UNIVERSITY OF CALFIORNIA, IRVINE

Genetic and Metabolic Mechanisms of Breast Cancer Metastasis

DISSERTATION

Submitted in partial satisfaction for the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Biomedical Sciences

by

Paige Vail Halas

Dissertation Committee:
Associate Professor, Devon A. Lawson, Chair
Associate Professor, Olga V. Razorenova
Associate Professor, Selma Masri
Assistant Professor, Francesco Marangoni
Associate Professor, Dorota Skowronska-Krawczyk

2024

Dedication

To

my family,

in gratitude for their constant support during my journey toward this degree,

my fiancé,

for being my unwavering support after long days of experiments, both successful and challenging,

and my friends,

for their patience in listening to the many trials and tribulations of my graduate school journey,

in recognition of their worth

motivation

“You are Stronger than your Challenges and your Challenges are Making you Stronger”

Karen Salmansohn

“And if you never bleed, you're never gonna grow. And it's alright now”

Taylor Swift

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to start by expressing my gratitude to my graduate advisor, Devon Lawson. When I joined the lab at the start of the COVID-19 pandemic, you took a chance on me, despite having never seen me pipet or dissect a mouse. Over the past five and a half years, you've guided my growth into an independent scientist, a collaborative teammate, and a supportive mentor. You've encouraged my curiosity and taught me the importance of knowing when it's time to pivot. Your support extended beyond the lab, fostering my passion for scientific communication and connecting me with impactful community organizations like Team Michelle, which truly add meaning to the work we do. I am so grateful for the mentorship you provided me and the environment you built in your lab. It's a special place that I am going to miss dearly.

I'm also grateful for the Lawson-Kessenbrock lab's unwavering support throughout my PhD journey. From the early training I received from Ryan, Dennis, Jacob, Maren, and Nick to the daily professional and personal encouragement from Aaron, Isam, Lincy, Pascal, Tim, Jacob, Sharmila, Jessica, Tatyana, Alex, Noah, and Carina, I feel incredibly fortunate to have been part of such an amazing team. I also want to thank Dennis for his guidance and for generously passing on his project during a time when my first PhD project faced significant challenges. I especially want to acknowledge Hannah—not only for your invaluable help on my thesis project but also for becoming one of my closest friends over the past two years. From long days working on experiments to the day you helped me pick out my wedding dress, our friendship has been a truly special part of this journey. This project wouldn't have become the story it is without you, and I'm so grateful to have had you by my side for all of it. Leaving this group is a bittersweet part of moving on, and I will truly cherish my time working with all of you. I'll miss our long lunches, coffee walks and brewery nights where we'd chat about anything and everything, and I'm especially grateful for you all sticking through my endless Taylor Swift chatter. Thank you for making graduate school so much more bearable, and for the friendships I've built with you all during my time at UCI.

I would like to extend my heartfelt thanks to all the faculty and staff who have contributed to my training and success at UCI. A special thank you goes to Pauline and Vanessa at the Stem Cell Research Core for always helping me navigate tricky flow gates when I struggled to find my cells, tackling crazy cleaning protocols for metabolomics, and keeping me company during long sorting days. I also want to express my gratitude to my collaborators, Cholsoon Jang and Chris Hughes, as well as their researchers, Johnny Le and Stephanie Hachey. You all have taught me so much and played a significant role in my project's success.

I am also grateful to my undergraduate mentee, Connie Chan. It was such an honor to mentor you during my PhD, watching you grow and learning from you in the process. I'm incredibly proud of all we accomplished together and so excited for you as you continue in Pharmacy School at USC.

I want to express my gratitude to my mentors, Dr. Cristie Andrade, Dr. Ulrich Schlecht, and Dr. Severine Margeridon-Thermet, who encouraged me to pursue graduate education. You have all been scientific role models, guiding me in my career and providing opportunities for which I'm incredibly thankful. I also want to acknowledge my committee members—Dr. Olga Razorenova, Dr. Selma Masri, Dr. Francesco Marangoni, and Dr. Dorota Skowronska-Krawczyk—for their valuable guidance and support throughout my journey. Thank you for reading my thesis and offering thoughtful feedback and advice.

I want to extend my deepest thanks to my friends for their unwavering support throughout this journey. I appreciate you all for proudly calling me "your science friend" and for always being there to listen when I needed to vent, even if the details were confusing or unfamiliar. You've celebrated my achievements, big and small—even when you didn't quite understand what they meant—and for the few who also took the wild leap into grad school, thank you for sharing in the craziness. And, of course, thanks for tolerating my rants about working with mice.

I want to express my heartfelt thanks to my parents for everything they've done for me throughout my academic career. You made sure I had a good education and provided the resources I needed to achieve my goals. Additionally, your support in my pursuit of higher-level education has been unwavering from deciding where to attend undergrad, to selecting a graduate program, to deciding to pursue a PhD. Additionally, I'm grateful for the way you encouraged me to explore creative outlets, which still keep me grounded to this day. I also want to thank my Aunt Karen for her unwavering support along my academic journey. From reading my college essays to answering countless career-related questions, you've always been there for me. As a woman in science, you've been a true role model, and I'm incredibly thankful for your guidance and encouragement.

I want to thank my fiancé, Matthew. We met just as I was finishing my Master's, and you've been there from the very start of my PhD journey—starting with making me my favorite dinner of tacos the night before my interviews. I'm so grateful for your constant support and for enduring the day-to-day craziness of graduate school. You've shown such patience and gentleness as I navigated the ups and downs, from experiments that didn't cooperate to building my self-confidence through it all. I'm so excited to take the next steps in our lives together—and now you can say you're officially marrying a doctor (just not the kind who can prescribe anything)!

Finally, I want to give a special acknowledgment to Taylor Swift, whose music has been my constant companion since high school. Entering my doctoral era alongside Taylor, now a fellow "doctor," has been especially meaningful—her music has brought comfort, inspiration, and motivation every step of the way.

I would like to thank the financial support of Team Michelle and the National Cancer Institute under Award Number T32CA009054 for this research. I would like I would like to thank the Gazzaniga Family for their support of my research and the ARCS Foundation for their personal support of my graduate work.

Vita

Paige Vail Halas

EDUCATION

University of California, Irvine School of Medicine Department of Physiology and Biophysics	Doctor of Philosophy, Biomedical Sciences	2019-2024
University of California, Irvine School of Biological Sciences Department of Molecular Biology and Biochemistry	Masters of Science, Biotechnology	2017-2019
Willamette University	Bachelors of Arts, Chemistry (Biochemistry Track)	2013-2017

PUBLICATIONS

1. In submission: **Halas PV**, Savage HL, et al. PHLDA2 promotes breast cancer metastasis by increasing extracellular matrix deposition and vascular permeability. 2024

2. Chakraborty M, Ramaiah A, Adolfi A, **Halas P,** Kaduskar B, Ngo LT, Jayaprasad S, Paul K, Whadgar S, Srinivasan S, Subramani S, Bier E, James AA, Emerson JJ. Hidden genomic features of an invasive malaria vector, Anopheles stephensi, revealed by a chromosome-level genome assembly. BMC Biol. 2021 Feb 10;19(1):28. PubMed Central PMCID: PMC7876825.

FUNDING AND AWARDS

Gazzaniga Family Medical Research Award ($6,000)	2024
ARCS Foundation Scholar Award: UCI School of Medicine ($10,000)	2023-2024
UC Irvine: Undergraduate Research Opportunities Graduate Fellowship ($5,000)	2022
UC Irvine Minority Serving Institution Enhancement Fellowship (MSIE) ($10,000)	2020-2021
Willamette University Academic Leader Award Scholarship	2013-2017

SELECTED PRESENTATIONS

"Investigating PHLDA2 as a Genetic Driver of Breast Cancer Metastasis" 2024
Metastasis Research Society 20th Biennial Congress, London, England
Poster Presentation

"PHLDA2, a travel agent for cancer cells" 2024
UC Irvine Grad Slam-Semi-Finals and Final Pitch Competition Irvine, CA
Oral Presentation

"*Investigating PHLDA2 as a Genetic Driver of Breast Cancer Metastasis"* 2023

AACR, Special Conferences, Advances in Breast Cancer, San Diego, CA
Poster Presentation

"*Investigating PHLDA2 as a Genetic Driver of Breast Cancer Metastasis"* 2023
Cancer Research Institute Symposium, University of California Irvine. Irvine, CA
Poster Presentation

"Metabolic Mediators of Breast Cancer Metastasis" 2022
Cancer Research Institute Retreat, University of California Irvine. San Diego, CA
Poster Presentation

"Metabolic Mediators of Breast Cancer Metastasis" 2022
Cancer Research Institute Retreat, University of California Irvine. San Diego, CA
Oral Presentation

TEACHING AND INVOLVEMENT

TEACHING ASSISTANT

2019

Bio99: Introductory Biology Teaching Assistant
Lead two discussion sections weekly that accompanied course lectures.
Graded exams and discussion assignments.

MB 118L: Microbiology Laboratory Teaching Assistant 2018-2019
Lead lab course for microbiology class. Managed about 30 students at a time in the lab space. Responsible for grading all laboratory assignments, administering practical exams and grading evaluations taken in accompanying lecture class.

LEADERSHIP ROLES AND CAMPUS INVOLVEMENT

UCI Graduate Professional Success in Stem (GPS-STEM) Program Member 2017-2024

Associated Women in Science (AWIS) UC Irvine Chapter Member 2017-2024

Social Media Managing Editor, Loh Down on Science Podcast 2023-2024

Graduate Mentor, Undergraduate Research Opportunities Program 2022-2024
University of California, Irvine

Writer, Loh Down on Science Podcast 2021-2023

Graduate Mentor, UCI Cancer Research Institute Youth Science Fellowship 2022
Program University of California, Irvine

Founder and Ambassador, UC Irvine Cellular and Molecular Biology Gateway 2020-2021
Peer Mentorship Program

Abstract of the Dissertation

Genetic and Metabolic Mechanisms of Breast Cancer Metastasis

By

Paige Vail Halas

Doctor of Philosophy in Biomedical Sciences

University of California, Irvine, 2024

Assistant Professor Devon A. Lawson, Chair

Metastasis is a major determinant of patient survival in cancer, yet the genetic and metabolic mechanisms underlying metastatic progression remain poorly understood. In our previous work, we identified *PHLDA2*, a gene not previously implicated in breast cancer metastasis, as a transcriptional marker of metastatic breast cancer. Analyzing patient datasets, we found this gene is more highly expressed in breast tumor tissue than normal tissue. This elevated expression is likely driven by methylation, as *PHLDA2* is an imprinted gene, indicating a potential role for epigenetic regulation in its overactivation in cancer. Patient datasets revealed that high *PHLDA2* expression is associated with reduced relapse-free and distal metastasis-free survival, underscoring its prognostic significance. Building on these findings, this thesis explores *PHLDA2* as a driver of metastatic burden, demonstrating that it promotes breast cancer metastasis by restructuring the extracellular matrix and increasing vessel permeability, thereby reshaping the metastatic microenvironment. Additionally, to better study the metabolic processes driving aggressive cancer cell behavior at metastatic sites, I developed a single-cell metabolomics approach. This work identifies citrate as a potential metabolic drivers of cancer metastasis using a novel tool to study cancer metabolism at a single-cell resolution

Chapter 1: Introduction

Chapter 1.1: Breast Cancer and the Metastatic Cascade

Breast cancer patient mortality is driven by untreatable metastatic lesions. Breast cancer is the most frequently diagnosed cancer in women globally and remains the leading cause of cancer-related mortality[1]. Patient survival declines significantly as breast cancer progresses from localized disease in the breast to systemic metastatic disease. Due to its aggressive and recurrent nature, breast cancer metastasis is responsible for 90% of patient deaths[2]. The metastatic process is quite complex and genetic and metabolic drivers of breast cancer metastasis are largely unidentified.

