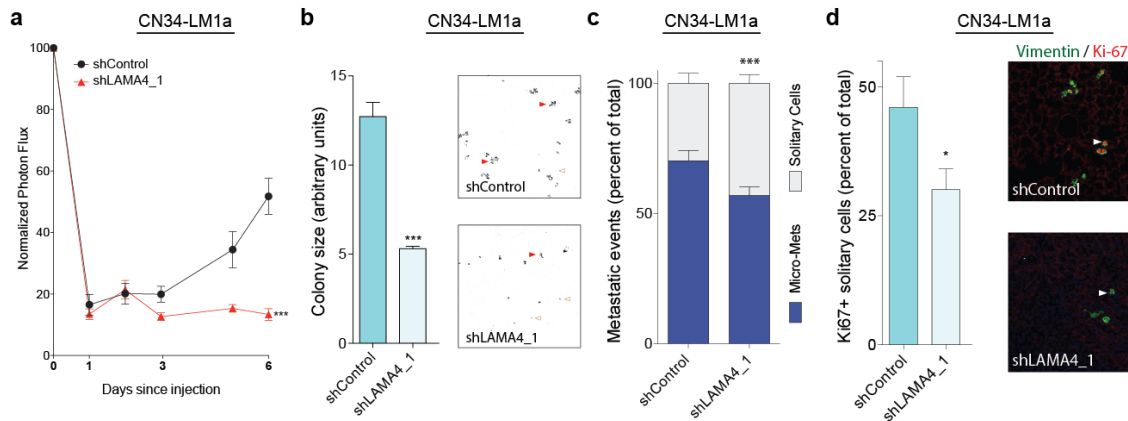


**Figure 3.13: *LAMA4* represses the fraction of cells in G0/G1 in the absence of substratum-attachment *in vitro*** (a) To assess the cell cycle phase of populations of cancer cells at low density in the absence of substratum-attachment, cells were seeded at low densities into low-attachment 6-well plates in high viscosity media containing methyl-cellulose and later isolated and subject to flow-cytometry and cell cycle analysis. (b) MDA-TE3 cells transduced with either control shRNA or an shRNA targeting *LAMA4* were assessed for cell cycle phase using a flow-cytometer (left). Representative DNA content histograms from DAPI staining (right).  $n = 6$  independent samples. (c) MDA-parental, MDA-TE3, or MDA-LM2 cells were assessed for cell cycle phase using a flow-cytometer (left). Representative DNA content histograms from DAPI staining (right).  $n = 4$  independent samples.  $**P < 0.01$ ,  $***P < 0.001$  were obtained using one-sided Student's t-test (b-c). All data are represented as mean  $\pm$  or  $\pm$  S.E.M.

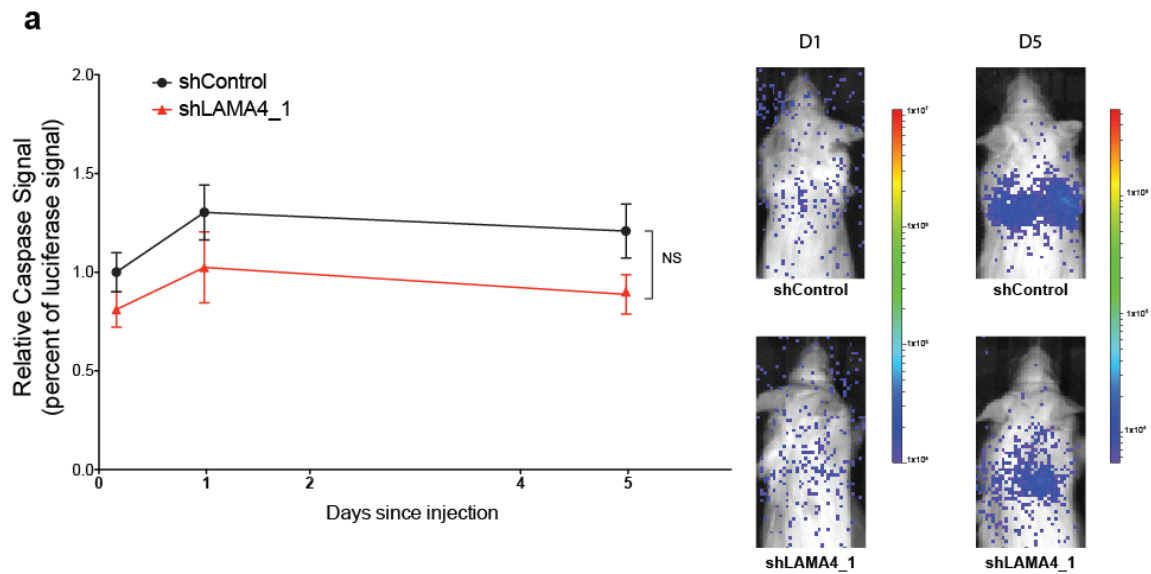
### ***LAMA4* promotes metastatic proliferation and incipient micro-metastasis formation**

During metastasis, cancer cells that have detached from the primary tumor subsequently arrive in a foreign microenvironment with a distinct ECM composition that is acutely devoid of proliferative and survival cues that permit the effective formation of colonies (Fidler, 2003, 2011; Giancotti, 2013). In addition, studies have shown that the introduction of cancer cells into the intravenous circulation leads to the pulmonary seeding of mainly solitary cells that initially lack contact with each other (Shibue and Weinberg, 2009). 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*. The promotion of clonal expansion by *LAMA4* under such *in vitro* conditions suggested that this gene might also regulate the initial proliferation of disseminated cancer cells and their subsequent formation of multi-cellular colonies *in vivo*. To address this, highly metastatic LM1a cells transduced with either control hairpins or hairpins targeting *LAMA4* were injected into the venous circulation of immunodeficient mice. After a latency period of several days, the bioluminescence signal from the lungs of mice injected with control cells began to increase while the signal from mice injected with *LAMA4*-knockdown cells showed a persistently reduced signal (Figure 3.14a). Measurement of cancer cell-derived caspase 3/7 activity through an *in vivo* reporter at several time points during this same period revealed no significant differences between the normalized signal from animals injected with control cells relative to those injected with *LAMA4*-knockdown cells (Figure 3.15a), demonstrating that the differences in early metastatic outgrowth were unlikely to be secondary to apoptosis. Rather, these

observations were consistent with *LAMA4* promoting the proliferation and expansion of cancer cells that had extravasated into the lungs. Analysis of lungs from these animals revealed that the knockdown of *LAMA4* in LM1a cells resulted in a significant decrease in the fraction of multi-cellular colonies relative to single cells in the lungs of mice several days after tail-vein injection (Figure 3.14c), demonstrating that *LAMA4* promotes the formation of micro-metastases *in vivo*. Additionally, *LAMA4* depletion also reduced the percentage of solitary cells that stained positively for the proliferative marker Ki67 (Figure 3.14d), demonstrating that a greater fraction of solitary *LAMA4*-knockdown cells were not actively proliferating. Further analysis revealed that knockdown of *LAMA4* led to a decrease in the average size of metastatic colonies formed (Figure 3.14b), consistent with the continued proliferation of control cells giving rise to larger colonies relative to knockdown cells.

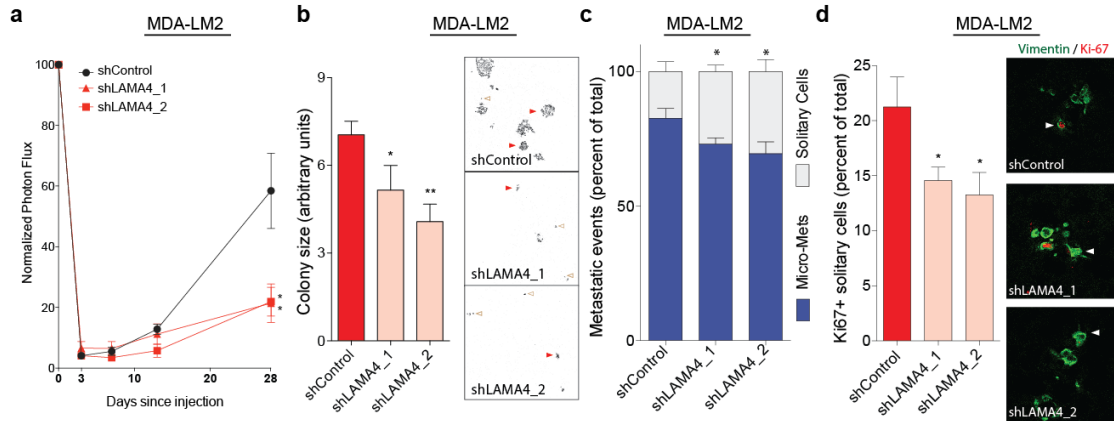


**Figure 3.14: *LAMA4* promotes the proliferation of disseminated metastatic cells and micro-metastasis formation *in vivo*** (a-d)  $3 \times 10^5$  CN34-LM1a cells transduced with either control shRNA or an shRNA targeting *LAMA4* were injected intravenously into immunodeficient mice and lung bioluminescence 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), 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).  $n = 5$  independent lungs. Micro-mets is micro-metastases.  $*P < 0.05$ ,  $***P < 0.001$  were obtained using a one-sided Student's t-test (a-d). All data are represented as mean  $\pm$  or  $\pm$  S.E.M.



**Figure 3.15: *LAMA4* does not significantly regulate apoptosis at early time points during metastatic colonization (a)**  $3 \times 10^5$  CN34-LM1a cells transduced with either control shRNA or an shRNA targeting *LAMA4* were injected intravenously into immunodeficient mice (see Figure 3.14). *In vivo* quantification of apoptotic cells was monitored by measurement of a luciferase-based caspase-3/7 reporter normalized to cancer cell luciferase signal over several days (left). Representative luciferase-based caspase-3/7 (non-normalized) bioluminescence signal (right).  $n = 5$  independent mice. NS is not significant.  $P$  value was obtained using a two-sided Student's  $t$ -test (a). All data are represented as mean  $\pm$  S.E.M.

An independent experiment employing the highly metastatic LM2 cell line confirmed these observations (Figure 3.16a-d).