During the metastatic cascade, cancer cells in the primary tumor in the breast will invade the surrounding tissue, allowing for intravasation into the microvasculature of the lymphatic and circulatory systems[2]. These cells survive and travel through the bloodstream to secondary sites in distal tissues[2]. At these secondary sites, cells extravasate from the blood stream into microvessels of distant tissues. These cells then adapt to the microenvironment in secondary sites and proliferate to colonize secondary lesions[2]. The metastatic cascade is largely inefficient due to numerous biological barriers; however, aggressive metastatic lesions possess the ability to overcome these challenges, enabling successful colonization at distant sites. In the context of breast cancer, cancer cells preferentially metastasize in the lung, liver, bone and brain[3]. In this study, we utilized models of breast cancer lung metastasis, incorporating both spontaneous and experimental metastasis approaches for experimentation.

Cancer cells possess key characteristics that drive their aggression such as immortality, genomic instability, cell death resistance, altered metabolism and ability to invade and metastasize[4,5]. These cells interact with their environment to promote processes such angiogenesis, inflammation, growth inhibition resistance and enabling immortality to promote abnormal growth[4,5]. These processes have been defined as the Hallmarks of Cancer[4,5]. The

differentiation between a benign tumor and cancer is the ability of at least one cell to invade through the basement membrane. However, for metastatic cancer cells, invasion alone is not sufficient to establish metastatic disease. These cells must be able to survive independently of the host and adapt a new microenvironment, driving the acquisition of hallmarks that drive invasive cancer cell behavior[4,5]. To successfully adapt to this new microenvironment, cancer cells rely on the extracellular matrix (ECM), which plays a pivotal role in supporting cellular functions and facilitating the signaling pathways essential for their survival and progression.

Chapter 1.2: Extracellular Matrix and Vasculature: Key Players in Cancer Progression and Metastasis

The ECM is made up of interstitial elements within tissues and organs to provide structural support for cell growth, cellular adhesion, production of growth factors and promote various signaling pathways during homeostasis[6,7]. As a major component of the tumor microenvironment (TME), the dysregulation of the ECM has been extensively studied for its role in cancer progression from primary tumor formation to metastasis[6–8]. Through biochemical and structural modifications, the ECM impacts cancer cell growth and metastasis[7]. The ECM is comprised of three layers; the glycocalyx, the basement membrane and the stromal ECM[7]. The glycocalyx plays a role in cellular interaction with its surroundings through glycoproteins, proteoglycans and glycolipids that cover the surface of epithelial and endothelial cells[9]. The basement membrane contains laminins and collagen IV for the purpose of separating epithelial layers from the stromal ECM[7]. The stromal ECM is made up of glycoproteins such as collagen, elastin, hyaluronic acid and fibronectin[7]. These three layers are all influenced and altered by the presence of tumor cells, which disrupt their normal structure and function, facilitating cancer progression and metastasis.

In cancer, the stomal ECM undergoes significant remodeling during disease which promotes tumor growth through increased density and stiffness[7]. ECM remodeling correlates with tumor invasiveness and poor clinical outcome[7]. As collagen is the most abundant component of

the ECM, it has been observed that increased collagen deposition provides a structural support for tumor formation in primary and secondary tumor sites[7]. Once cancer cells metastasize to a secondary site, to form a secondary lesion these cells must acclimate to a new microenvironment. To do this, the ECM, particularly through deposition of collagen, establishes a new environment for cancer cells in a secondary organ[7]. This has been demonstrated in both the pre-metastatic and metastatic niche of secondary tumor sites[7]. Through mechanisms of production/deposition, cross-linking, degradation and remodeling of the ECM, cancer cells can be supported in their new environments, promoting secondary tumor growth[7].

However, these ECM changes are not limited to being an advantage in tumor sites. In the metastatic cascade, cancer cells must invade the vasculature, survive in the blood and exit the blood vessels at distal locations[7]. In cancer, the basement membrane integrity is disrupted, allowing for disruption of the vascular basement membrane and increased vessel permeability[7]. Angiogenesis, the formation of new blood vessels, is often promoted by tumors to ensure a supply of nutrients and oxygen[10]. This newly formed vasculature, which is frequently leaky, facilitates metastasis by enabling cancer cells to enter the circulation[10]. To enter circulation through invasion, cancer cells degrade basement membrane vessel junctions and ECM components using the secretion of matrix metalloproteinases (MMPs)[7]. The glycocalyx additionally regulates cell adhesion before intravasation by allowing tumor cell binding to endothelial cells[7]. These interactions assist cancer cells in attaching to the epithelial lining before entering the blood[7]. In extravasation when cancer cells exit the blood stream and invade surrounding tissues, the glycocalyx is again altered causing increased vessel permeability[7]. During extravasation, the basement membrane is further degraded by MMPs to allow for cancer cell escape into secondary sites[7]. Overall, the dysregulation of the ECM during cancer supports tumor growth and metastasis.

Chapter 1.3: Heterogeneity and Therapeutic Challenges in Breast Cancer Metastasis

Given the pivotal role of the ECM in supporting tumor progression, its influence is particularly evident in the context of breast cancer, which is highly heterogeneous, comprising four main molecular subtypes: Luminal A, Luminal B, HER2+, and Triple Negative (TNBC). Along with primary tumor heterogeneity, metastatic capability and location are also highly variable among breast cancer subtypes[3]. Part of this heterogeneity can be attributed to the differences in ECM alteration between breast cancer subtypes[11]. For example, more aggressive subtypes, Basal and HER2+, been demonstrated to the have increased stiffness compared to less aggressive Luminal A and B subtypes[11]. Breast cancer subtypes are determined according to a gene expression pattern based on the presence or absence of estrogen receptor (ER), progesterone receptor (PR) or human epidermal growth factor receptor positive (HER2)[12,13]. These molecular markers can be leveraged for targeted treatments of primary tumors and metastatic lesions such as estrogen targeted hormone therapy or Herceptin targeting HER2+ cancers[14]. Given the lack of targetable receptors in TNBC, this model was employed in our previous work to identify transcriptional markers of metastasis and continues to serve as a foundation for this current work.

Regardless of the metastatic site, once TNBC tumors establish themselves in secondary locations, there are no targeted approaches to treat these lesions. Current treatments for receptor negative cancer cells primarily rely on systemic chemotherapeutic strategies, which not only target cancer cells but also harm healthy tissues, leading to significant side effects[14]. Targeted therapeutics that specifically address markers of metastatic breast cancer cells are crucial for improving patient outcomes while minimizing toxicity[15]. However, to discover novel therapeutic targets for metastatic lesions, a deeper understanding of the key drivers of breast cancer metastasis is required. The purpose of this work was to identify genetic and metabolic drivers of breast cancer metastasis to better understand how TNBC breast cancer cells metastasize to the

lung. To achieve this, we evaluated the role of the gene *PHLDA2* in promoting breast cancer lung metastasis. Additionally, we took an untargeted approach to identify metabolic markers of breast cancer metastasis using a novel single cell metabolomics approach.

Chapter 1.4: The Role of PHLDA2 in Placental Development and Disease

Our previous work defined the unique transcriptomic programs of primary tumor and metastatic cancer cells in previously established human patient-derived-xenograft models (PDX) of TNBC using single-cell RNA sequencing[16,17]. Using three PDX models of TNBC donated by the Huntsman Cancer Institute, cancer cells were injected into the mammary fat pad of immune-deficient NOD-SCID Gamma (NSG) mice[17]. Primary tumors were grown and allowed to metastasize to secondary sites including the lung and lymph nodes[17]. Using flow sorting, human cells were isolated from each tissue and single cell RNA-Sequencing was conducted[17]. In all three PDX models, gene pleckstrin homology like domain family A member 2 (*PHLDA2*) was highly expressed in metastatic tissues, but was absent in primary tumors[17]. RNA scope confirmed that *PHLDA2* transcripts were twofold higher in metastatic lung tissues than primary tumors[17]. Using a biomarker identification algorithm, *PHLDA2* was a top biomarker for identifying metastatic cells in eight of ten computational models[17]. However, the data did not establish whether *PHLDA2* is correlated with metastasis or contributes to its progression.

PHLDA2, also referred to as *TSSC3,* is in the tumor suppressor region of human chromosome 11. This location also contains many imprinted genes. Both parents contribute genetically to their offspring, but in the process of genomic imprinting, monoallelic gene expression occurs according to parental origin[18]. *PHLDA2* is a maternally imprinted gene, meaning that the maternal allele is transcriptionally active, and the paternal allele is silenced[19]. Imprinting, achieved through epigenetic modifications, ensures the regulated expression of genes[18]. This regulation is critical as these genes play a vital role in fetal development and

placental biology, supporting the rapid growth required during these processes [18]. Reflecting this function, *PHLDA2* has been found to be highly expressed in the murine placenta[19].

Disruption of expression of imprinted genes has been demonstrated to cause human diseases such as growth restriction, obesity and cancer[18]. A murine study determined that global loss of imprinting using embryonic stem cells leads to tumorigenesis in a variety of tissues[20]. Additionally, using chimeric models this study demonstrated that loss of imprinting alone can promote tumorigenesis[20]. In humans, dysregulated imprinting that results from somatic or germline events are associated with cancer risk[21,22]. These diseases often arise from dysregulation of DNA methylation as it plays a critical role in establishment and maintenance of imprinted genes[22].

Epigenetic changes in breast cancer subtypes are associated with subtype classification, cancer stages, and the prediction of clinical outcomes[23]. TNBC is considered broadly hypomethylated based on low levels of methylation compared to other subtypes and normal breast tissue[23]. Additionally, metastatic risk and poor survival were correlated with TNBC tumors lacking Cpg island methylation sites[23,24]. It has also been shown that patients with hypomethylated TNBC tumors have worse overall survival[23,25]. DNA methylation is regulated by DNA methylases (DNMTs) and DNA demethylases (TETs)[23]. When overexpressed, the TET family of enzymes, have been found to demethylate genes[23]. In investigating what drives hypomethylation in TNBC tumors, it was found that TET1 is upregulated in about 40% of TNBC patients and has been shown to increase oncogenic signaling through demethylation[23,25]. Experimental knock out of TET1 in TNBC cell lines, MDA-MB-231 and Hs578T, supports the oncogenic potential of this enzyme *in vitro* through reduction of cell migration and proliferation[23,25]. This suggests that the aggressive nature of TNBC can be attributed to its hypomethylation that drives its metastatic capabilities.

Due to its imprinted nature and abundant expression in the placenta, *PHLDA2* has largely been studied in the context of placental development using murine models. The murine placenta

contains two complementary layers to support fetal growth, the spongiotrophoblast and the labyrinth. The spongiotrophoblast is a cellular layer in the placenta that supports fetal development through the exchange of hormones and glycogen energy between mother and fetus[26–28]. *PHLDA2* has been shown to regulate placental demands for maternal resources in the spongiotrophoblast, functioning as a rheostat in placental development[28]. Loss of *PHLDA2* imprinting caused by the deletion of *Kvdmr1*—a regulatory region for imprinted, maternally expressed genes—leads to a significant reduction in the spongiotrophoblast layer[26,29]. *PHLDA2* overexpression reduces the size of the placental layer as well as reduces energy allocated to the fetus, while gene deletion leads to placental overgrowth[28]. Using murine models, it has been determined that *PHLDA2* is the cause of fetal growth restriction[19]. This finding is reflected in human patients with whom PHLDA2 overexpression correlates with fetal growth restriction[19]. Overall, these studies demonstrate that *PHLDA2* is critical for proper placental development.

Additionally, *PHLDA2* has been shown to be a transcriptomic marker of the labyrinth, a highly vascularized layer of the placenta that facilitates blood exchange[30]. This layer of the placenta acts as a filter, allowing for transfer of oxygen from mother to fetus and transfer of waste from fetus back to the mother[30]. As embryonic development progressed, this study identified that the labyrinth exhibited an upregulation of transcriptomic pathways related to vessel development and angiogenesis[30]. Highlighting these pathways supports the function of the labyrinth as a complex matrix of blood vessels[30].

PHLDA2 encodes a small protein (17 kDa) of the same name which contains a pleckstrin homology (PH) domain and an alanine repeat chain. Alanine repeats have been shown to marginally increase protein stability, but can also assist with protein-protein interactions[31,32]. PHLDA2 is part of a pleckstrin homology domain (PHLD) class of proteins[33]. The pleckstrin homology domain (PH) is highly conserved among eukaryotic species and is abundant in the human proteome[33]. PH domain-containing proteins were initially discovered through their ability to bind specifically to the phosphorylated head group of phosphatidylinositol 4,5-bisphosphate

(PtdIns(4,5)P2) *in vitro*[33]. However, less than 10% of proteins with PH domains show strong specificity for phosphorylated phosphatidylinositol lipids[33]. Instead, most of these proteins bind to various phosphorylated phosphatidylinositol lipids, as long as the lipids have adjacent phosphate groups on their inositol head such as PtdIns(1,4,5)P3, PtdIns(3,4,5)P3)[33]. The PHLD class of proteins is divided into two families, PHLDA and PHLDB [33]. These families differ based on the position of the PH domain on the N or C-terminus of the protein, with PHLDA PH domains positioned on the N-terminus and PHLDB PH domains located on the C-terminus [33]. The PHLDA family proteins, particularly PHLDA1, are more well studied than PHLDB proteins[33]. Within the PHLDA family, there are three paralogous genes that have diverted through evolution, *PHLDA1*, *PHLDA2* and *PHLDA3*, which are all located on different chromosomes[33,34]. However, they all share similar gene organization with a PH domain two exons, one of which is coding, and a small intron[34].

PHLDA1, is about 45 kDa and contains a N-terminal PH domain in addition to a polyglutamine and proline histidine repeat[33]. This gene is highly expressed in all human tissues but is most elevated in salivary glands[33]. This gene is located on human chromosome 12 and has not been reported to be imprinted, like its paralogous relative *PHLDA2*[34]. This protein is the most studied in the PHLDA family and has been studied in cancer. PHLDA1 was first identified as a transcription factor in murine T cell hybridomas where it was found to be necessary for Fas expression and T-cell activation induced apoptosis[34,35]. However, since it's identification, there has been conflicting data on its role in cancer, as it demonstrates a context dependent role as an oncogene or tumor suppressor[33]. It has been found that *PHLDA1* is overexpressed in highly metastastic cell lines and is correlated with patient poor prognosis in osteosarcoma[36]. Polyglutamine repeats, known to contribute to genomic instability, may impact the genetic function of *PHLDA1*, potentially driving tumorigenesis[34,37]. Alternatively, *PHLDA1* has been reported to have reduced expression in many cancers and has found to be associated with tumor-suppressive effects such as reduction of cell growth, colony formation, and migration along with

increased apoptosis[34,36,38–40,40–43]. Although initially unexplored, the mechanism by which PHLDA1 promotes tumor suppression was eventually linked to its regulation of AKT signaling.

PHLDA1 has been demonstrated to have a tumor suppressive function through inhibition of AKT signaling[38]. AKT signaling plays a central role in cellular growth and survival as well as migration and metabolism[44]. This pathway has been widely demonstrated to be dysregulated leading to cancer growth[44]. Growth factor binding to receptor tyrosine kinases on the cellular membrane activate phosphatidylinositol 3-kinase (PI3K)[44]. This activation causes production of phosphatidylinositol (3,4,5)-triphosphate on the cell membrane (PIP_3)[44]. PIP_3 production on the inner cellular membrane triggers the translocation and recruitment of AKT, which contains a PH domain, to the plasma membrane[44]. AKT is then activated at the cell membrane through phosphorylation at Thr308 and Ser407 residues. Activation of AKT leads to increased cell survival and proliferation, characteristic with cancer behavior[44].

Due to their PH domain, PHLDA family proteins are thought to competitively bind to membrane-bound phosphatidylinositol phosphates (PIPs), and deactivate oncogenic pathways such as AKT signaling[33,45,46]. *PHLDA1* has been determined to be a novel target for p53, a protein that regulates genomic stability[38]. It has been found that the PH domain of *PHLDA1* localizes the protein to the cell membrane facilitating its binding to PIPs[38]. This competitive binding represses AKT signaling, suggesting a role of *PHLDA1* in tumor suppression. In breast and ovarian cancers, genetic expression of *PHLDA1* confirms this role[38]. *PHLDA3* has also been studied in the context of AKT signaling.

PHLDA3 is a 13.9 kDa protein that is expressed in all human tissues but is elevated in fat tissue. Located on chromosome 1, *PHLDA3* has been demonstrated to have sequence similarity to an imprinted gene[47]. Additionally, it has been found that the genetic region of *PHLDA3* loses heterozygosity (LOH) at a high frequency in pancreatic neuroendocrine tumors (PanNETs) which lead to disease and poor prognosis[48].This study also demonstrates that both LOH as well as methylation are necessary for *PHLDA3* driven PanNET tumor development[48]. *PHLDA3* has been

determined to be a p53 target gene, similar to *PHLDA1* in pancreatic and lung cancers[44,48,49]. By competitively binding with PIP_3, AKT activation is inhibited, leading to tumor suppression[44,48,49]. However, it was also identified that in PanNETs *PHLDA3* frequently undergoes LOH and methylation at the gene locus, leading to suppression of *PHLDA3* transcription and downstream tumorigenesis[49]. These studies highlight the complexity of the context dependent roles of *PHLDA1* and *PHLDA3* in cancer.

PHLDA2 has been studied far less extensively than *PHLDA1* and *PHLDA3* in relation to its effect on AKT signaling. The ability of the PHLDA2 PH domain to bind to PIPs has been predicted through *in silico* analysis and confirmed using a novel single molecule pull down assay. to bind PIPs[45]. This binding capability has also been demonstrated in the context of trophoblast stem cells in mice[50]. While the paralogous genes *PHLDA1* and *PHLDA3* are known to reduce AKT signaling, underexplored whether *PHLDA2* directly affects this pathway. However, the mechanism of PIP binding and its chromosomal location suggest that *PHLDA2* may play a tumor-suppressive role in cancer.

The role of PHLDA2 in cancer remains highly controversial, with studies revealing both tumor-suppressing and tumor-promoting characteristics depending on the cancer type and context. PHLDA2 is located within a tumor suppressor region on the human chromosome, and its increased expression in human osteosarcoma has been associated with enhanced autophagy[51]. This protective function maintains cellular homeostasis by preventing the accumulation of damaged proteins and organelles which lead to DNA damage, therefore classifying it as tumor suppressing[51]. This protective role in osteosarcoma suggests that *PHLDA2* may act as a tumor suppressor by mitigating DNA damage and preventing the onset of genomic instability. In support of this, *in vivo* studies have shown that when human osteosarcoma cell lines overexpressing *PHLDA2* are injected into nude mice, these animals exhibit fewer metastatic occurrences compared to controls[51]. Thus, further implicating *PHLDA2* in suppressing cancer spread. In lung cancer models, PHLDA2 has been characterized as a mediator of ferroptosis, a form of regulated

cell death that has emerged as a key mechanism in tumor suppression[52]. These studies highlight *PHLDA2*'s potential role in limiting tumor growth and metastasis.

However, despite these tumor-suppressive functions, contrasting studies have indicated that PHLDA2 may also play a role in promoting cancer in certain contexts. In lung cancer, for instance, PHLDA2 expression has been shown to positively correlate with AKT activation, a critical signaling pathway involved in cell survival, growth, and metabolism[53]. The activation of AKT, particularly through phosphorylation (p-AKT), is essential for promoting cancer cell proliferation and survival, suggesting that in this context, *PHLDA2* may contribute to cancer progression rather than suppression[53]. Additionally, in hepatocellular carcinoma (HCC), a common and aggressive form of liver cancer, PHLDA2 expression is associated with poor patient survival[54]. In this context, *PHLDA2* is linked to several cancer-promoting processes, including angiogenesis, which facilitates tumor growth by enhancing blood vessel formation, and stemness, which is associated with the ability of cancer cells to self-renew and differentiate[54]. Furthermore, PHLDA2 has been implicated in treatment resistance in HCC, adding another layer of complexity to its role in cancer biology[54].

These findings underscore the multifaceted and context-dependent mechanisms by which PHLDA2 influences cancer progression. While it may act as a tumor suppressor in some cancers, it appears to have a dual role in others, contributing to cancer survival and treatment resistance. Importantly, the precise function of *PHLDA2* in breast cancer progression and metastasis remains largely unexplored. Given its involvement in various forms of cancer, further research into *PHLDA2*'s role in breast cancer could unveil its potential as a novel genetic driver of metastatic disease. This research would offer new insights into cancer biology and potential therapeutic targets for metastatic breast cancer. Identifying a new genetic marker of metastasis would be especially valuable because metastatic cancer cells are notoriously difficult to treat. This is particularly true for TNBC, a subtype of breast cancer known for its aggressive nature and lack of targeted therapies. The absence of specific molecular targets in TNBC makes it challenging to

treat effectively, and the identification of markers like *PHLDA2* could potentially lead to new therapeutic strategies or diagnostic tools that address this gap in treatment.

Chapter 1.6: Metabolic Drivers of Breast Cancer Metastasis

In addition to identifying genetic markers of metastasis to improve patient outcome, it is important to also understand the metabolic drivers of metastasis. It has been long understood that metabolic alteration occurs during tumorigenesis[55]. Due to their rapid growth, cancer cells sustain their functions through high metabolic demands[55]. A long understood metabolic hallmark of cancer cells through the Warburg Effect describes the use of glycolysis for energy production necessary for cellular survival[56,57]. In glycolysis, cells obtain energy from the uptake of glycose in the cytoplasm to produce two ATP molecules per glucose molecule digested[56,57]. This process is fast; however, it is energetically unfavorable[56,57]. In contrast, Oxidative Phosphorylation (OXPHOS), is energetically more favorable generating thirty to thirty-six ATP molecules[56,57]. However, OXPHOS needs oxygen to function, whereas glycolysis can occur under anaerobic conditions[56,57]. However, even with oxygen availability, the Warburg Effect described that cancer cells preferentially perform glycolysis to generate energy, despite being energetically unfavorable[56,57]. However, there has been significant work in the field challenging this dogma in the context of metastasis, suggesting the involvement of fatty acids, glutamine, proline and pyruvate carboxylase-mediated metabolism in metastasis[58–63]. Our prior work utilized a single cell transcriptomics approach to determine what metabolic process promotes metastasis, specifically during early metastatic seeding.

Our lab performed single cell sequencing on primary tumors and metastatic lungs and lymph nodes from immunocompromised NOD scid gamma (NSG) mice carrying patient derived primary tumors (PDX)[17]. We found in this work using gene ontology (GO) analysis of differentially expressed genes in primary tumors and metastatic lesions that primary tumors have high

expression of glycolytic markers[17]. This agrees with the Warburg Effect. However, we also found a clear switch in this metabolism in metastatic lesions[17]. In early metastatic cells, OXPHOS pathways were significantly expressed[17]. This work determines that breast cancer cells switch their metabolism from glycolysis to oxidative phosphorylation (OXPHOS) during metastasis[17]. To further determine that metastatic lesions were reliant on OXPHOS, we pharmacologically inhibited OXPHOS *in vivo*[17]. Through inhibition of complex V in the electron transport chain using oligomycin, we were able to selectively affect OXPHOS while shifting cells to glycolysis for ATP production.[17] We demonstrated a significant reduction in metastasis with complex V inhibition in two independent models[17]. Additionally, the inhibition of complex V showed no significant impact on primary tumor growth, indicating that OXPHOS is not essential for the progression of primary tumors[17]. This work provided evidence that OXPHOS is critical for metastasis[17]. However, it remains unclear what nutrient source allows for metastatic cancer cells to make this metabolic switch during the metastatic cascade.