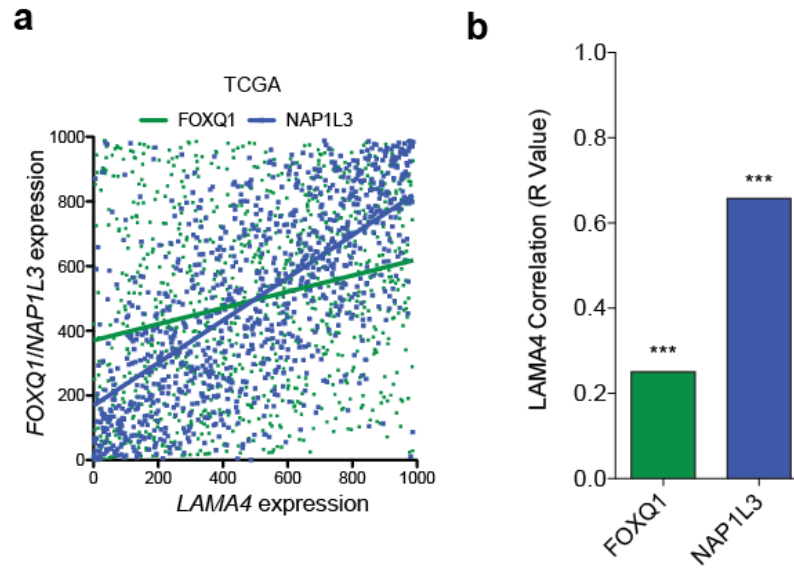
**Figure 3.16: Independent validation of *LAMA4*'s promotion of disseminated metastatic cell proliferation and colony formation *in vivo* (a-d)**  $3 \times 10^5$  MDA-LM2 cells transduced with either control shRNA or two independent shRNAs targeting *LAMA4* were injected intravenously into immunodeficient mice and lung bioluminescence was measured over time (a).  $n = 6$  (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 (b; representative background-subtracted vimentin stained images, right), 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).  $n = 4$  independent lungs. Micro-mets is micro-metastases. \* $P < 0.05$ , \*\* $P < 0.01$  were obtained using a one-sided Student's t-test (a-d). All data are represented as mean + or  $\pm$  S.E.M.

Collectively, these findings demonstrate that *LAMA4* promotes the active proliferation of disseminated solitary cells and the formation and expansion of incipient multi-cellular micro-metastases *in vivo*.

### ***FOXQ1* promotes the expression of *LAMA4***

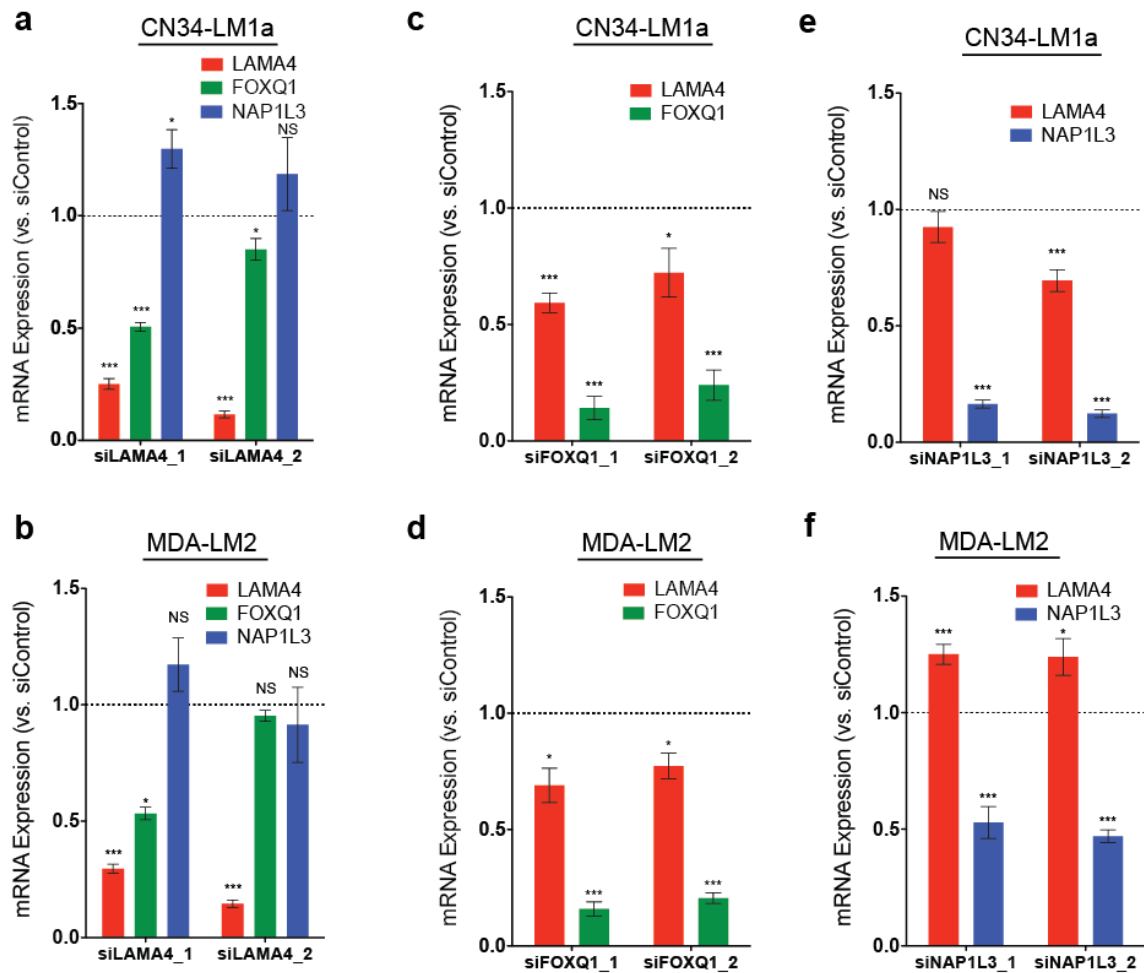
The increased expression of *FOXQ1* and *NAPIL3* along with *LAMA4* in independent *in vivo* selected derivatives suggested that these three genes might be co-expressed in other breast cancers. Interrogation of a large set of primary breast tumor samples from The Cancer Genome Atlas (TCGA) ( $n = 988$ ; see Methods) revealed that *LAMA4* expression was positively correlated with the expression of both *FOXQ1* and *NAPIL3* (Figure 3.17a-b).





**Figure 3.17: *LAMA4* expression co-varies with the expression of *FOXQ1* and *NAP1L3* in primary breast tumors (a-b)** *LAMA4* expression was positively correlated with the expression of *FOXQ1* and *NAP1L3* in a large set of primary breast tumor samples. Ranked *LAMA4* (x-axis) expression values were plotted against ranked *FOXQ1* (y-axis, green) and *NAP1L3* (y-axis, blue) expression values (a). R Spearman correlation co-efficient between *LAMA4* and *FOXQ1* or *NAP1L3* (b).  $n = 988$  independent primary breast cancer patient samples. \*\*\* $P < 0.001$  was obtained using a linear regression (b).

The co-variance of *LAMA4* with *FOXQ1* and *NAPIL3* suggested that these genes might functionally interact. Experiments revealed that the depletion of *FOXQ1* using independent hairpins in multiple cell lines decreased *LAMA4* levels (Figure 3.18c-d)—consistent with *FOXQ1* being upstream of *LAMA4*. The depletion of *LAMA4*, however, did not lead to consistent or robust changes in the expression of either *FOXQ1* or *NAPIL3* (Figure 3.18a-b). Additional experiments revealed that depletion of *NAPIL3* did not consistently affect *LAMA4* expression across the different cell lines tested (Figure 3.18e-f). Collectively, these results demonstrate that *LAMA4* expression co-varies with *FOXQ1* and *NAPIL3* in primary human tumors, and that *FOXQ1* promotes the expression of *LAMA4*.



**Figure 3.18: *FOXQ1* promotes the expression of *LAMA4* (a-b)** Relative expression of *LAMA4*, *FOXQ1* and *NAP1L3* following siRNA knockdown of *LAMA4* in LM1as (**a**;  $n = 6$ ) or LM2s (**b**;  $n = 6$ ) compared to control siRNA. (**c-d**) Relative expression of *LAMA4* and *FOXQ1* following siRNA knockdown of *FOXQ1* in LM1as (**c**;  $n = 9$ ) or LM2s (**d**;  $n = 6$ ) compared to control siRNA (**e-f**) Relative expression of *LAMA4* and *NAP1L3* following siRNA knockdown of *NAP1L3* in LM1as (**a**;  $n = 12$ ) or LM2s (**b**;  $n = 10$ ) compared to control siRNA. NS is not significant. \* $P < 0.05$ , \*\*\* $P < 0.001$  were obtained using a two-sided one-sample t-test (**a-f**).

### Summary of findings in Chapter III

This chapter describes a systematic approach to identify genes that promote tumor re-initiation through molecular characterization of tumorigenic-enriched derivatives. Transcriptomic profiling was used to identify genes that were differentially expressed between TE cells and their parental populations. Many of these genes were also found to be upregulated in highly metastatic cells that had been previously derived from the same parental populations. Three genes (*LAMA4*, *FOXQ1*, and *NAPIL3*) demonstrated greater expression in tumorigenic-enriched and metastatic derivatives relative to their parental populations. Each of these genes promoted metastatic efficiency *in vivo*, and one of these genes, *LAMA4*, was selected for further study and characterization. Functional studies revealed *LAMA4* to be a robust promoter of both tumor re-initiation and metastasis. Over-expression of *LAMA4* in parental populations was sufficient to enhance their tumor re-initiating and metastatic capacity, while endogenous expression of *LAMA4* by tumorigenic-enriched cells promoted their capacity to re-initiate tumors in multiple microenvironments. Further characterization of the phenotypes regulated by *LAMA4* revealed that it promotes disseminated metastatic cell proliferation and colonization *in vivo*, and in conditions of non-attachment to substratum *in vitro*, *LAMA4* decreases the fraction of cells in G0/G1 and promotes the clonal expansion of cancer cells into multicellular units. The identification of *LAMA4* as a gene that promotes primary and metastatic tumor re-initiation in multiple micro-environments suggests its importance in human breast cancer biology, and the following chapter provides clinical evidence to further support the relevance of this gene (Chapter IV).

## CHAPTER IV

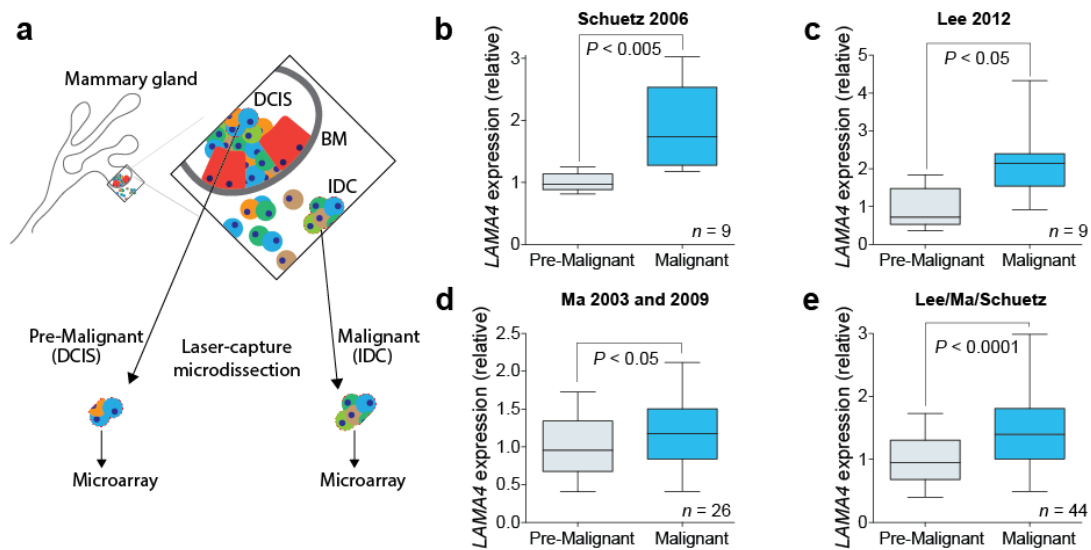
### THE CLINICAL RELEVANCE OF THESE FINDINGS TO BREAST CANCER

This chapter presents clinical data to support the role of *LAMA4* in breast cancer progression. Multiple clinical datasets were examined to assess the association of *LAMA4* with early breast cancer progression and clinical relapse.