Chapter 1.7: Experimental Limitations of Single Cell Metabolomics

Understanding what nutrient sources promote this invasive cancer cell behavior could lead to treatment of metastasis through cellular nutrient starvation mechanisms or dietary intervention. To answer this question, we have adapted and optimized a novel methodology for tissue digestion and flow cytometry sorting of cells from solid lung tissues that is compatible with liquid chromatography-mass spectrometry (LC-MS). Assessing the metabolic differences using LC-MS between primary tumor and metastatic cells has been greatly limited by sample preparation. Metabolomics relies on flash frozen tissue samples, which preserves metabolic signature while preventing recovery of minority metastatic cells from a tissue using fluorescence-activated cell sorting (FACs)[64]. When tissues are dissociated into single cell suspension for FACs, tissues are often digested under physiological conditions at 37°C. However, it has been demonstrated that physiological conditions alter the metabolomic profile of tissue[64]. To determine what nutrient

sources, drive this OXPHOS energy production, it was necessary to develop a method for digesting tissues at cold temperatures to isolate rare metastatic cells using flow cytometry.

To reduce RNA sequencing artifacts, a study employed the use of a cold activated protease to digest murine kidneys[65]. This protocol could be adapted to digest lung tissue into single cell suspension suitable for FACs while preserving their metabolic signature at cold temperatures. As metabolomics is typically performed on whole tissues, there lacks an ability to identify metabolic changes in subpopulations of organs, such as metastatic cells. Due to this limitation, there is little knowledge of metabolic heterogeneity in tissues. However, a protocol was developed to investigate the metabolic profile of rare circulating cancer cells using flow cytometry[66]. Using a cold activated protease and a protocol sufficient for obtaining metabolic profiles from rare cells using flow cytometry, we can use an unbiased approach to determine the nutrient source that drives OXPHOS metabolism in metastatic breast cancer cells.

Due to these limitations, there have been few studies conducted to determine the metabolic drivers of metastasis. In melanoma, it has been reported that metabolic differences drive metastatic potential due to alterations in lactate transport[67]. Using radiolabeled nutrients, a study determined that metastatic melanoma cells lines exhibit enhanced lactate uptake[67]. Through inhibition of MCT1, a lactate transporter, there was little effect on primary tumor growth[67]. However, with inhibition there was a significant depletion of circulating melanoma cells resulting in reduced metastatic disease[67]. In a study assessing breast cancer brain metastasis, mice injected with cancer cells in the breast or brain were fed radio-labeled glucose[68]. In comparing these two tissues, it was determined that fatty acid synthesis is elevated in breast tumors growing in the brain[68]. This differential metabolism is likely attributed to decreased lipid availability in the brain[68]. By inhibiting fatty acid synthesis genetically and pharmacologically, it was demonstrated that breast cancer tumor growth in the brain was reduced[68]. These studies demonstrate that nutrient sources are critical drivers of cancer metastasis and are additionally targetable as a mechanism for improving patient survival.

In our PDX model of breast cancer metastasis, we have previously demonstrated that metastatic cells rely heavily on OXPHOS for their energy needs. However, the specific nutrient source fueling this metabolic process remains unclear. Identifying this nutrient dependency is crucial, as it can reveal vulnerabilities in the metabolic network of metastatic cells. By employing a novel single-cell metabolomics approach, we aim to uncover the metabolic drivers of breast cancer lung metastasis. This knowledge could provide critical insights into the mechanisms of metastasis and pave the way for the development of targeted therapeutic strategies to improve treatment outcomes.

Chapter 2: PHLDA2 promotes breast cancer metastasis by increasing extracellular matrix deposition and vascular permeability

Chapter 2.1. Overview

In this chapter, we aimed to investigate whether *PHLDA2* actively promotes metastasis and modulates breast cancer cell behavior to enhance metastatic potential. By doing so, we sought to uncover its previously uncharacterized role in breast cancer progression and metastasis. *PHLDA2* is a maternally imprinted gene that has been demonstrated to mark the highly vascularized placental labyrinth and regulate placental development[26,28,30,69]. This gene acts as a rheostat, as deletion promotes placental overgrowth, but loss of imprinting impairs placental growth[26,28,30,69].

PHLDA2 is located within a tumor suppressor region of the chromosome, however, there is a lack of consensus of its role in cancer. *PHLDA2* expression has been correlated with enhanced autophagy and reduced metastatic burden in osteosarcoma[51]. In lung cancer models, *PHLDA2* has been defined as a ferroptosis mediator, which drives tumor suppression[52]. These studies support a role for *PHLDA2* in tumor suppression. Conversely, *PHLDA2* was found to promote AKT activation in lung cancer tissue, demonstrating tumor promoting function[53]. In a human hepatocellular carcinoma patient dataset, PHLDA2 was found to be indicative of poor prognosis and increased angiogenesis[54]. Together, these studies suggest that *PHLDA2* 's role in cancer progression is context-dependent, with conflicting evidence showing that it can both promote and suppress tumor growth depending on the biological and experimental conditions. However, the role of *PHLDA2* in breast cancer, particularly its functional impact on metastasis *in vivo*, remains unexplored. This study is novel in addressing this critical gap, providing insights into the molecular underpinnings of PHLDA2-mediated metastasis. Furthermore, these findings hold significant promise for advancing our understanding of metastatic TNBC which could improve patient survival through diagnostic tools and treatment strategies.

Our study revealed that *PHLDA2* promotes breast cancer metastasis, demonstrating its role in driving metastatic progression by modulating the extracellular matrix and enhancing vascular permeability. By genetically modifying *PHLDA2* expression levels in PDX and human cell line models, we observed a correlation between *PHLDA2* expression and metastatic burden. We have determined that *PHLDA2* promotes collagen deposition in the metastatic lungs, without altering the primary tumor. Additionally, we identified other ECM modulators, *MFAP5* and *SPARC*, whose expression is correlated with increased *PHLDA2* expression. Finally, we determined with an *in vitro* vascularized micro tumor system and *in vivo* staining of metastatic lungs, that *PHLDA2* promotes vascular permeability. Understanding the roles of pro-metastatic genes like *PHLDA2* in breast cancer metastasis can improve patient outcomes by enabling targeted treatment of metastatic cancer cells and identifying potential biomarkers for earlier detection and personalized therapeutic strategies.

Chapter 2.2. Significance

By identifying *PHLDA2* as a gene not previously linked to breast cancer spread, we provide new insights into the genetic drivers of breast cancer metastasis. Our data reveal that *PHLDA2* is more highly expressed in cancer tissues and suggest that its activation is influenced by methylation, consistent with its status as an imprinted gene. Furthermore, the association of elevated *PHLDA2* expression with poor relapse-free and distal metastasis-free survival in patient datasets establishes it as a prognostic marker for aggressive disease. Functionally, we demonstrate that *PHLDA2* contributes to metastasis by altering the tumor microenvironment through extracellular matrix restructuring and increased vessel permeability, which has been demonstrated to facilitate cancer cell invasion and colonization at distant sites. These findings not only uncover a novel role for *PHLDA2* in breast cancer biology but also allow for potential avenues for therapeutic targeting and biomarker development aimed at mitigating metastatic burden and improving patient outcomes.

Chapter 2.3. Innovation

The discovery of *PHLDA2* as a novel player in breast cancer metastasis marks a significant advancement in our understanding of the genetic mechanisms that drive this complex process. While many genes have been studied in the context of breast cancer metastasis, *PHLDA2* had not previously been implicated, making its identification particularly noteworthy. Our research reveals that elevated expression of *PHLDA2* in metastatic tissues is associated with poor clinical outcomes, suggesting it may serve as a valuable prognostic marker.

To further investigate the functional role of PHLDA2 in metastatic behavior, we employed a vessel-on-a-chip system to evaluate vascular permeability. This innovative approach allows for the precise recreation of the tumor microenvironment, enabling us to model and quantify how *PHLDA2* influences endothelial cell permeability *in vivo*. By using this advanced technology, we provide critical insights into the mechanisms by which PHLDA2 alters the vascular landscape, facilitating the invasive capabilities of cancer cells. This combination of novel gene discovery and cutting-edge experimental methodology highlights the potential for *PHLDA2* to be a key target for future therapeutic interventions aimed at mitigating metastatic spread.

Chapter 2.4. Materials and Methods

2.4.1. Patient Survival Analysis

To assess relapse-free and distal-metastasis free survival, we used the KM plotter database (KMplot.com) to generate Kaplan-Meier survival curves using a primary tumor microarray dataset (mRNA Breast Cancer Gene Chip) of breast cancer patients. All Kaplan-Meier plots use patient split ‘auto select best cutoff’ parameter. Subtypes were determined using PAM50 subtypes.

2.4.2. TCGA Gene Expression Analysis

To assess gene expression in primary tumor versus normal tissue, the UALCAN analysis portal was used to mine the TCGA Breast Invasive Carcinoma dataset. Statistics were provided by the portal.

2.4.3 Analysis of epigenetic dysregulation of PHLDA2 in breast cancer

Bulk methylation and gene expression profiles for the *PHLDA2* gene were obtained from the GDC TCGA Breast Cancer repository. Exploratory data analysis and data download was performed with the UCSC Xena platform[70]. In-depth analysis was performed on the singularity container server "Bioportal1" with a JupyterLab instance created by the UC Irvine Research Cyberinfrastructure Center. An R notebook with the following packages was used: R (v4.1.2), ggplot2 (v3.3.5), dplyr (v.1.0.8), reshape (1.4.4). Primary tumors and normal samples which contained missing data were filtered out using "na.omit". For single-modality plotting, only data for that modality was considered in the filtration. For joint methylation and gene expression analysis, samples with missing data in either modality were excluded.

To select the candidate regulatory sites for PHLDA2, the methylation beta values of all 32 associated loci from Infinium HumanMethylation450 BeadChip were visualized as box and whisker plots. Welch's t-test from UCSC-Xena as well as visual examination of the differences in the means in the tumor and normal samples for each locus identified two loci as most differentially demethylated in breast cancer (cg05167973 and cg04720330). These loci were replotted with their adjusted p-values after a Bonferroni multiple testing adjustment and retained for further analysis.

To test whether the two demethylated loci have a role in upregulating *PHLDA2* expression in breast cancer, the methylation beta values of each locus was plotted against the RNA

expression for PHLDA2 on a scatter plot. The negative sign of the slope for the best-fit line as well as Spearman's rank correlation rho were used to infer the inverse relationship between methylation levels at the candidate regulatory loci and the levels of gene expression. T-tests and Bonferroni p-value adjustments were performed again.

2.4.4. PDX Sample

A patient derived xenograft (PDX) sample was provided by A. L. Welm at the Department of Oncological Sciences at the Huntsman Cancer Institute (HCI). Tissues were collected from individuals being treated at the Huntsman Cancer Hospital and University of Utah with informed consent under a protocol approved by the Institutional Review Board of the University of Utah. A pleural effusion of a female patient diagnosed with Stage IIIC ER−PR−Her2− basal-like (PAM50) invasive ductal carcinoma and treated with several rounds of chemotherapies is noted as HCI-010. This sample was collected and de-identified by the Huntsman Cancer Institute Tissue Resource and Application Core facility before being obtained for implantation. The study is compliant with all relevant ethical regulations regarding research involving human participants.

2.4.5. Cell Culture

Breast cancer cell line MDA-MB-231, was cultured in DMEM, 1X (Corning, 10-013-CV) media respectively with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich, cat. no. 12306 C) and 1% penicillin/streptomycin (Hyclone, SC30010) 100 X solution. The cells were passaged with 0.05% trypsin (Corning, 25-052-Cl) after reaching 70% confluency. MDA-MB-231 cells are mycoplasma-free and STR profiled with a 93% match to the ATCC reference.

2.4.6. Viral Transduction (PDX and Cell Lines)

PDX cells were transduced as described in our previous Nature Communications biology publication.[1] GFP control (+GFP) and human PHLDA2-GFP (+PHLDA2-GFP) lentiviral

expression vectors were packaged into lentiviral particles purchased from VectorBuilder Inc., Chicago, IL, USA (+GFP Cat. no. (VB190812-1255tza), +PHLDA2-GFP Cat. no. (VB180802-1084xrz)). MDA-MB-231 cell line was transduced with Scramble GFP control (+Scramble-GFP) and human PHLDA2 shRNA-GFP (+shPHLDA2-GFP) lentiviral expression vectors which were packaged into lentiviral particles and purchased from VectorBuilder Inc., Chicago, IL, USA (+Scramble-GFP Cat. no. (VB010000-0009mxc), +shPHLDA2-GFP Cat. no. (VB220307-1326nwx)). All cell lines were transduced with 10 mg/mL polybrene. Cells were incubated in virus overnight before switching to fresh media to allow cells to grow out. GFP positive transduced cells were selected using FACs sorting. Alterations in PHLDA2 expression were confirmed by qPCR and Western Blot Analysis.

2.4.7. Animal Experiments

For spontaneous metastasis generation, HCI010 (120,000) cells were mixed 1:1 with sterile PBS and Matrigel (Corning 356230) and injected orthotopically into the number four mammary fat pad area of female NSG mice using a 0.5 mL insulin syringe. For experimental metastasis generation, cultured MDA-MB-231 (500,000 cells) cells were resuspended in 100 uL of sterile PBS and injected intravenously into the tail vein of female NSG mice. Primary tumors were measured using a caliper and volumes were calculated using the tumor volume = 1/2 (length x width^2) equation. All animal experiments were reviewed and approved by The Institutional Animal Care and Use Committee of the University of California, Irvine.

2.4.8. Tissue Harvest and Dissociation

At the endpoint, animals were euthanized by asphyxiation with CO_2 followed by cervical dislocation and perfusion with 10 mM EDTA in PBS. Tumors and lungs were removed and dissociated for flow cytometry by mechanical chopping with razor blades. The chopped tissues were digested in 2 ug/mL Collagenase IV (Sigma-Aldrich cat. No. C5138-1G) in DMEM with 300 mg/mL DNAseI (Catalog Number) for 20 minutes at 37°C. The cell suspensions were washed

with Hanks Balanced Salt Solution (HBSS) and passed through a 70 μm cell strainer. Lung and primary tumor cells were treated with 1X Red Blood Cell Lysis buffer followed by resuspension in PBS for FACs analysis.

2.4.9. Flow Cytometry and Florescence Activated Cell Sorting (FACs)

Cells were stained with a fluorescently conjugated antibodies for CD298/MHCI to isolate human cells. Human specific-CD298 (diluted 1:100; PE; BioLegend, cat. no. 341704) and the mouse-specific antibody MHC-I (diluted 1:150; APC; Thermo Fisher Scientific, cat. no. 17-5957-80) antibodies were added to samples. Cell viability was determined using SYTOX blue stain (diluted 1:1000, Thermo Fisher Scientific, cat. No. S34857). Cell viability was determined using SYTOX blue stain (diluted 1:1000, Thermo Fisher Scientific, cat. No. S34857). The samples were resuspended in PBS for FACs analysis using the BD FACSAria Fusion cell sorter (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). To isolate single cells from doublet and multiplet cells, forward scatter area by forward scatter width (FSC-A x FSC-H) and side scatter area by side scatter width (SSC-A x SSC-H) were used. GFP labeled human primary tumor and metastatic lung cells were selected by gating on $Sytox^{-}GFP+$ cells.

2.4.10. qPCR

RNA was extracted from cells using Quick-RNA MicroPrep (Zymo Research, R1050). cDNA was obtained from iScript cDNA Synthesis Kit (Bio Rad, 1708891). q-PCR amplification was performed using Power SYBR Green PCR Master Mix (Applied Biosystems, A25742). Gene specific primers were utilized to amplify *PHLDA2* (Forward Primer: CCGCCGCGGGCCATA, Reverse Primer: CCACAGCCGGATGGTAGAAA), *MFAP5* (Forward Primer: AGTCAACGAGGAGACGATGTG, Reverse Primer: CATCCCAGCACTCCAAGTCA) and *SPARC* (Forward Primer: TGAGAATGAGAAGCGCCTGG, Reverse Primer: TGGGAGAGGTACCCGTCAAT). Each sample was standardized to the housekeeping gene *GAPDH* (Forward Primer: CTCTCTGCTCCTCCTGTTCGACGAC, Reverse Primer:

TGAGCGATGTGGCTCGGCT). Primers were generated by Integrated DNA Technologies. Relative quantification of transcripts for all samples was performed as triplicates and analyzed by the CT values for each gene group.

2.4.11. Western Blot

Cells were lysed in RIPA buffer (150mM NaCl, 50mM Tris-HCl pH 8.0, 1% Triton-X, 0.5% sodium deoxycholate, 0.1% SDS, diluted in H_2O) with protease and phosphatase inhibitors (Thermo Fisher, Cat No. 78430). Lysates were incubated with agitation for 45 minutes and then centrifuged at 12,000g for 10 minutes. Supernatant was removed and protein was quantified using Pierce BCA Protein Assay Kit (Thermo Fisher, Cat No. 23225) according to manufacturer's instructions. 8-20ug protein samples were run on a 12% SDS PAGE gel (BioRad, Cat No. 4568046) and proteins were wet transferred to a PDVF membrane. After blocking with 5% BSA, primary antibody was incubated at 4°C overnight. Primary antibodies included: 1:500 rabbit Anti-TSCCE (PHLD2) antibody (Abcam, Cat no. ab234669), 1:1000 SPARC polyclonal antibody (Thermo Fisher, Cat no. PA5-78178), and 1:1000 mouse anti-mouse β-actin (Santa Cruz, sc-47778). Secondary antibodies were diluted in blocking solution and placed on membranes for one hour at room temperature. Secondary antibodies included: 1:1000 Goat Anti-Mouse HRP (Thermo Fisher, Cat No. 31430) or 1:1000 Goat Anti-Rabbit HRP (Thermo Fisher, Cat No. 31460). Protein bands were visualized using Thermo Fisher Chemiluminescent Substrate kit (Thermo Fisher, Cat No. 34579) and imaged BioRad ChemiDoc system. Densitometry was performed using ImageJ and normalized to β-actin control.

2.4.12. Tissue Preparation

Tumor and lung tissues for Mason’s trichrome staining were fixed in 10% neutral-buffered formaldehyde at room temperature for 24 hours, placed in freshly prepared 70% ethanol, and process for paraffin embedding in a Leica tissue processor using standard protocols. FFPE

blocks were sectioned into 5 μm sections using a Reichert-Jung 2040 Autocut Rotory Microtome and stained with Mason's trichrome protocol.

Tumor and lung issues for immunofluorescence staining were fixed in formaldehyde for 6-12 hours at 4℃ and placed in 30% sucrose overnight. Tissues were embedded with O.C.T Compound (Tissue-Tek, Cat no. 4583) in disposable base mold at – 80℃. Tissue sections of 8 μm were sectioned using a Leica Cryostat CM1950 and stored in a closed container at -80℃.

2.4.12. Immunofluorescence

Slides with OCT-embedded tissue sections were washed, blocked with BlockAid Blocking Solution (Life Technologies Corporation, 810710) at room temperature for an hour, and incubated overnight with primary antibodies (1: 100 Ki67, Genetex, Cat No. GTX16667, 1:100 MFAP5 Rabbit Polyclonal Antibody, Proteintech, Cat no. 15727-1-AP, CD31 Rat Anti-mouse, 1:75, BD Bioscience, Cat no. 741740) at 4℃. Albumin Rabbit Polyclonal Antibody (1:100, Proteintech, Cat no. 16475-1-AP) was used in the same method but blocking occurred in 5% Fish Gelatin. Tissues were incubated with appropriate secondary antibodies for 1 hour in the dark at room temperature. Tissues were mounted with Antifade Mounting Medium with DAPI (Epredia, 8312-4) and sealed with a coverslip. Slides were imaged on Keyence BZ-X series inverted microscope.

2.4.13. Immunofluorescence Image Quantification

Five-ten 20x regions of interest were captured per tissue stained. Each image was quantified and all images corresponding to a sample were averaged to report one value per mouse on graphs. The following analyses were performed using NIH open-source image software ImageJ (**http://rsbweb.nih.gov/ij/**).

$Ki67^+$ GFP^+ area

$Ki67^+$ GFP^+ double positive area was quantified by using thresholding to measure the area of GFP^+. GFP^+ area was converted to a mask and projected onto the same image's $Ki67^+$ channel. $Ki67^+$ area across the image was visualized using a constant threshold. Next, $Ki67^+$ area within the GFP^+ mask was measured. Lastly, $Ki67^+$ GFP^+ area was normalized to total GFP^+ area and multiplied by 100 to measure the percent of $Ki67^+$ area within GFP^+ metastase or the $Ki67^+$ GFP^+ area. We selected this method specifically to quantify Ki67-positive tumor cells, while avoiding the inclusion of other proliferating cell types in the metastatic lung.

MFAP5 expression and Albumin leakage

MFAP5 and albumin staining were quantified as the ratio of $MFAP^+$ or $albumin^+$ area normalized to $DAPI^+$ area and multiplied by 100 to quantify percent of MFAP or albumin area relative to DAPI area. We determined MFAP5, albumin, or DAPI area using consistent threshold values across samples. We chose this method of quantification to determine the expression of MFAP5 or leakage of albumin relative to total lung tissue area.

Vessel Structure Analysis

Tumor vessel structural analysis was performed as described previously[71]. Briefly, vessel parameters quantified included the total number of $CD31^+$ vessels, the average vessel length of all vessels (reported in μm), the number of vessels with a length greater than 50 μm (elongated vessels), and the number of 100 μm^2 regions per image occupied by $CD31^+$ vessels (microvessel density). The number of open lumens was also quantified but none were observed in images and therefore is not reported.

2.4.14. Cell Growth Assays (Cell Titer-Glo 3D Cell Viability Assay, MTT Cell Proliferation Assay)

After 7 days of culturing PDX primary tumor organoid cells, 200,000, 500,000 and 100,000 cells were seeded into rows of a 24 well ultra-low attachment plate. Cell Titer-Glo 3D Cell Viability

Assay (Promega, Cat no. G9681) protocol was followed as directed by the manufacturer and imaged using the luminescence function on Biotek Cytation 5 plate reader. MDA-MB-231 cells were seeded at 5,000 cells/well in a standard flat bottom 96 well plate. MTT Cell Proliferation Assay Kit (Cayman Chemical, Cat no. 10009365) protocol was followed as directed by manufacturer and imaged using absorbance function on Biotek Cytation 5 plate reader.

2.4.15. Bulk RNA-Sequencing Library Prep

RNA was extracted from sorted primary tumor cells (Viable GFP^+) using Quick-RNA MicroPrep (Zymo Research, R1050) as directed by the manufacturer. Total RNA was monitored for quality control using the Agilent Bioanalyzer Nano RNA chip and Nanodrop absorbance ratios for 260/280nm and 260/230nm. Library construction was performed according to the Illumina TruSeq mRNA stranded protocol. The input quantity for total RNA within the recommended range and mRNA was enriched using oligo dT magnetic beads. The enriched mRNA was chemically fragmented. First strand synthesis used random primers and reverse transcriptase to make cDNA. After second strand synthesis the ds cDNA was cleaned using AMPure XP beads and the cDNA was end repaired and then the 3' ends were adenylated. Illumina barcoded adapters were ligated on the ends and the adapter ligated fragments were enriched by nine cycles of PCR. The resulting libraries were validated by qPCR and sized by Agilent Bioanalyzer DNA high sensitivity chip. The concentrations for the libraries were normalized and then multiplexed together. The multiplexed libraries were sequenced using paired end 100 cycles chemistry for the Illumina NovaSeq 6000. 50 million reads were sequenced per sample.