#### **Increased expression of *LAMA4* marks early breast cancer progression**

Based on the increased expression of *LAMA4* by tumorigenic-enriched derivatives and its sufficiency in promoting tumor re-initiation *in vivo*, the role of *LAMA4* in human clinical samples was evaluated. As breast cancers transition from the pre-malignant stage (DCIS) to the malignant stage (IDC), cancer cells acquire the ability to proliferate and survive in the absence of, or with the proper engagement to, proteins in the ECM as they pass through the basement membrane and enter into the surrounding stromal microenvironment (Espina and Liotta, 2011). The role of *LAMA4* in driving the expansion of cancer cells in the settings of non-attachment *in vitro*, and in multiple microenvironments *in vivo*, suggested that expression of *LAMA4* might be increased during the transition from pre-malignancy to malignancy in breast cancer. The ideal dataset would be one wherein malignant carcinoma cells and pre-malignant cells were specifically isolated and separated from the surrounding tissue from the same patient's breast cancer lesion and molecularly interrogated. Multiple datasets (Lee et al., 2012; Ma et al., 2009; Ma et al., 2003; Schuetz et al., 2006) of laser capture micro-dissected breast

cancer tissues from pre-malignant (DCIS) or malignant (IDC) areas of disease obtained from independent patients that had been subjected to transcriptomic microarray analysis (Figure 4.1a, see Methods) were identified. Examination of these datasets revealed that within the same patients, *LAMA4* expression was significantly increased in malignant breast cancer cells relative to nearby pre-malignant cancer cells (Figure 4.1b-e;  $P < 0.0001$  for combined datasets). These findings are consistent with the increased expression of *LAMA4* in populations of cells enriched for tumor-forming potential and support functional studies demonstrating a role for *LAMA4* in promoting breast cancer tumor re-initiation.



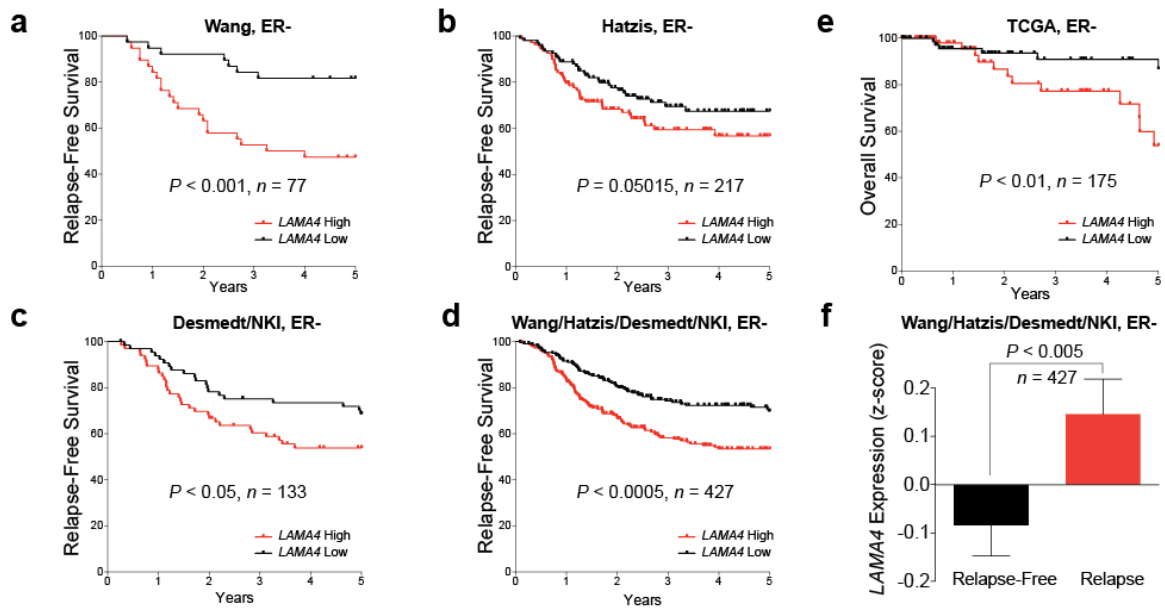
**Figure 4.1: Increased expression of *LAMA4* marks early breast cancer progression**

(a-e) *LAMA4* mRNA expression was assessed in multiple gene expression datasets (see Methods) from laser capture micro-dissected 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.  $P$  values were obtained using a one-sided paired t-test (b-e).

### **Increased expression of *LAMA4* is associated with poor relapse-free survival**

In addition to being associated with early breast cancer progression, the functional experiments demonstrating the role of *LAMA4* in tumor re-initiation suggested that this gene might also be associated with primary or metastatic relapse. Analysis of multiple independent datasets (Desmedt et al., 2007; Hatzis et al., 2011; van de Vijver et al., 2002; Wang et al., 2005) 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* (Figure 4.2a-d,  $P < 0.0005$  for combined datasets). Furthermore, tumors from patients that relapsed expressed higher levels of *LAMA4* independent of stratification (Figure 4.2f) and increased expression of *LAMA4* was also associated with shorter overall survival in an independent cohort of ER-negative breast cancer patients (Figure 4.2e).





**Figure 4.2: Increased expression of *LAMA4* is correlated with clinical relapse**

(a-d) 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*. Patient's 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 (a,  $n = 77$ ), Hatzis (b,  $n = 217$ ), Desmedt/NKI (c,  $n = 133$ ), Wang/Hatzis/Desmedt/NKI (d;  $n = 427$ ).  $n$  is the number of independent patient samples. ER- is ER-negative. (e) Kaplan-Meier curve depicting overall survival of patients with ER-negative breast cancer as a function of their primary tumor's expression of *LAMA4*. Patient's 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. (f) 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.  $P$  values were obtained using a one-sided Mantel-Cox test (a-e), or a one-tailed Mann-Whitney test (f).

#### **Summary of findings from Chapter IV**

This chapter presents evidence from clinical samples that support the role of *LAMA4* in human breast cancer progression. In early breast cancer lesions, malignant breast cancer cells express higher levels of *LAMA4* relative to adjacent pre-malignant breast cancer cells obtained from matched patient samples. These findings support the role of *LAMA4* in early breast cancer progression. The role of *LAMA4* in breast cancer relapse was also assessed. In patients with ER-negative breast cancer, those patients whose primary tumors expressed high levels of *LAMA4* succumbed to relapse earlier than patients whose tumors expressed low levels of *LAMA4*.

## CHAPTER V

### SUMMARY AND PERSPECTIVES

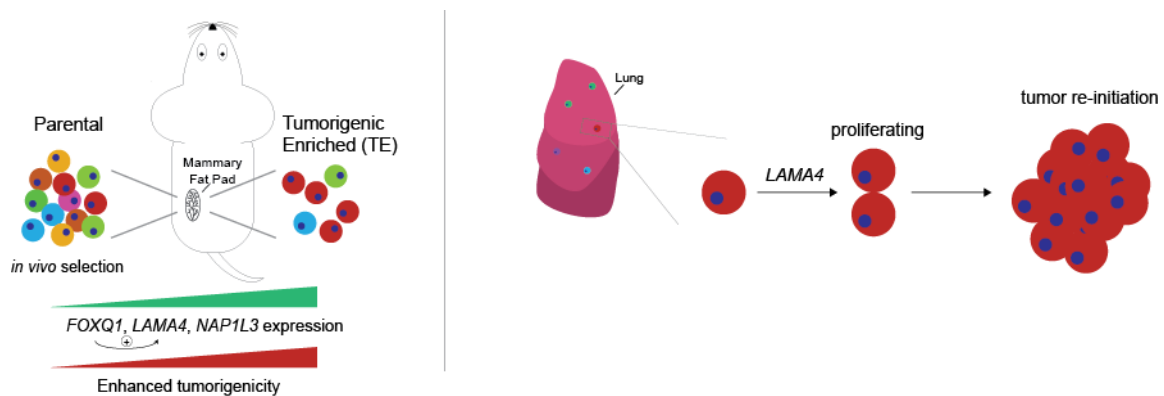
#### Overall summary

A molecular and cellular understanding of the features that govern tumor-forming potential is of great interest to the scientific and biomedical communities (Williams et al., 2013). This thesis describes a strategy to elucidate the genes and cellular biology that govern tumor re-initiation by the application of an unbiased approach to select for populations of cells with enhanced tumorigenic capacity. Tumorigenic-enriched (TE) populations derived through *in vivo* selection demonstrated enhanced tumor re-initiation capacity in orthotopic as well as ectopic contexts. Transcriptomic profiling revealed a set of genes (*LAMA4*, *FOXQ1*, and *NAPIL3*) for which expression levels were increased in TE cells relative to their parental populations. These genes were also found to be expressed at greater levels by highly metastatic cells, and each of these genes promotes metastatic efficiency *in vivo*. Characterization of one of these genes, *LAMA4*, revealed it to be sufficient to promote the enhancement of both the tumorigenic and metastatic potential of breast cancer populations. Endogenous expression of *LAMA4* by TE cells promoted tumor re-initiation in multiple organ microenvironments. Importantly, *LAMA4*'s regulation of tumor re-initiation emerged in the context of limiting cell numbers, a situation that mirrors the pathophysiological progression of human cancers.

The ability of cancer cells to survive and to proliferate in the absence of attachment to an underlying substrate throughout multiple stages of cancer progression (Pavlova et al., 2013) led to the hypothesis that cancer-derived *LAMA4* could enhance the clonal expansion of cancer cells in the absence of substratum-attachment *in vitro*, and during metastatic colonization *in vivo*, when solitary cells are exposed to a foreign micro-environment that may lack requisite extracellular cues (Aguirre-Ghiso, 2007; Giancotti, 2013). *In vivo* experiments revealed that *LAMA4* promotes the fraction of solitary disseminated metastatic cells that are proliferating during the initial phase of metastatic colonization and enhances the relative numbers and sizes of incipient micro-metastases that are generated. Based on this model, the shifting of cells from a resting, non-proliferative state to a proliferative state would yield a greater number of metastases that are re-initiated, while the continued proliferation of cells expressing *LAMA4* would increase the size of micro-metastases that are formed. Consistent with this, in conditions of non-attachment to substratum *in vitro* that model this process, *LAMA4* decreased the fraction of cells in G0/G1 and promoted the clonal expansion of cancer cells into multicellular units.