2.4.16. Bulk Sequencing Analysis

Bulk RNA sequencing data was processed following a standardized workflow. Initial quality control of raw sequencing reads was conducted using FastQC, followed by alignment to the Homo sapiens reference genome (hg38) using HISAT2. Gene-level read counts were generated using featureCounts, counting only exonic regions. Differential gene expression analysis was

performed with DESeq2, identifying differentially expressed genes (DEGs) with an FDR-corrected adjusted p-value < 0.05 and $\log_2$ fold change ≥ 0.5 for the PHLDA2 overexpression samples compared to GFP control samples. Low-count genes were filtered out before calculating reads per kilobase per million mapped reads (RPKM). For further analysis, only protein-coding genes with an average reads per kilobase per million mapped reads (RPKM) above 2 were included in volcano plots. Volcano plot was plotted with R package EnhancedVolcano (v1.20.0)[72]. The gene expression heatmap displays normalized gene expression levels in transcripts per million (TPM). Heatmap was made using pheatmap (v1.0.12)[73]. Gene set enrichment analysis (GSEA) was performed using R packages clusterProfiler (v4.10.1) and org.Hs.eg.db (v3.18.0) using the Gene Ontology Biological Processes gene sets[74,75]. Top ten significant (adjusted p value > 0.05) pathways of interest were selected. Normalized enrichment score bar plots were generated with ggplot2 (v3.5.1). Enrichment plots were generated with enrichplot (v1.22.0).

2.4.17. Mason's Trichrome Staining

Tissues were deparaffinized with Histo-Clear and decreasing concentrations of ethanol (100%, 95%, 70%) for one minute each. Each step was done twice. Masson's trichrome staining was performed using Trichrome Stain Kit (Abcam, Cat no. ab150686) according to manufacturer instructions. Briefly, Bouin's fluid was added onto tissues and incubated for one hour, followed by a wash step. Equal parts of Weigert's reagents A and B were mixed to create Weigert's Iron Hematoxylin and added onto tissues for five minutes and then washed. Tissues were then incubated with a Biebrich Scarlet/Acid Fuchsin solution for 15 minutes followed by a wash step. After, tissues were incubated in Phosphomolybdic/Phosphotungstic Acid solution for 15 minutes. Without washing Aniline Blue was immediately added to the slides and incubated for 10 minutes. Slides were then washed and incubated in a 1% Acetic Acid solution for 5 minutes. Tissues were then dipped in 95% ethanol and 100% ethanol and left in Histo-Clear (National

Diagnostics, HS-200) for one minute. Tissues were mounted with a mounting medium (Cytoseal XYL, Epredia, 8212-4) and coverslip. 10x objective images were acquired on the Keyence BZ-X series inverted microscope. Open access image software, QuPath (downloaded from https://qupath.github.io/) was used for analysis. In QuPath, a boarder was drawn around a metastatic lesion to create a region of interest (ROI) and area of the ROI was measured. Blue pixels (collagen) were then isolated from other colors using the QuPath's color deconvolution stains and the area of collagen within the metastatic lesion ROI was measured. Lastly, the area of collagen within a metastatic lesion was normalized to total area of metastatic lesion and multiplied by 100 to quantify the percent of collagen deposition in a metastatic lesion. This process was repeated for every lesion identified in 5-10 randomly acquired 10x objective brightfield images. Each lesion was plotted as a point on the graph. In total, 3 lungs were imaged per group.

2.4.18. Microfluidic Device Fabrication

Microfluidic device fabrication and loading has been previously described[76–81]. In summary, a custom polyurethane master mold is created using a 2-part polyurethane liquid plastic (Smooth Cast 310, Smooth-On Inc.). Subsequently, a PDMS layer is replicated from this master mold, and holes are punched to create inlets and outlets. The platform is assembled in two stages: first, the PDMS layer is chemically glued and subjected to 2 minutes of oxygen plasma treatment to affix it to the bottom of a 96-well plate. Following this, a 150 μm thin transparent membrane is bonded to the bottom of the PDMS device layer through an additional 2-minute treatment with oxygen plasma. The fully assembled platform is then placed in a 60°C oven overnight, covered with a standard 96-well plate polystyrene lid, and sterilized using UV light for 30 minutes before cell loading.

2.4.19. Cell Culture and Microfluidic Device Loading

To establish the vascularized microtumor (VMT), normal human lung fibroblasts and ECFC-ECs (endothelial colony forming cell endothelial cells) were harvested and resuspended in fibrinogen solution at a concentration of 3×10^6 cells/mL and 7×10^6 cells/mL, respectively. MDA-MB-231 cells (GFP, PHLDA2, Scramble or shPHLDA2) were introduced into the mixture at a concentration of 1×10^5 cells/mL fibrinogen solution. Fibrinogen solution was prepared by dissolving 70% clottable bovine fibrinogen (Sigma-Aldrich) in EBM2 basal media (Lonza) to a final concentration of 5 mg/mL. The cell-matrix suspension (6 µL) was mixed with thrombin (50 U/mL, Sigma-Aldrich) at a concentration of 3 U/mL, quickly seeded into each microtissue chamber, and allowed to polymerize in a 37◦C incubator for 15 minutes. Laminin (1 mg/mL, LifeTechnologies) was then introduced into the microfluidic channels through medium inlets and incubated at 37°C for an additional 15 minutes. After incubation, culture medium (EGM-2, Lonza) was introduced into the microfluidic channels and medium wells. Medium was changed every other day and hydrostatic pressure head re-established daily to maintain interstitial flow.

2.4.20. VMT fluorescence imaging and analyses

Fluorescence images were acquired with a Biotek Lionheart fluorescent inverted microscope using automated acquisition and standard 10x air objective. To test vessel perfusion, 25 µg/mL rhodamine-conjugated 70 kDa dextran was added to the medium inlet for 10 minutes. ImageJ software (National Institutes of Health) was utilized to determine the total extravascular fluorescence intensity (mean grey value) for each VMT. Subtraction of background fluorescence measurements was performed for each chamber.

2.4.21. Statistics and Reproducibility

Data are presented as the mean ± s.d. from at least three independent experiments, unless stated otherwise. Statistics were conducted using Student's *t* test or Mann Whitney. Statistical test results are represented on graphs according to the following: ns, $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$. No statistical method was used to determine

sample size and age-matched female NSG mice were randomly assigned into experiments. Investigators were aware which conditions applied to each experimental group while performing experiments and analyzing data.

Chapter 2.5. Results

2.5.1. PHLDA2 expression is associated with increased metastasis and poor survival in breast cancer patients

PHLDA2 expression in breast cancer patient tumors and its association with metastasis remain unclear. We investigated *PHLDA2* expression in healthy breast (n=114) and invasive breast carcinoma (n=1097) from the Cancer Genome Atlas (TCGA) dataset [82,83]. This revealed that *PHLDA2* expression is 3-16-fold higher in patient breast tumors relative to healthy breast cancer subtypes **(Figure 2-1a)**. We further find that *PHLDA2* shows the highest expression in the HER2+ breast cancer subtype (**Figure 2-2a**), which is consistent with prior reports that *PHLDA2* is regulated by HER2 signaling[53]. We further determined whether PHLDA2 expression in patient primary breast tumors is associated with relapse free survival (RFS, n=4929) or distant metastasis free survival (DMFS, n=2765) using the KMplotter gene chip database containing 1,809 patients [84,85]. Remarkably, high *PHLDA2* expression is associated with poor RFS (HR=1.42) and DFMS (HR=1.52) in nearly all types of patients and breast cancer subtypes (**Figure 2-1b,c**) (**Figure 2-2b,c**). This suggests that increased PHLDA2 expression is in primary tumors is a robust predictor of both metastatic propensity, as well as poor survival across breast cancer patient populations.

2.5.2. PHLDA2 hypomethylation is associated with increased RNA expression and metastasis

PHLDA2 is a maternally imprinted gene that shows strong gene dosage effects on placental growth following hypomethylation and expression of the paternal allele[26,28,69]. Knowing imprinted genes are regulated through methylation, we investigated whether *PHLDA2* overexpression in breast cancer may also occur through demethylation. *PHLDA2* is a 1,148 kb gene on chromosome 11 that is comprised of two exons (one untranslated) one intron, and 5' upstream regulatory sequences that include the promoter region. There are 32 CpG loci known to be associated with *PHLDA2*, including six exonic, two intronic, and 24 in the 5' regulatory

sequence (**Figure 2-1d**). We compared mean methylation levels for each locus in patient breast cancer samples compared to healthy breast tissues from the TCGA dataset. We found that most loci exhibit low beta values (mean: 0.05), suggesting they remain unmethylated in both healthy breast and tumor tissues (**Figure 2-2d**). However, two loci located in the untranslated exon region (cg04720330 and cg051679730) show higher beta values and appear methylated in healthy breast tissue (**Figure 2-1e, Figure 2-2d**). Importantly, we found an inverse correlation between *PHLDA2* RNA expression and methylation at these two specific loci (**Figure 2-1f**), suggesting that methylation at these sites may preferentially regulate *PHLDA2* expression. *PHLDA2* expression has also been reported to be regulated by hypomethylation in clear cell renal cell carcinoma[86]. These loci also display significantly lower methylation in primary and metastatic tumor samples compared to healthy tissue (**Figure 2-1e**). These findings suggest that the increased *PHLDA2* expression observed in primary breast and metastatic tumor samples may occur through hypomethylation, particularly at two loci located in the untranslated exon of the gene.

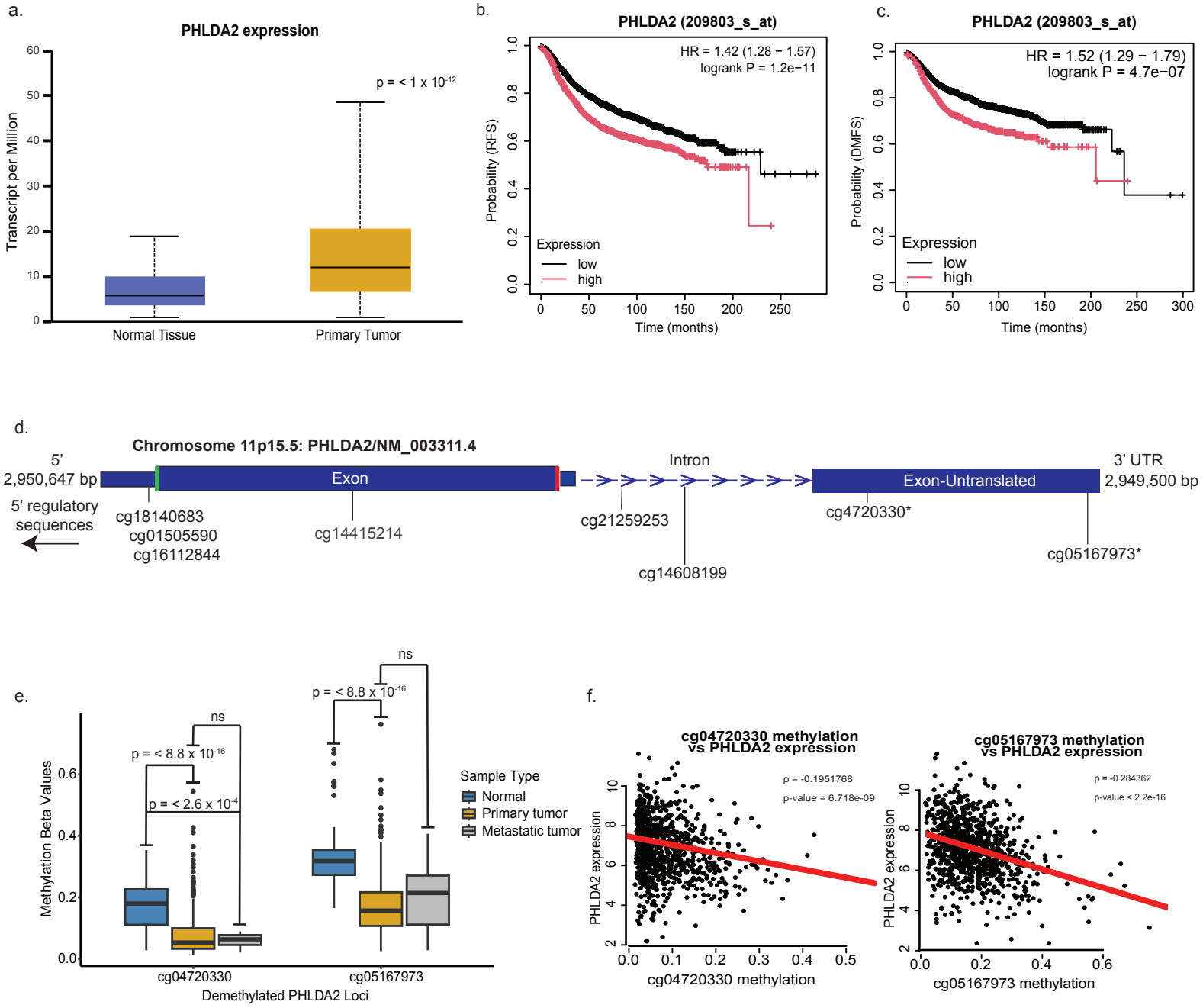


Figure 2-1: Elevated PHLDA2 expression in paitent breast tumors is associated with increased metastasis, poor survival and hypomethylation at specific CpG loci.

a.PHLDA2 expression in healthy breast (n=114) and primary breast tumor (n = 1097) tissues. Bar plot shows RNA expression in tissues from the TCGA dataset plotted using the UALCAN tool (ualcan.path.uab.edu). *P* value was calculated using a student's t-test. b. Kaplan-Meier plot shows probability of RFS in breast cancer patients with high (n=2208) vs. low (n=2721) expression of *PHLDA2* in their primary tumor tissue. Plot was generated using the KM plotter compiled dataset of mRNA breast cancer gene chip and visualized using the KM plotter tool (KMplot.com). c. Kaplan-Meier plot shows probability of DMFS in breast cancer patients with high (n=717) vs. low (n=2048) expression of *PHLDA2* in their primary tumor tissue. Plot was generated using the KM plotter compiled dataset (mRNA breast cancer genechip) and visualized using the KM plotter tool (KMplot.com). d. Schematic shows CpG loci in the exonic and intronic regions of the *PHLDA2* gene. Loci located in upstream 5' regulatory sequences are not shown. Loci were annotated using the GRCh37/hg19 human genome and the Infinium HumanMethylation450 assay. *Loci hypomethylated in primary and metastatic tumor tissues relative to healthy breast tissues. e. Bar plot shows methylation values at select CpG loci (cg4720330, cg05167973) in healthy breast (n = 139), primary (n = 1101), and metastatic (n = 5) tumor tissues from breast cancer patients from the TCGA dataset. Select loci were identified in a screen for loci that display hypomethylation in tumor tissues and correlation with increased *PHLDA2* expression. *P* values were calculated using student's t test with Bonferroni multiple-testing adjustment. f. Correlation plots show relationship between *PHLDA2* RNA expression and methylation at select CpG loci (cg4720330, cg05167973). *P* value was calculated using a students t-test with Bonferroni multiple testing adjustment. The rho value was calculated using Spearman's rank correlation.

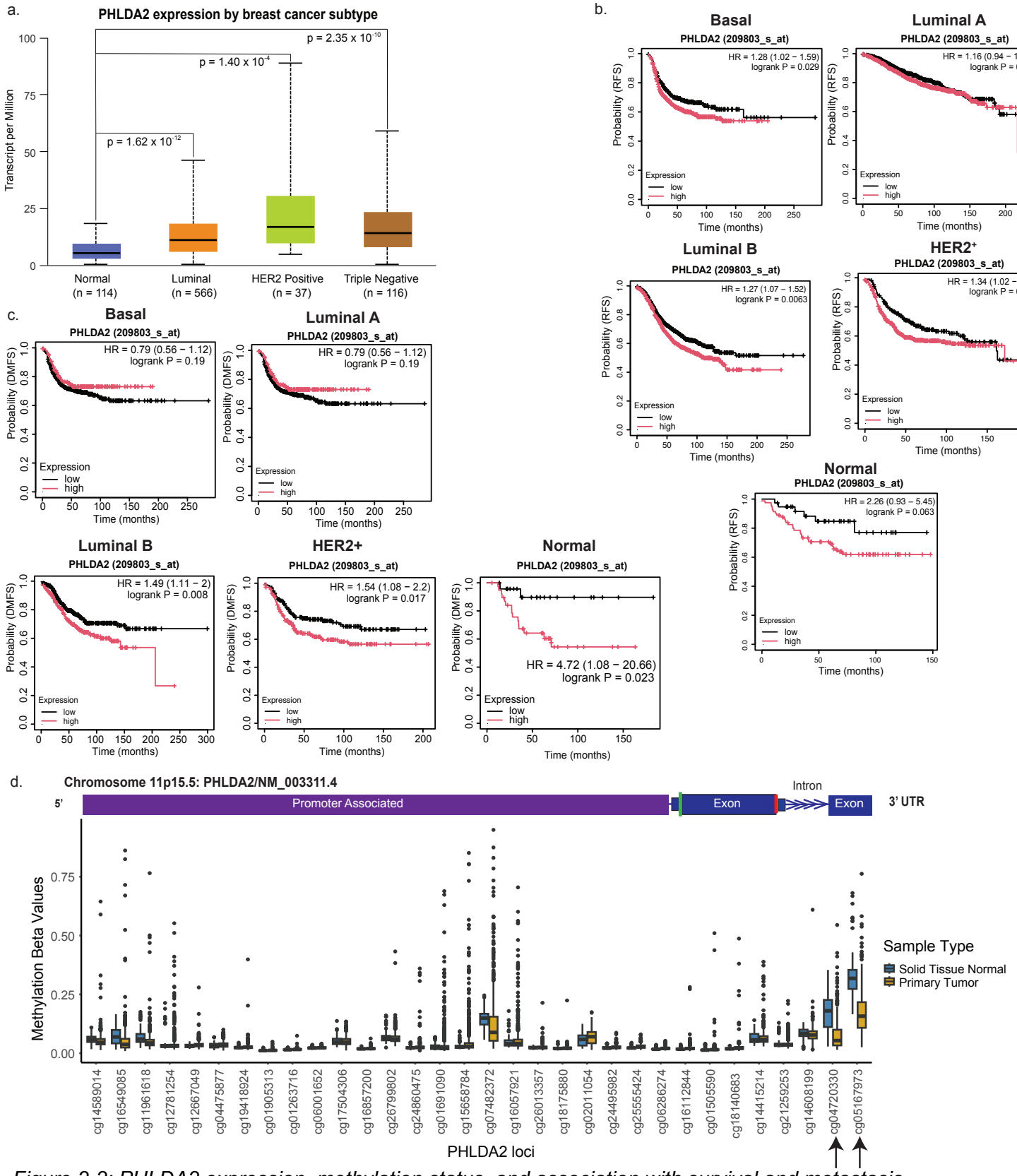


Figure 2-2: PHLDA2 expression, methylation status, and association with survival and metastasis in breast cancer patients.

a.*PHLDA2* expression in healthy breast (n=114) and primary breast tumors representing different breast cancer subtypes Luminal, HER2 positive and Triple Negative. Bar plot shows RNA expression in tissues from the TCGA dataset plotted using the UALCAN tool (ualcan.path.uab.edu). *P* value was calculated using a student's t-test. b. Kaplan-Meier plots show probability of RFS in breast cancer patients with high vs. low expression of *PHLDA2* in their primary tumor tissue, separated by breast cancer subtype. Plots were generated using the KM plotter compiled dataset of mRNA breast cancer genechip and visualized using the KM plotter tool (KMplot.com). Basal (n=392 low, 561 high), Luminal A (n=852 low, 957 high), Luminal B (n=710 low, 643 high), HER2+ (n=206 low, 489 high), Normal (n=38 low, 81 high).c. Kaplan-Meier plots show probability of DMFS in breast cancer patients with high vs. low expression of *PHLDA2* in their primary tumor tissue, separated by breast cancer subtype. Plots were generated using the KM plotter compiled dataset of mRNA breast cancer genechip and visualized using the KM plotter tool (KMplot.com). Basal (n=469 low, 161 high), Luminal A (n=579 low, 419 high), Luminal B (n=443 low, 230 high), HER2+ (n=197 low, 204 high), Normal (n=24 low, 31 high).d. Methylation values for all *PHLDA2* loci in solid normal tissue (n = 139) and tumor tissue (n = 1101). Significantly methylated loci between primary tumor and normal tissue are marked with arrows. Diagram of *PHLDA2* gene was generated with UCSC Genome browser and denotes location of loci in the chromosome using the hg37 reference genome.

2.5.3. PHLDA2 promotes breast cancer lung metastasis in vivo

We determined whether *PHLDA2* functionally promotes metastasis by assessing the effects of genetic modulation of *PHLDA2* expression on breast cancer metastasis *in vivo*. We utilized a human patient-derived xenograft (PDX) model of triple negative breast cancer (HCI010) that spontaneously metastasizes to the lung following orthotopic transplantation[16]. We previously found that cells from this model display increased *PHLDA2* expression (2-3-fold) during metastatic seeding in the lung[17]. HCI010 cells were propagated using a 3D tumor sphere culture system, followed by lentiviral infection using constructs to overexpress (*PHLDA2* GFP) or knockdown (*PHLDA2* shRNA GFP) *PHLDA2* [87]. *PHLDA2* overexpression did not result in significant changes *in vitro*, indicating that the cells are viable and show similar growth kinetics after genetic modulation. (**Figure 2-4e**). The HCI010 cells were subsequently injected orthotopically into NOD SCID gamma (NSG) mice and tissues were harvested at endpoint (~two months) (**Figure 2-3a**). Western blot analysis confirmed overexpression and knockdown of PHLDA2 in experimental relative to control tumors (**Figure 2-4 f,g**). Genetic modulation of PHLDA2 expression resulted in limited changes in primary tumor growth (**Figure 2-3b, Figure 2-4b**), demonstrating that primary tumor size should not have significant confounding effects on metastatic burden.

We determined the effects of changes in PHLDA2 expression on lung metastasis by whole mount fluorescence microscopy and flow cytometry for GFP (**Figure 2-3c,d**). Remarkably, we found that PHLDA2 overexpression leads to a dramatic increase in GFP+ metastatic foci (**Figure 2-3c**), and 6-10 times more metastatic cells in the lung than control (**Figure 2-3d**). Immunofluorescence staining for Ki67 in metastatic lesions revealed limited differences in cellular proliferation between PHLDA2 overexpressing and control conditions (**Figure 2-4h**). These results suggest that PHLDA2 drives metastasis through enhancing seeding efficiency, since we observe a dramatic increase in metastatic foci but limited changes in proliferation.

PHLDA2 knockdown resulted in a reduced but non-significant change in lung metastasis (**Figure 2-4c,d**). This may be due to low baseline levels of *PHLDA2* expression (**Figure 2-3e**) and lung metastasis in the HCI010 model, making it difficult to measure a statistically significant reduction in *PHLDA2* expression or metastasis. We therefore performed analogous experiments using MDA-MB-231 breast cancer cells, which have ~2-fold higher levels of endogenous *PHLDA2* expression than HCI010 cells (**Figure 2-3e,f**). Western blot analysis confirmed efficient knockdown of PHLDA2 following lentiviral infection in vitro (**Figure 2-4i**). PHLDA2 knockdown did not result in significant changes in MDA-MB-231 cell growth as measured by MTT assay *in vitro* (**Figure 2-4j**). MDA-MB-231 cells were injected via tail vein, and tissues were harvested 21 days later (**Figure 2-3f**). Analysis of lung tissues showed that PHLDA2 knockdown results in a significant reduction in metastatic burden in the lung (**Figure 2-3g,h**). Taken together, these findings show that PHLDA2 is not only associated with metastasis in patients, but also functionally promotes breast cancer lung metastasis *in vivo*.