The final phase of this study presents clinical data that support the role of *LAMA4* in multiple stages of human breast cancer progression. In early breast lesions, malignant cancer cells express greater levels of *LAMA4* relative to adjacent pre-malignant cancer cells obtained from matched patient samples. In addition, increased *LAMA4* expression in ER-negative human tumors was found to be associated with reduced relapse-free survival. Collectively, this thesis describes the application of an unbiased approach to

select for populations of cells with enhanced tumorigenic capacity, whose characterization led to the identification of a molecular determinant, *LAMA4*, that promotes tumor re-initiation in multiple microenvironments and for which increased expression is associated with early human breast cancer progression and clinical relapse (Illustration 5.1).



**Illustration 5.1: *In vivo* selection for tumorigenic-enriched cells uncovers *LAMA4* as a promoter of tumor re-initiation in breast cancer**

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).

### **Derivation and characterization of tumorigenic-enriched populations**

This study describes the derivation of populations of cells with enhanced tumor-forming capacity from multiple breast cancer cell lines using *in vivo* selection. Tumorigenic-enriched populations derived through this method could be propagated *in vitro*, demonstrated enhanced tumor re-initiation capacity in the primary and metastatic microenvironment, and surprisingly did not demonstrate enhancement of multiple *in vitro* phenotypes thought to be associated with increased tumorigenic potential. These results reveal several important insights into the underlying biology of cell populations with enhanced tumorigenic capacity.

The enhanced capacity of tumorigenic-enriched derivatives to give rise to tumors despite being propagated *in vitro* reveals that their enhancement is not a transient phenotype that is lost with passage. The further characterization of tumorigenic-enriched populations using cellular, molecular, and biochemical techniques requiring large numbers of cells was significantly aided by the ability to propagate and expand these populations *in vitro*. Many other methods that have been used to study highly tumorigenic populations involve the isolation of small subsets of cells and as a result often have technical limitations that can prevent the application of certain experimental techniques (for example, biochemistry typically requires hundreds of thousands or millions of cells) (Gedye and Ailles, 2013). The derivation of highly tumorigenic cells from breast cancer cell lines using *in vivo* selection represents a methodological advance that may further the understanding of the biological basis for tumorigenicity. In addition, the success of this approach in multiple

breast cancer cell lines supports its potential to derive populations of cells with enhanced tumor-forming capacity from other cancers.

While it is thought that primary tumorigenic capacity is also associated with enhanced metastatic capacity, this question has not been fully explored (Magee et al., 2012; Vanharanta and Massague, 2013). This study reveals that *in vivo* selection for enhanced tumor-forming capacity in an orthotopic microenvironment co-selects for i) enhanced tumorigenic capacity in the lungs and ii) enhanced metastatic capacity to the liver. The enhanced capacity of tumorigenic-enriched derivatives to generate tumors in multiple metastatic organ microenvironments relative to their parental populations reveals a functional link between primary tumorigenicity and metastatic re-initiation. This link is further supported by the finding that multiple genes (*LAMA4*, *FOXQ1*, and *NAPIL3*) are commonly expressed at greater levels by populations independently *in vivo* selected for primary tumor re-initiation or metastasis.

Multiple *in vitro* phenotypes, including proliferation and colony formation under adherent conditions, the recruitment of endothelial cells, and attachment to substratum, are thought to be associated with cancer cell aggressiveness and enhanced tumor-forming capacity (Tabaries et al., 2012; Tavazoie et al., 2008; Zhou et al., 2014). The lack of enhancement of these multiple *in vitro* phenotypes by tumorigenic-enriched populations demonstrates that selection for these phenotypes is not required for selection of populations of cells with enhanced tumorigenic capacity by ER-negative breast cancer

cells. In addition, these results underscore the limitations of forming conclusions on *in vivo* properties based on *in vitro* findings.

### **Identification of molecular determinants of tumor re-initiation in breast cancer**

*LAMA4, FOXQ1, NAPIL3 are a co-varying set of genes expressed at greater levels by tumorigenic-enriched and metastatic derivatives*

Transcriptomic profiling revealed a set of genes—*LAMA4*, *FOXQ1*, and *NAPIL3*—to be expressed at greater levels by tumorigenic-enriched and metastatic derivatives relative to their parental populations. The greater levels of expression of these genes in independently derived *in vivo* selected populations from multiple breast cancer cells lines prompted the question of whether they were also co-expressed in other human breast cancers. Analysis of a large cohort of primary human breast tumors revealed that *LAMA4* expression co-varied with both *FOXQ1* and *NAPIL3*, suggesting that these genes may functionally interact or be co-regulated. Experiments revealed that *FOXQ1* promotes the expression of *LAMA4*, suggesting that *FOXQ1* acts upstream of *LAMA4*. Additional experiments revealed no interaction between *LAMA4* and *NAPIL3*. Taken together, these results suggest that *LAMA4* may be a downstream transcriptional target of *FOXQ1*. The strong correlation in the expression of *NAPIL3* and *LAMA4* suggests that these genes may share a common upstream regulator.

### *FOXQ1*

*FOXQ1* is a member of the FOX family of transcription factors which contain a specific DNA-binding motif known as a ‘winged helix’ or ‘forkhead’ domain (Bieller et al.,



2001). Multiple studies have demonstrated the role of *FOXQ1* in hair follicle development and in cancer progression (Goering et al., 2008; Zhang et al., 2011a). *FOXQ1* exerts these effects through the regulation of downstream transcriptional targets, although the overlap of these targets in different the phenotypes regulated by *FOXQ1* is unknown (Abba et al., 2013; Qiao et al., 2011; Sun et al., 2013). This study identifies *FOXQ1* as a gene whose expression is increased in populations of cells with enhanced tumorigenic and metastatic capacity. Depletion of *FOXQ1* led to decreased lung-metastatic colonization as evidenced by a decrease in the numbers of tumor nodules formed in the lungs of immunodeficient mice. Co-expression studies revealed a positive correlation between *FOXQ1* and *LAMA4*, and functional experiments demonstrated that *FOXQ1* promotes the expression of *LAMA4*. Future experiments could give insight into whether the promotion of *LAMA4* expression by *FOXQ1* is through direct transcriptional activation. In addition, the epistatic relationship between *FOXQ1* and *LAMA4* could be explored to determine if the phenotypes regulated by *FOXQ1* are dependent on the expression of *LAMA4*. Furthermore, the full transcriptional regulatory network impacted by *FOXQ1* could be assessed in order to identify additional candidate genes that may promote tumor re-initiation or metastasis in breast cancer.

### *NAPIL3*

*NAPIL3* is a poorly characterized protein that is a member of a family of nucleosome assembly 1-like (NAP1L) proteins that share significant homology to nucleosome assembly proteins (NAPs) (Watanabe et al., 1996). NAP1L proteins are thought to be involved in chromatin remodeling, and several have been implicated in the nervous

system and neuronal differentiation (Cho et al., 2013). Some NAP1L proteins have been shown to play roles in cell cycle progression, including *NAPIL2*, which has been shown to impact the cell cycle by regulating the downstream target *CDKN1C* through an effect on histone acetylation (Attia et al., 2011). While *NAPIL3* is thought to be expressed predominantly in neurons, this study identifies *NAPIL3* as a gene expressed by breast cancer cells that promotes metastatic efficiency. It would be interesting to determine if *NAPIL3* is also expressed endogenously by mammary epithelial cells, or if this gene is only expressed after their oncogenic transformation into cancer cells.

#### *LAMA4*

*LAMA4* is a member of a family of laminin genes that encode proteins normally present in the extracellular matrix (ECM) (Hynes, 2009; Nelson and Bissell, 2006). The extracellular matrix is responsible for the structural integrity of many organs (Ghajar and Bissell, 2008). The ECM also provides critical instructive signals that direct normal tissue development and homeostasis, as well as cancer (Artacho-Cordon et al., 2012). Laminins are composed of distinct heterotrimers with diverse functions (Hamill et al., 2009; Scheele et al., 2007). In humans, there are five alpha chain laminins, three beta chain laminins, and four gamma chain laminins, and each laminin heterotrimer is composed of one alpha chain, one beta chain, and one gamma chain (Domogatskaya et al., 2012). Most laminins polymerize to form long continuous networks that lay the foundation for BM formation (Moyano et al., 2010). The conformation of most laminin heterotrimers exposes the N-terminal of each laminin chain, which can bind to nearby chains to enable the polymerization of long sheets (Domogatskaya et al., 2012). During development, the

secretion and deposition of laminins into the extracellular matrix lays the foundation for other ECM proteins to bind, such as collagens and fibronectins. While laminins are secreted as soluble proteins, cross-linking of proteins to the ECM leads to the formation of an insoluble matrix. In addition to providing physical anchoring of cells to the ECM, laminins also provide critical signaling cues that regulate cell growth, proliferation, and survival (DeHahn et al., 2004; Grassian et al., 2011; Stenzel et al., 2011; Wondimu et al., 2004). Several groups have interrogated the importance of various stromal and cancer-derived laminins to cancer progression (Kusuma et al., 2012; Mori et al., 2011; Zhou et al., 2004). In many of these studies, 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 (Fujita et al., 2005; Kusuma et al., 2012; Mori et al., 2011; Yurchenco, 2011).

The function of each distinct laminin heterotrimer can largely be attributed to the identity of its  $\alpha$ -chain. Laminins generated from almost all  $\alpha$ -chains can form long continuous networks that lay the foundation for BM formation (Moyano et al., 2010), except for chains that lack an N-terminal ‘head’ region that prevents their *de novo* polymerization, such as the chain encoded by *LAMA4* (Hamill et al., 2009). Previous studies have revealed that *LAMA4* requires enzymatic cross-linking in order for it to be incorporated into ECM generated *in vitro* (Ilani et al., 2013). These features raise the possibility that laminin- $\alpha 4$  could bind to receptors on cells and lead to signaling in the absence of its attachment and cross-linkage to the extracellular matrix. Candidate receptors for laminin- $\alpha 4$  include integrins, which are known receptors for laminins and have well-described

roles in cell proliferation and survival (Miranti and Brugge, 2002; Moreno-Layseca and Streuli, 2014). The cell-autonomous expression by cancer cells of this specific laminin gene that lacks an anchoring head region provides a potential mechanism for these cells to decouple proliferative capacity from the physiologically constrained requirement of attachment to an immobile extracellular matrix.

Future studies could use genetic mouse models to further interrogate the importance of *LAMA4* to tumorigenesis. In order to study its potential role in development, investigators have generated *LAMA4* knockout mice, which have been found to have deregulated angiogenesis (Stenzel et al., 2011; Thyboll et al., 2002) and abnormalities in the peripheral nervous system (Wallquist et al., 2005). Further experiments could be conducted by crossing mice lacking *LAMA4* with mouse models of spontaneous or oncogene-driven mammary tumor formation. It could also be beneficial to generate inducible or conditional knockout mouse models of *LAMA4* in order to study the effects of acute loss of *LAMA4* just prior to, or during, the onset of tumorigenesis.

### **Phenotypes regulated by *LAMA4***

#### *Organ-independent tumor re-initiation*

While organ-specific factors can be critical determinants of tumorigenic capacity (Joyce and Pollard, 2009; Langley and Fidler, 2011; Welm, 2008), the focus of this study was to attempt to identify molecules that regulated general features of tumor re-initiation that were not dependent on a particular micro-environment. In order to identify such potential general regulators, cells that were independently *in vivo* selected for enhanced tumorigenic or metastatic capacity were transcriptomically compared to reveal a set of

common genes more highly expressed in these cells relative to their parental populations. It was hypothesized that genes within this overlapping set might function independent of the organ used for selection and not be restricted to features only relevant to tumorigenic capacity in the primary site or to the metastatic cascade. While the ability of some genes to modulate tumor re-initiation may be specific to a particular organ microenvironment, functional experiments revealed that *LAMA4* promotes tumor re-initiation capacity in orthotopic as well as ectopic settings. These results are consistent with *LAMA4* governing a general feature of tumor re-initiation that is independent of a particular microenvironment. Enhanced expression of organ selective regulators of colonization by subsets of disseminated cells during metastasis would be expected to provide further advantage to incipient metastatic cells as they colonize various organs (Nguyen et al., 2009a).

#### *Clonal expansion*

It is thought that cancer cell interaction with, and remodeling of, the local microenvironment during multiple stages of cancer progression (Chou et al., 2013; Gao et al., 2012; Kessenbrock et al., 2010; Peduto et al., 2006; Zhang et al., 2013b) can enable cells to resume cell division and expand into colonies when they lack important extracellular cues (Aguirre-Ghiso, 2007; Giancotti, 2013). An unfavorable microenvironment that does not allow for optimum interaction of cancer cell with the ECM can lead to proliferative suppression (Goss and Chambers, 2010) or cellular quiescence, a period of protracted G0/G1 arrest (Agami and Bernards, 2002; Aguirre-Ghiso, 2007). The findings that the cell-autonomous expression of *LAMA4* promotes the proliferation of cells in the absence of substratum-attachment *in vitro* and during incipient

metastatic outgrowth *in vivo* suggests that *LAMA4* antagonizes the suppressive effects that these contexts can have on proliferative potential.

## **Implications of this study to human breast cancer**

### *Therapeutic applications*

The identification of *LAMA4* as a functional promoter of tumor re-initiation in multiple ER-negative breast cancer cell lines reveals a candidate target for therapeutic intervention. Functional blocking of the secreted *LAMA4* protein product could potentially suppress the outgrowth of disseminated metastatic cells or prevent the re-activation of disseminated solitary cells that have not yet gained the capacity to re-activate proliferative growth. Blocking antibodies against *LAMA4* could be generated by immunizing mice with laminin- $\alpha 4$  protein and screening for monoclonal antibodies that demonstrates a phenotypic response. This study's identification of *LAMA4* as a potential therapeutic target in ER-negative breast cancer has important implications for patients diagnosed with ER-negative subtypes, an aggressive subset of breast cancers that are refractory to treatment with multiple anti-estrogenic therapies (Ali and Coombes, 2002; Andre and Puszta, 2006).

### *Primary and Metastatic Relapse*

While aggressive cancers can lead to rapid death due to metastatic burden to multiple organs, relapse of primary or metastatic disease can occur even after the surgical resection of the primary tumor and without any evidence of metastases (Giancotti, 2013). Such relapse is thought to be the result of small numbers of residual cells that escape treatment which subsequently re-activate proliferative potential in an inhospitable

microenvironment after varying periods of time (Aguirre-Ghiso, 2007; Bragado et al., 2013; Ghajar et al., 2013; Goss and Chambers, 2010; Vanharanta and Massague, 2013). The functional experiments from this study that demonstrate the role of *LAMA4* in the re-initiation of tumors in the primary and metastatic site suggested its expression in clinical samples might correlate with primary and metastatic relapse. In patients with ER-negative breast cancer, those patients whose primary tumors expressed higher levels of *LAMA4* were more likely to suffer from relapse relative to those patients whose tumors expressed low levels of *LAMA4*. The correlation of *LAMA4* with primary and metastatic relapse suggests its potential as a predictive marker in patients with ER-negative breast cancer. Future prospective studies would be needed to establish the prognostic utility of *LAMA4* expression in patients with ER-negative breast cancer.

## **Conclusion**

This thesis describes an unbiased system to select and characterize populations of breast cancer cells with enhanced tumorigenic potential. Molecular characterization of tumorigenic-enriched populations led to the identification of *LAMA4* as a gene that promotes tumor re-initiation in multiple microenvironments. Characterization of the phenotypes governed by *LAMA4* revealed it to promote clonal expansion during substratum-detachment *in vitro* and to promote incipient metastatic proliferation and colony formation *in vivo*. Expression of *LAMA4* marks the transition from human pre-malignant to malignant breast cancer, and higher levels of *LAMA4* in primary ER-negative tumors was correlated with poor relapse free-survival. These findings reveal cell biological and molecular insights into the processes that underlie tumor re-initiation and may have future clinical application in the treatment of breast cancer.

## MATERIALS AND 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 (Ponomarev et al., 2004) 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 (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 (Biserni, 2010). 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-enhanced (TE) derivatives from the MDA-MB-231 and CN34 human breast cancer cell lines, moderate numbers of parental populations ( $1 \times 10^4$  MDA-231-parental or  $2 \times 10^4$  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 2<sup>nd</sup> and 4<sup>th</sup> mammary glands of NOD scid mice. Tumors that were generated were dissociated into single cells (see below) and propagated *in vitro* to yield 1<sup>st</sup> generation tumorigenic-enhanced (TE1) derivatives. TE1 cells were then subjected to another round of *in vivo* selection by injecting 10-fold less ( $1 \times 10^3$  MDA-TE1 or  $2 \times 10^3$  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 2<sup>nd</sup> generation tumorigenic-enhanced (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 ( $1 \times 10^2$  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 3<sup>rd</sup> generation tumorigenic-enhanced (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 then 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 2<sup>nd</sup> and 4<sup>th</sup> 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 as the number of independent mammary fat 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 3<sup>rd</sup> intercostal space, 3mm deep, directly into the lung parenchyma of age-matched NOD scid or NOD scid gamma mice. For ectopic tumor re-initiation assays,  $n$  was as 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 as 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 (Minn et al., 2005; Tavazoie et al., 2008). The MDA-MB-231 cell line (ATCC) was derived from the pleural effusion of a patient suffering from metastatic breast cancer (Cailleau et al., 1978) and the derivation of the MDA-LM2 subpopulation through *in vivo* selection has been described previously (Minn et al., 2005). 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 (Bos et al., 2009) 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 (Tavazoie et al.,

2008). Both cell lines are Estrogen Receptor-negative (ER-negative) (Bos et al., 2009). Cells in culture were routinely tested for mycoplasma contamination. For proliferation assays in the absence of substratum-attachment, cells were sorted using a FACS Aria 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 cell cycle 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, stained with DAPI, analyzed using an LSR II (Beckton Dickinson), and cell cycle phase was determined using FlowJo (TreeStar) software. For flow cytometry, for 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 (Beckton Dickinson) and FlowJo (TreeStar) software.

### **Cell proliferation and colony formation assay**

For cell proliferation assays,  $2.5 \times 10^4$  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,  $1 \times 10^2$  cells were seeded in triplicate into tissue-culture treated adherent 10cm plates (Falcon). On day 14, plates were washed with PBS, fixed in 6% gluteraldehyde (Sigma-Aldrich), stained with 0.5% crystal violet (Sigma-Aldrich), and washed with H<sub>2</sub>O to better enable the visualization of colonies that were formed.

### **Endothelial recruitment assay**

For endothelial recruitment assays,  $5 \times 10^4$  cancer cells were seeded overnight onto tissue-culture treated adherent 24-well plates (Falcon) while human umbilical vein endothelial cells (HUVECs; Lonza) grown to 70-80% confluence were serum-starved overnight. The next day, HUVECs were labelled with CellTracker Red<sup>TM</sup> CMTPX (Molecular Probes) according to manufacturer's protocol and  $5 \times 10^4$  HUVECs were seeded onto the top layer of 3um HTS Fluroblock transwell inserts (BD Falcon) that were 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 in a tissue-culture incubator 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 (Schneider et al., 2012).

### **Cell attachment assay**

The cell attachment assay was performed by labelling cancer cells with CellTracker Red<sup>TM</sup> CMTPX (Molecular Probes) according to manufacturer's protocol and seeding  $1 \times 10^5$  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 software (Schneider et al., 2012).

### **Generation of retrovirus-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<sup>®</sup> 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.

The following shRNA sequences were used:

shControl: 5'CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTG

shFOXQ1\_1: 5'CCGGGCTGGCGGAGATCAACGAGTACTCGAGTACTCGTTGATCTCCGCCAGCTTTTTG

shFOXQ1\_2: 5'CCGGCCAGCTCCTTCGCCATCGACACTCGAGTGTGATGGCGAAGGAGCTGGTTTTTG

shLAMA4\_1: 5'CCGGGCCTAAAGCAAGTCAGAATAACTCGAGTTATTCTGACTTGCTTTAGGCTTTTTG

shLAMA4\_2: 5'CCGGCCCATGATATGAGGGATAACTCGAGTTATCCCTCATATCATGGGTTTTTG

shNAP1L3\_1: 5'CCGGGCCTGTAAGTTACACCTTTGACTCGAGTCAAAGGTGTAACCTACAGGCTTTTTG

shNAP1L3\_2: 5'CCGGCCTGTAAGTTACACCTTTGACTCGAGTCAAAGGTGTAACCTACAGGTTTTTG

### **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. The following primer sequences were used to clone full-length *LAMA4*:



*LAMA4* Forward: 5'-CCGGCCGAATTCATGGCTTTGAGCTCAGCCTG-3'

*LAMA4* Reverse: 5'-CCGGCCGTCGACTCAGGCTGCTGGACAGGAGT-3'

### **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. The following primer sequences were used:

*GAPDH* Forward: 5'- AGCCACATCGCTCAGACAC-3'

*GAPDH* Reverse: 5'- GCCCAATACGACCAAATCC-3'

*HPRT* Forward: 5'- GACCAGTCAACAGGGGACAT-3'

*HPRT* Reverse: 5'- CCTGACCAAGGAAAGCAAAG-3'

*SMAD4* Forward: 5'- TGGCCCAGGATCAGTAGGT -3'

*SMAD4* Reverse: 5'- CATCAACACCAATTCCAGCA-3'

*FOXQ1* Forward: 5'- GCGGACTTTGCACTTTGAA-3'

*FOXQ1* Reverse: 5'- TTTAAGGCACGTTTGATGGA-3'

*LAMA4* Forward: 5'- TTCGAACACCAGCTGACAAC-3'

*LAMA4* Reverse: 5'- AGGTAACCATTGCGCATTTTC-3'

*NAPIL3* Forward: 5'- CTCCCTGGTAGAGTTTGTAGGC-3'

*NAPIL3* Reverse: 5'- GACAGGTTCGAGACCATTTT-3'

### **siRNA knockdown experiments**

For siRNA knockdown experiments, cells were seeded into 6-well 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. Cells were collected and RNA was harvested 24-30hr post-transfection. Control and target gene siRNAs were obtained from Dharmacon (Thermo Fisher Scientific). The following siRNA sequences with corresponding catalogue numbers were used:

siControl: 5'-UAGCGACUAAACACAUCAA (cat: D-001210-01-05)

siFOXQ1\_1: 5'-ACGAGUACCUCAUGGGCAA (cat: D-008705-01-0005)

siFOXQ1\_2: 5'-UGGCGGAGAUCAACGAGUA (cat: D-008705-02-0005)

siLAMA4\_1: 5'-CUCAGCGGUUGGCAGGCAA (cat: D-011712-19-0005)

siLAMA4\_2: 5'-GAAGGGAGCUCAGCGGUUG (cat: D-011712-20-0005)

siNAP1L3\_1: 5'-GAACAAAUUUCGUGGAUAG (cat: D-011879-01-0005)

siNAP1L3\_2: 5'-GGGCAGAUUUUACAUGAUA (cat: D-011879-02-0005)

### **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 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 were stained with mouse primary antibodies against Vimentin (1:100, Vector Laboratories) and with Ki-67 (1:200, ab15580, abcam). 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). For solitary cell/micro-met comparisons, unprocessed images were inspected with ImageJ software(Schneider et al., 2012), 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 software(Schneider et al., 2012). 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 software(Schneider et al., 2012); 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**

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. Protein from conditioned media was separated using SDS–polyacrylamide gel electrophoresis, transferred to Immobilon-P Transfer Membrane (Millipore), and probed with an antibody against human laminin- $\alpha$ 4 (1:200; clone 6C3, sc-130541, Santa Cruz Biotechnology). 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.

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

To identify mRNAs whose levels were increased across tumorigenic-enhanced (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 core facility. The raw signal intensities for each probe were median-normalized and replicates were averaged. Candidate promoters of tumor re-initiation were identified by overlapping 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*, *FOXQ1*, and *NAPIL3* 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) (Schuetz et al., 2006), Lee 2012 (GSE41197) (Lee et al., 2012), Ma 2003: (Table 4) (Ma et al., 2003), Ma 2009 (GSE14548) (Ma et al., 2009). 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) (Wang et al., 2005), Desmedt (GSE7390) (Desmedt et al., 2007), Hatzis (GSE25066) (Hatzis et al., 2011), NKI (van de Vijver et al., 2002), and The Cancer Genome Atlas (TCGA) Breast Cancer Project provisional dataset ([/www.cBioPortal.org/](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 (Li et al., 2011) 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  $\pm$  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 exclude 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.

### **Accession Numbers**

Microarray gene expression data from the parental MDA-MB-231 and CN34 cell lines and their respective tumorigenic-enriched (TE) and lung-metastatic (LM) derivatives are deposited at GEO under the accession number GSE52629.

## REFERENCES

- Abba, M., Patil, N., Rasheed, K., Nelson, L.D., Mudduluru, G., Leupold, J.H., and Allgayer, H. (2013). Unraveling the role of FOXQ1 in colorectal cancer metastasis. *Molecular cancer research : MCR 11*, 1017-1028.
- Agami, R., and Bernards, R. (2002). Convergence of mitogenic and DNA damage signaling in the G1 phase of the cell cycle. *Cancer letters 177*, 111-118.
- Aguirre-Ghiso, J.A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. *Nature reviews Cancer 7*, 834-846.
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America 100*, 3983-3988.
- Ali, S., and Coombes, R.C. (2002). Endocrine-responsive breast cancer and strategies for combating resistance. *Nature reviews Cancer 2*, 101-112.
- Andre, F., and Puzstai, L. (2006). Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy. *Nature clinical practice Oncology 3*, 621-632.



Artacho-Cordon, A., Artacho-Cordon, F., Rios-Arrabal, S., Calvente, I., and Nunez, M.I. (2012). Tumor microenvironment and breast cancer progression: a complex scenario. *Cancer biology & therapy* 13, 14-24.

Attia, M., Forster, A., Rachez, C., Freemont, P., Avner, P., and Rogner, U.C. (2011). Interaction between nucleosome assembly protein 1-like family members. *Journal of molecular biology* 407, 647-660.

Bieller, A., Pasche, B., Frank, S., Glaser, B., Kunz, J., Witt, K., and Zoll, B. (2001). Isolation and characterization of the human forkhead gene FOXQ1. *DNA and cell biology* 20, 555-561.

Biserni, A., Martorana, F., Roncoroni, C., Klaubert, D., Maggi, A., Ciana, P. (2010). Identification of Apoptotic Cells in Reporter Mice Using Modified Luciferin. . TOP (TRANSGENIC OPERATIVE PRODUCTS) SRL.

Bos, P.D., Zhang, X.H., Nadal, C., Shu, W., Gomis, R.R., Nguyen, D.X., Minn, A.J., van de Vijver, M.J., Gerald, W.L., Foekens, J.A., *et al.* (2009). Genes that mediate breast cancer metastasis to the brain. *Nature* 459, 1005-1009.

Brabletz, T. (2012). EMT and MET in metastasis: where are the cancer stem cells? *Cancer cell* 22, 699-701.

Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nature reviews Cancer* 5, 744-749.

Bragado, P., Estrada, Y., Parikh, F., Krause, S., Capobianco, C., Farina, H.G., Schewe, D.M., and Aguirre-Ghiso, J.A. (2013). TGF-beta2 dictates disseminated tumour cell fate in target organs through TGF-beta-RIII and p38alpha/beta signalling. *Nature cell biology* 15, 1351-1361.

Briskin, C. (2013). Progesterone signalling in breast cancer: a neglected hormone coming into the limelight. *Nature reviews Cancer* 13, 385-396.

Buchstaller, J., McKeever, P.E., and Morrison, S.J. (2012). Tumorigenic cells are common in mouse MPNSTs but their frequency depends upon tumor genotype and assay conditions. *Cancer cell* 21, 240-252.

Byler, S., Goldgar, S., Heerboth, S., Leary, M., Housman, G., Moulton, K., and Sarkar, S. (2014). Genetic and epigenetic aspects of breast cancer progression and therapy. *Anticancer research* 34, 1071-1077.

Cailleau, R., Olive, M., and Cruciger, Q.V. (1978). Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In vitro* 14, 911-915.

Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* *154*, 61-74.

Chaffer, C.L., and Weinberg, R.A. (2011). A perspective on cancer cell metastasis. *Science* *331*, 1559-1564.

Chakrabarti, R., Hwang, J., Andres Blanco, M., Wei, Y., Lukacisin, M., Romano, R.A., Smalley, K., Liu, S., Yang, Q., Ibrahim, T., *et al.* (2012). Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nature cell biology* *14*, 1212-1222.

Cho, I., Tsai, P.F., Lake, R.J., Basheer, A., and Fan, H.Y. (2013). ATP-dependent chromatin remodeling by Cockayne syndrome protein B and NAP1-like histone chaperones is required for efficient transcription-coupled DNA repair. *PLoS genetics* *9*, e1003407.

Chou, J., Lin, J.H., Brenot, A., Kim, J.W., Provot, S., and Werb, Z. (2013). GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. *Nature cell biology* *15*, 201-213.

Clark, E.A., Golub, T.R., Lander, E.S., and Hynes, R.O. (2000). Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* *406*, 532-535.

Cordenonsi, M., Zanconato, F., Azzolin, L., Forcato, M., Rosato, A., Frasson, C., Inui, M., Montagner, M., Parenti, A.R., Poletti, A., *et al.* (2011). The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* 147, 759-772.

De Vito, C., Riggi, N., Cornaz, S., Suva, M.L., Baumer, K., Provero, P., and Stamenkovic, I. (2012). A TARBP2-dependent miRNA expression profile underlies cancer stem cell properties and provides candidate therapeutic reagents in Ewing sarcoma. *Cancer cell* 21, 807-821.

DeHahn, K.C., Gonzales, M., Gonzalez, A.M., Hopkinson, S.B., Chandel, N.S., Brunelle, J.K., and Jones, J.C. (2004). The alpha4 laminin subunit regulates endothelial cell survival. *Experimental cell research* 294, 281-289.

Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., Viale, G., Delorenzi, M., Zhang, Y., d'Assignies, M.S., *et al.* (2007). 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* 13, 3207-3214.

Diehn, M., and Majeti, R. (2010). Metastatic cancer stem cells: an opportunity for improving cancer treatment? *Cell stem cell* 6, 502-503.

Domogatskaya, A., Rodin, S., and Tryggvason, K. (2012). Functional diversity of laminins. *Annual review of cell and developmental biology* 28, 523-553.

Ercan, C., van Diest, P.J., and Vooijs, M. (2011). Mammary development and breast cancer: the role of stem cells. *Current molecular medicine* 11, 270-285.

Espina, V., and Liotta, L.A. (2011). What is the malignant nature of human ductal carcinoma in situ? *Nature reviews Cancer* 11, 68-75.

Fidler, I.J. (1973). Selection of successive tumour lines for metastasis. *Nature: New biology* 242, 148-149.

Fidler, I.J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nature reviews Cancer* 3, 453-458.

Fidler, I.J. (2011). The biology of cancer metastasis. *Seminars in cancer biology* 21, 71.

Folkman, J. (2002). Role of angiogenesis in tumor growth and metastasis. *Seminars in oncology* 29, 15-18.

Fujita, M., Khazenzon, N.M., Bose, S., Sekiguchi, K., Sasaki, T., Carter, W.G., Ljubimov, A.V., Black, K.L., and Ljubimova, J.Y. (2005). Overexpression of beta1-chain-containing laminins in capillary basement membranes of human breast cancer and its metastases. *Breast cancer research : BCR* 7, R411-421.

Gao, H., Chakraborty, G., Lee-Lim, A.P., Mo, Q., Decker, M., Vonica, A., Shen, R., Brogi, E., Brivanlou, A.H., and Giancotti, F.G. (2012). The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell* 150, 764-779.

Gedye, C., and Ailles, L. (2013). Isolation and characterization of cancer stem cells in vitro. *Methods in molecular biology* 946, 181-204.

Ghajar, C.M., and Bissell, M.J. (2008). Extracellular matrix control of mammary gland morphogenesis and tumorigenesis: insights from imaging. *Histochemistry and cell biology* 130, 1105-1118.

Ghajar, C.M., Peinado, H., Mori, H., Matei, I.R., Evason, K.J., Brazier, H., Almeida, D., Koller, A., Hajjar, K.A., Stainier, D.Y., *et al.* (2013). The perivascular niche regulates breast tumour dormancy. *Nature cell biology* 15, 807-817.

Giancotti, F.G. (2013). Mechanisms governing metastatic dormancy and reactivation. *Cell* 155, 750-764.

Goering, W., Adham, I.M., Pasche, B., Manner, J., Ochs, M., Engel, W., and Zoll, B. (2008). Impairment of gastric acid secretion and increase of embryonic lethality in Foxq1-deficient mice. *Cytogenetic and genome research* 121, 88-95.

Gomis, R.R., Alarcon, C., Nadal, C., Van Poznak, C., and Massague, J. (2006). C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. *Cancer cell* 10, 203-214.

Goss, P.E., and Chambers, A.F. (2010). Does tumour dormancy offer a therapeutic target? *Nature reviews Cancer* 10, 871-877.

Grassian, A.R., Coloff, J.L., and Brugge, J.S. (2011). Extracellular matrix regulation of metabolism and implications for tumorigenesis. *Cold Spring Harbor symposia on quantitative biology* 76, 313-324.

Gumireddy, K., Li, A., Gimotty, P.A., Klein-Szanto, A.J., Showe, L.C., Katsaros, D., Coukos, G., Zhang, L., and Huang, Q. (2009). KLF17 is a negative regulator of epithelial-mesenchymal transition and metastasis in breast cancer. *Nature cell biology* 11, 1297-1304.

Gumireddy, K., Sun, F., Klein-Szanto, A.J., Gibbins, J.M., Gimotty, P.A., Saunders, A.J., Schultz, P.G., and Huang, Q. (2007). In vivo selection for metastasis promoting genes in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* 104, 6696-6701.

Gupta, G.P., Nguyen, D.X., Chiang, A.C., Bos, P.D., Kim, J.Y., Nadal, C., Gomis, R.R., Manova-Todorova, K., and Massague, J. (2007). Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* 446, 765-770.

Gusterson, B. (2009). Do 'basal-like' breast cancers really exist? *Nature reviews Cancer* 9, 128-134.

Hamill, K.J., Kligys, K., Hopkinson, S.B., and Jones, J.C. (2009). Laminin deposition in the extracellular matrix: a complex picture emerges. *Journal of cell science* 122, 4409-4417.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* 100, 57-70.

Hatzis, C., Pusztai, L., Valero, V., Booser, D.J., Esserman, L., Lluch, A., Vidaurre, T., Holmes, F., Souchon, E., Wang, H., *et al.* (2011). A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA : the journal of the American Medical Association* 305, 1873-1881.

He, G., Dhar, D., Nakagawa, H., Font-Burgada, J., Ogata, H., Jiang, Y., Shalapour, S., Seki, E., Yost, S.E., Jepsen, K., *et al.* (2013). Identification of liver cancer progenitors whose malignant progression depends on autocrine IL-6 signaling. *Cell* 155, 384-396.

Hedley, B.D., and Chambers, A.F. (2009). Tumor dormancy and metastasis. *Advances in cancer research* 102, 67-101.

Hu, Y., Ru, N., Xiao, H., Chaturvedi, A., Hoa, N.T., Tian, X.J., Zhang, H., Ke, C., Yan, F., Nelson, J., *et al.* (2013). Tumor-specific chromosome mis-segregation controls cancer plasticity by maintaining tumor heterogeneity. *PloS one* 8, e80898.



Hynes, R.O. (2009). The extracellular matrix: not just pretty fibrils. *Science* 326, 1216-1219.

Ilani, T., Alon, A., Grossman, I., Horowitz, B., Kartvelishvily, E., Cohen, S.R., and Fass, D. (2013). A secreted disulfide catalyst controls extracellular matrix composition and function. *Science* 341, 74-76.

Ishizawa, K., Rasheed, Z.A., Karisch, R., Wang, Q., Kowalski, J., Susky, E., Pereira, K., Karamboulas, C., Moghal, N., Rajeshkumar, N.V., *et al.* (2010). Tumor-initiating cells are rare in many human tumors. *Cell stem cell* 7, 279-282.

Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA: a cancer journal for clinicians* 61, 69-90.

Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010). Cancer statistics, 2010. *CA: a cancer journal for clinicians* 60, 277-300.

Joyce, J.A., and Pollard, J.W. (2009). Microenvironmental regulation of metastasis. *Nature reviews Cancer* 9, 239-252.

Kang, Y., Siegel, P.M., Shu, W., Drobnjak, M., Kakonen, S.M., Cordon-Cardo, C., Guise, T.A., and Massague, J. (2003). A multigenic program mediating breast cancer metastasis to bone. *Cancer cell* 3, 537-549.

Kessenbrock, K., Plaks, V., and Werb, Z. (2010). Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell* 141, 52-67.

Khanna, C., and Hunter, K. (2005). Modeling metastasis in vivo. *Carcinogenesis* 26, 513-523.

Kim, R.S., Avivar-Valderas, A., Estrada, Y., Bragado, P., Sosa, M.S., Aguirre-Ghiso, J.A., and Segall, J.E. (2012). Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. *PloS one* 7, e35569.

Korpai, M., Ell, B.J., Buffa, F.M., Ibrahim, T., Blanco, M.A., Celia-Terrassa, T., Mercatali, L., Khan, Z., Goodarzi, H., Hua, Y., *et al.* (2011). Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nature medicine* 17, 1101-1108.

Kusuma, N., Denoyer, D., Eble, J.A., Redvers, R.P., Parker, B.S., Pelzer, R., Anderson, R.L., and Pouliot, N. (2012). Integrin-dependent response to laminin-511 regulates breast tumor cell invasion and metastasis. *International journal of cancer Journal international du cancer* 130, 555-566.

Langley, R.R., and Fidler, I.J. (2011). The seed and soil hypothesis revisited--the role of tumor-stroma interactions in metastasis to different organs. *International journal of cancer Journal international du cancer* 128, 2527-2535.

Lee, S., Stewart, S., Nagtegaal, I., Luo, J., Wu, Y., Colditz, G., Medina, D., and Allred, D.C. (2012). Differentially expressed genes regulating the progression of ductal carcinoma in situ to invasive breast cancer. *Cancer research* 72, 4574-4586.

Leonard, G.D., and Swain, S.M. (2004). Ductal carcinoma in situ, complexities and challenges. *Journal of the National Cancer Institute* 96, 906-920.

Li, Q., Birkbak, N.J., Györfy, B., Szallasi, Z., and Eklund, A.C. (2011). Jetset: selecting the optimal microarray probe set to represent a gene. *BMC bioinformatics* 12, 474.

Llobet-Navas, D., Rodriguez-Barrueco, R., Castro, V., Ugalde, A.P., Sumazin, P., Jacob-Sendler, D., Demircan, B., Castillo-Martin, M., Putcha, P., Marshall, N., *et al.* (2014). The miR-424(322)/503 cluster orchestrates remodeling of the epithelium in the involuting mammary gland. *Genes & development* 28, 765-782.

Ma, X.J., Dahiya, S., Richardson, E., Erlander, M., and Sgroi, D.C. (2009). Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast cancer research : BCR* 11, R7.

Ma, X.J., Salunga, R., Tuggle, J.T., Gaudet, J., Enright, E., McQuary, P., Payette, T., Pistone, M., Stecker, K., Zhang, B.M., *et al.* (2003). Gene expression profiles of human breast cancer progression. *Proceedings of the National Academy of Sciences of the United States of America* 100, 5974-5979.

Magee, J.A., Piskounova, E., and Morrison, S.J. (2012). Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer cell* 21, 283-296.

Mehlen, P., and Puisieux, A. (2006). Metastasis: a question of life or death. *Nature reviews Cancer* 6, 449-458.

Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., Shu, W., Giri, D.D., Viale, A., Olshen, A.B., Gerald, W.L., and Massague, J. (2005). Genes that mediate breast cancer metastasis to lung. *Nature* 436, 518-524.

Miranti, C.K., and Brugge, J.S. (2002). Sensing the environment: a historical perspective on integrin signal transduction. *Nature cell biology* 4, E83-90.

Moreno-Layseca, P., and Streuli, C.H. (2014). Signalling pathways linking integrins with cell cycle progression. *Matrix biology : journal of the International Society for Matrix Biology* 34, 144-153.

Mori, T., Kariya, Y., Komiya, E., Higashi, S., Miyagi, Y., Sekiguchi, K., and Miyazaki, K. (2011). Downregulation of a newly identified laminin, laminin-3B11, in vascular basement membranes of invasive human breast cancers. *Cancer science* 102, 1095-1100.

Moyano, J.V., Greciano, P.G., Buschmann, M.M., Koch, M., and Matlin, K.S. (2010). Autocrine transforming growth factor- $\beta$ 1 activation mediated by integrin

$\alpha$ V $\beta$ 3 regulates transcriptional expression of laminin-332 in Madin-Darby canine kidney epithelial cells. *Molecular biology of the cell* 21, 3654-3668.

Nelson, C.M., and Bissell, M.J. (2006). Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. *Annual review of cell and developmental biology* 22, 287-309.

Nguyen, D.X., Bos, P.D., and Massague, J. (2009a). Metastasis: from dissemination to organ-specific colonization. *Nature reviews Cancer* 9, 274-284.

Nguyen, D.X., Chiang, A.C., Zhang, X.H., Kim, J.Y., Kris, M.G., Ladanyi, M., Gerald, W.L., and Massague, J. (2009b). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 138, 51-62.

Nguyen, L.V., Vanner, R., Dirks, P., and Eaves, C.J. (2012). Cancer stem cells: an evolving concept. *Nature reviews Cancer* 12, 133-143.

O'Brien, C.A., Kreso, A., and Jamieson, C.H. (2010). Cancer stem cells and self-renewal. *Clinical cancer research : an official journal of the American Association for Cancer Research* 16, 3113-3120.

Owens, T.W., and Naylor, M.J. (2013). Breast cancer stem cells. *Frontiers in physiology* 4, 225.

Padua, D., Zhang, X.H., Wang, Q., Nadal, C., Gerald, W.L., Gomis, R.R., and Massague, J. (2008). TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* 133, 66-77.

Pavlova, N.N., Pallasch, C., Elia, A.E., Braun, C.J., Westbrook, T.F., Hemann, M., and Elledge, S.J. (2013). A role for PVRL4-driven cell-cell interactions in tumorigenesis. *eLife* 2, e00358.

Peduto, L., Reuter, V.E., Sehara-Fujisawa, A., Shaffer, D.R., Scher, H.I., and Blobel, C.P. (2006). ADAM12 is highly expressed in carcinoma-associated stroma and is required for mouse prostate tumor progression. *Oncogene* 25, 5462-5466.

Pencheva, N., Tran, H., Buss, C., Huh, D., Drobnjak, M., Busam, K., and Tavazoie, S.F. (2012). Convergent multi-miRNA targeting of ApoE drives LRP1/LRP8-dependent melanoma metastasis and angiogenesis. *Cell* 151, 1068-1082.

Piao, H.L., Yuan, Y., Wang, M., Sun, Y., Liang, H., and Ma, L. (2014). alpha-catenin acts as a tumour suppressor in E-cadherin-negative basal-like breast cancer by inhibiting NF-kappaB signalling. *Nature cell biology* 16, 245-254.

Png, K.J., Halberg, N., Yoshida, M., and Tavazoie, S.F. (2012). A microRNA regulon that mediates endothelial recruitment and metastasis by cancer cells. *Nature* 481, 190-194.

Polyak, K. (2010). Molecular markers for the diagnosis and management of ductal carcinoma in situ. *Journal of the National Cancer Institute Monographs* 2010, 210-213.

Polyak, K., and Kalluri, R. (2010). The role of the microenvironment in mammary gland development and cancer. *Cold Spring Harbor perspectives in biology* 2, a003244.

Ponomarev, V., Doubrovin, M., Serganova, I., Vider, J., Shavrin, A., Beresten, T., Ivanova, A., Ageyeva, L., Tourkova, V., Balatoni, J., *et al.* (2004). A novel triple-modality reporter gene for whole-body fluorescent, bioluminescent, and nuclear noninvasive imaging. *European journal of nuclear medicine and molecular imaging* 31, 740-751.

Powles, T.J. (2002). Anti-oestrogenic prevention of breast cancer--the make or break point. *Nature reviews Cancer* 2, 787-794.

Prince, M.E., Sivanandan, R., Kaczorowski, A., Wolf, G.T., Kaplan, M.J., Dalerba, P., Weissman, I.L., Clarke, M.F., and Ailles, L.E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 104, 973-978.

Putti, T.C., El-Rehim, D.M., Rakha, E.A., Paish, C.E., Lee, A.H., Pinder, S.E., and Ellis, I.O. (2005). 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* 18, 26-35.

Qiao, Y., Jiang, X., Lee, S.T., Karuturi, R.K., Hooi, S.C., and Yu, Q. (2011). FOXQ1 regulates epithelial-mesenchymal transition in human cancers. *Cancer research* 71, 3076-3086.

Quintana, E., Shackleton, M., Sabel, M.S., Fullen, D.R., Johnson, T.M., and Morrison, S.J. (2008). Efficient tumour formation by single human melanoma cells. *Nature* 456, 593-598.

Reymond, N., d'Agua, B.B., and Ridley, A.J. (2013). Crossing the endothelial barrier during metastasis. *Nature reviews Cancer* 13, 858-870.

Rhim, A.D., Mirek, E.T., Aiello, N.M., Maitra, A., Bailey, J.M., McAllister, F., Reichert, M., Beatty, G.L., Rustgi, A.K., Vonderheide, R.H., *et al.* (2012). EMT and dissemination precede pancreatic tumor formation. *Cell* 148, 349-361.

Schatten, H. (2013). *Cell and molecular biology of breast cancer* (New York: Springer).

Scheele, S., Nystrom, A., Durbeej, M., Talts, J.F., Ekblom, M., and Ekblom, P. (2007). Laminin isoforms in development and disease. *Journal of molecular medicine* 85, 825-836.



Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods* 9, 671-675.

Schuetz, C.S., Bonin, M., Clare, S.E., Nieselt, K., Sotlar, K., Walter, M., Fehm, T., Solomayer, E., Riess, O., Wallwiener, D., *et al.* (2006). 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* 66, 5278-5286.

Shibue, T., and Weinberg, R.A. (2009). 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* 106, 10290-10295.

Shimono, Y., Zabala, M., Cho, R.W., Lobo, N., Dalerba, P., Qian, D., Diehn, M., Liu, H., Panula, S.P., Chiao, E., *et al.* (2009). Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* 138, 592-603.

Simpson, C.D., Anyiwe, K., and Schimmer, A.D. (2008). Anoikis resistance and tumor metastasis. *Cancer letters* 272, 177-185.

Song, S.J., Poliseno, L., Song, M.S., Ala, U., Webster, K., Ng, C., Beringer, G., Brikbak, N.J., Yuan, X., Cantley, L.C., *et al.* (2013). MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* 154, 311-324.

Stenzel, D., Franco, C.A., Estrach, S., Mettouchi, A., Sauvaget, D., Rosewell, I., Schertel, A., Armer, H., Domogatskaya, A., Rodin, S., *et al.* (2011). Endothelial basement membrane limits tip cell formation by inducing Dll4/Notch signalling in vivo. *EMBO reports* 12, 1135-1143.

Sun, H.T., Cheng, S.X., Tu, Y., Li, X.H., and Zhang, S. (2013). FoxQ1 promotes glioma cells proliferation and migration by regulating NRXN3 expression. *PloS one* 8, e55693.

Tabaries, S., Dupuy, F., Dong, Z., Monast, A., Annis, M.G., Spicer, J., Ferri, L.E., Omeroglu, A., Basik, M., Amir, E., *et al.* (2012). Claudin-2 promotes breast cancer liver metastasis by facilitating tumor cell interactions with hepatocytes. *Molecular and cellular biology* 32, 2979-2991.

Talmadge, J.E., and Fidler, I.J. (2010). AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer research* 70, 5649-5669.

Tavazoie, S.F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P.D., Gerald, W.L., and Massague, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 451, 147-152.

Thomas, C., and Gustafsson, J.A. (2011). The different roles of ER subtypes in cancer biology and therapy. *Nature reviews Cancer* 11, 597-608.

Thomas, P.A. (2011). Breast cancer and its precursor lesions : making sense and making it early (New York: Humana).

Thyboll, J., Kortessmaa, J., Cao, R., Soininen, R., Wang, L., Iivanainen, A., Sorokin, L., Risling, M., Cao, Y., and Tryggvason, K. (2002). Deletion of the laminin alpha4 chain leads to impaired microvessel maturation. *Molecular and cellular biology* 22, 1194-1202.

van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., *et al.* (2002). A gene-expression signature as a predictor of survival in breast cancer. *The New England journal of medicine* 347, 1999-2009.

Vanharanta, S., and Massague, J. (2013). Origins of metastatic traits. *Cancer cell* 24, 410-421.

Vargo-Gogola, T., and Rosen, J.M. (2007). Modelling breast cancer: one size does not fit all. *Nature reviews Cancer* 7, 659-672.

Vermeulen, L., De Sousa, E.M.F., van der Heijden, M., Cameron, K., de Jong, J.H., Borovski, T., Tuynman, J.B., Todaro, M., Merz, C., Rodermond, H., *et al.* (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nature cell biology* 12, 468-476.

Visvader, J.E., and Lindeman, G.J. (2012). Cancer stem cells: current status and evolving complexities. *Cell stem cell* 10, 717-728.

Wallquist, W., Plantman, S., Thams, S., Thyboll, J., Kortessmaa, J., Lannergren, J., Domogatskaya, A., Ogren, S.O., Risling, M., Hammarberg, H., *et al.* (2005). Impeded interaction between Schwann cells and axons in the absence of laminin alpha4. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 25, 3692-3700.

Wan, L., Pantel, K., and Kang, Y. (2013). Tumor metastasis: moving new biological insights into the clinic. *Nature medicine* 19, 1450-1464.

Wang, Y., Klijn, J.G., Zhang, Y., Sieuwerts, A.M., Look, M.P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M.E., Yu, J., *et al.* (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365, 671-679.

Watanabe, T.K., Fujiwara, T., Nakamura, Y., Hirai, Y., Maekawa, H., and Takahashi, E. (1996). Cloning, expression pattern and mapping to Xq of NAP1L3, a gene encoding a peptide homologous to human and yeast nucleosome assembly proteins. *Cytogenetics and cell genetics* 74, 281-285.

Welm, A.L. (2008). TGFbeta primes breast tumor cells for metastasis. *Cell* 133, 27-28.

Williams, S.A., Anderson, W.C., Santaguida, M.T., and Dylla, S.J. (2013). Patient-derived xenografts, the cancer stem cell paradigm, and cancer pathobiology in the 21st century. *Laboratory investigation; a journal of technical methods and pathology* 93, 970-982.

Wirtz, D., Konstantopoulos, K., and Searson, P.C. (2011). The physics of cancer: the role of physical interactions and mechanical forces in metastasis. *Nature reviews Cancer* 11, 512-522.

Wondimu, Z., Geberhiwot, T., Ingerpuu, S., Juronen, E., Xie, X., Lindbom, L., Doi, M., Kortessmaa, J., Thyboll, J., Tryggvason, K., *et al.* (2004). An endothelial laminin isoform, laminin 8 (alpha4beta1gamma1), is secreted by blood neutrophils, promotes neutrophil migration and extravasation, and protects neutrophils from apoptosis. *Blood* 104, 1859-1866.

Xu, L., Shen, S.S., Hoshida, Y., Subramanian, A., Ross, K., Brunet, J.P., Wagner, S.N., Ramaswamy, S., Mesirov, J.P., and Hynes, R.O. (2008). Gene expression changes in an animal melanoma model correlate with aggressiveness of human melanoma metastases. *Molecular cancer research : MCR* 6, 760-769.

Yu, M., Stott, S., Toner, M., Maheswaran, S., and Haber, D.A. (2011). Circulating tumor cells: approaches to isolation and characterization. *The Journal of cell biology* 192, 373-382.

Yurchenco, P.D. (2011). Basement membranes: cell scaffoldings and signaling platforms. Cold Spring Harbor perspectives in biology 3.

Zhang, H., Meng, F., Liu, G., Zhang, B., Zhu, J., Wu, F., Ethier, S.P., Miller, F., and Wu, G. (2011a). Forkhead transcription factor foxq1 promotes epithelial-mesenchymal transition and breast cancer metastasis. Cancer research 71, 1292-1301.

Zhang, X.H., Jin, X., Malladi, S., Zou, Y., Wen, Y.H., Brogi, E., Smid, M., Foekens, J.A., and Massague, J. (2013a). Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. Cell 154, 1060-1073.

Zhang, X.T., Kang, L.G., Ding, L., Vranic, S., Gatalica, Z., and Wang, Z.Y. (2011b). A positive feedback loop of ER- $\alpha$ 36/EGFR promotes malignant growth of ER-negative breast cancer cells. Oncogene 30, 770-780.

Zhang, Y., Yang, P., Sun, T., Li, D., Xu, X., Rui, Y., Li, C., Chong, M., Ibrahim, T., Mercatali, L., *et al.* (2013b). miR-126 and miR-126\* repress recruitment of mesenchymal stem cells and inflammatory monocytes to inhibit breast cancer metastasis. Nature cell biology 15, 284-294.

Zhou, W., Lv, R., Qi, W., Wu, D., Xu, Y., Liu, W., Mou, Y., and Wang, L. (2014). Snail contributes to the maintenance of stem cell-like phenotype cells in human pancreatic cancer. PloS one 9, e87409.