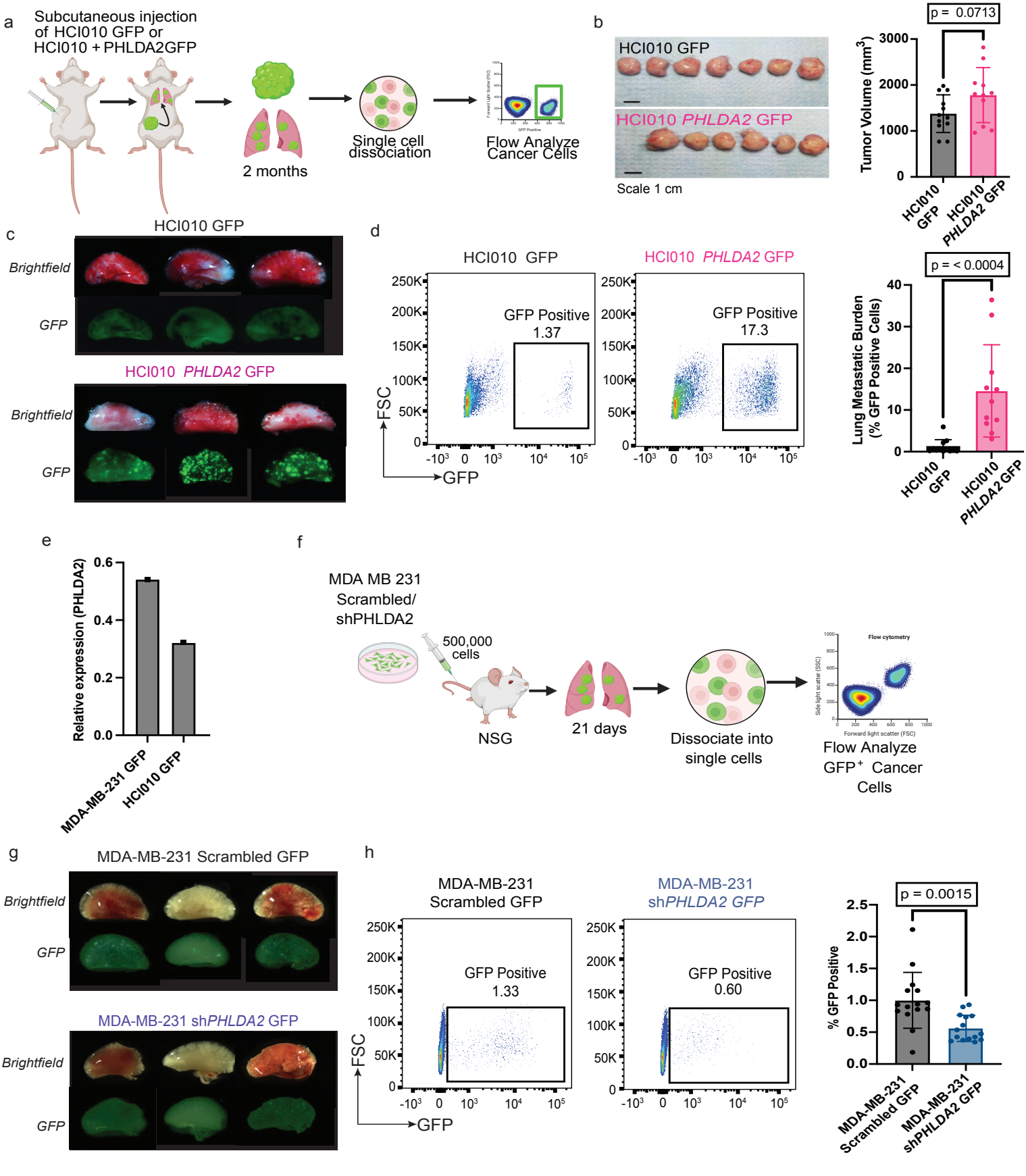


Figure 2-3: Overexpression of PHLDA2 promotes breast cancer lung metastasis in vivo.

a. Experimental schematic of spontaneous metastasis model for HCI010 PDX tumor cells expressing GFP (control) or PHLDA2 GFP in NSG mice. b. Images show primary tumors from mice injected orthotopically in the mammary fat pad region with HCI010 PDX tumor cells expressing GFP (control) or *PHLDA2* GFP (n= 12 GFP/11 *PHLDA2* GFP) Scale bar = 1 cm. Bar plot shows tumor volume (cm^3). *P* value was calculated using a student's t-test. c. Representative images show whole mount fluorescence microscopy of metastatic lungs from mice injected orthotopically with HCI010 PDX tumor cells expressing GFP (control) or *PHLDA2* GFP. Images are shown with brightfield and GFP filters. d. Representative flow cytometry plots of metastatic lungs from mice injected orthotopically in the mammary fat pad region with HCI010 PDX tumor cells expressing GFP (control) or *PHLDA2* GFP. Bar plot shows quantification of lung metastatic burden using cancer cell GFP marker (n=16/group). *P* value was calculated using a student's t-test. e. qPCR was performed to quantify mRNA expression of PHLDA2 in metastatic breast cancer models, MDA-MB-231 GFP and HCI010 GFP. Gene expression was normalized to housekeeping gene control, GAPDH. f. Experimental schematic of experimental metastasis model for IV delivery of MDA-MB-231 tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP in NSG mice. g. Images show whole mount fluorescence microscopy of metastatic lungs from mice injected with MDA-MB-231 tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP. Images are shown with brightfield and GFP filters. h. Representative flow cytometry plots of metastatic lungs from mice intravenously injected with MDA-MB-231 cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP. Bar plot shows quantification of lung metastatic burden using cancer cell GFP marker (n=15/group). *P* value was calculated using a student's t-test.

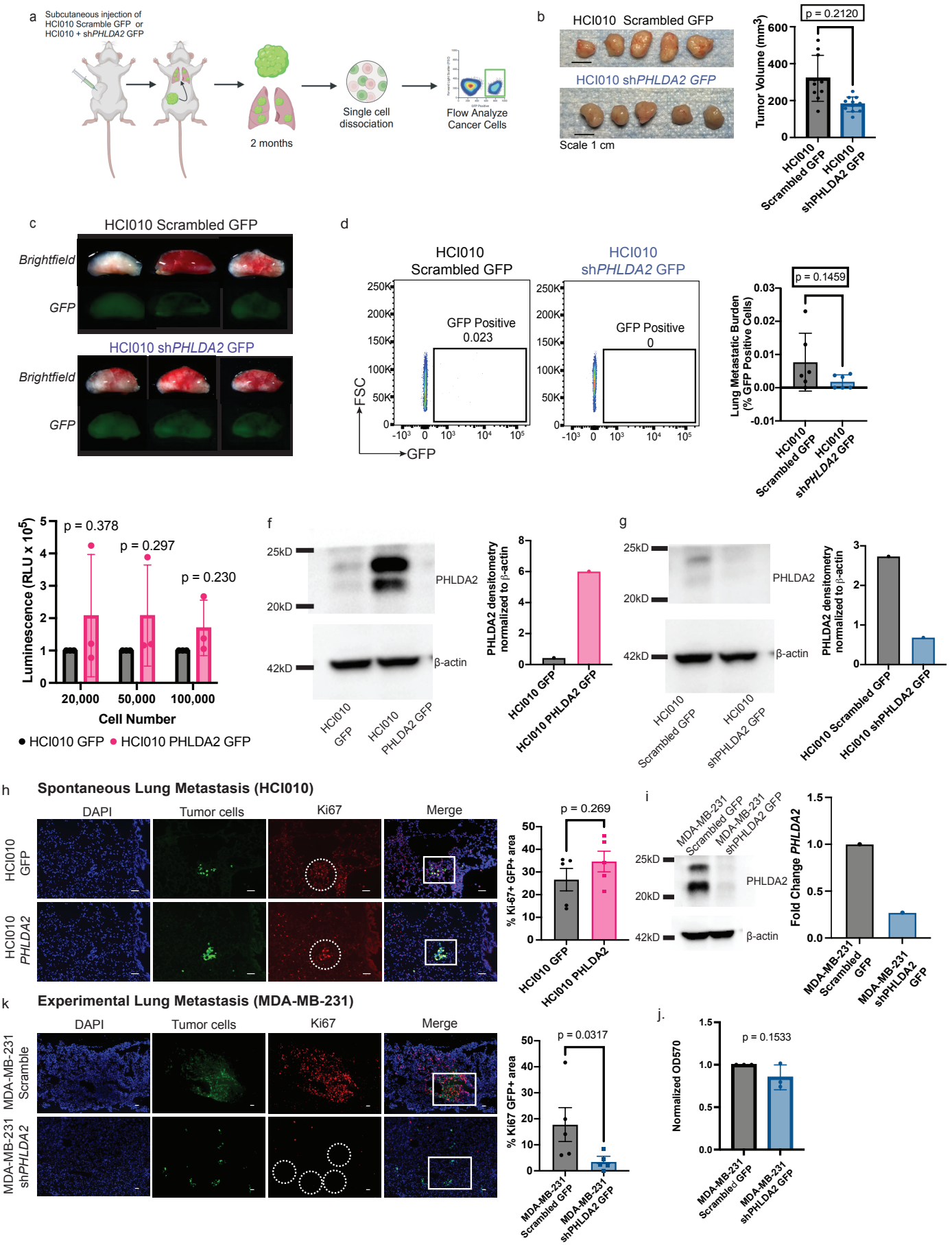


Figure 2-4: Modulation of PHLDA2 expression has limited effects on cancer cell proliferation

a. Experimental schematic of spontaneous metastasis model for HCI010 PDX tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP in NSG mice. b. Images show primary tumors from mice injected orthotopically in the mammary fat pad region with HCI010 PDX tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP (n = 11/group). Scale bar = 1 cm. Bar plot shows tumor volume (cm^3). *P* value was calculated using a student's t-test. c. Representative images show whole mount fluorescence microscopy of metastatic lungs from mice injected subcutaneously in the mammary fat pad region with HCI010 PDX tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP. Images are shown with brightfield and GFP filters. d. Representative flow cytometry plots of metastatic lungs from mice injected orthotopically in the mammary fat pad region with HCI010 PDX tumor cells expressing Scrambled GFP (control) or sh*PHLDA2* GFP. Bar plot shows quantification of lung metastatic burden using cancer cell GFP marker (n=10/group). *P* value was calculated using a student's t-test. e. Cell viability represented by Cell Titer Glo demonstrates in vitro proliferation of HCI010 GFP (control) or *PHLDA2* GFP organoids. Bar plots represent average luminescence values of organoids at three seeding densities (20,000, 50,000, 100,000 cells). Each point represents an average of 4 technical replicates. *P* values were calculated using multiple student's t-tests. f. Western blot image for PHLDA2 protein expression in HCI010 GFP (control) and PHLDA2 GFP sorted tumor cells. Bar plot represents densitometry of PHLDA2 normalized to densitometry of β-actin control. Image densitometry was performed using ImageJ. g. Western blot image for PHLDA2 protein expression in HCI010 Scrambled GFP (control) and shPHLDA2 GFP sorted tumor cells. Bar plot represents densitometry of PHLDA2 normalized to densitometry of β-actin control. Image densitometry was performed using ImageJ. h. Representative images of immunofluorescence staining for DAPI (blue), tumor cells (green) and Ki67 (red) of HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Scale bar = 50 μm. Quantification was performed for $Ki67^+$ GFP^+ area. See methods for detailed description of quantification. Graph is displayed as mean ± SD where each point represents one lung value obtained by the average of 5-10 20x microscopic fields. *P* value was calculated using a student's t-test. i. Western blot image for PHLDA2 protein expression in MDA-MB-231 Scrambled GFP (control) and shPHLDA2 GFP primary tumors. Bar plot represents densitometry of PHLDA2 normalized to densitometry of β-actin control. Image densitometry was performed using ImageJ. j. Cell viability represented by MTT demonstrated in vitro proliferation of MDA-MB-231 Scrambled GFP (control) or sh*PHLDA2* GFP cells. Bar plot represents Normalized OD_{570} which is related to cell proliferation. Each point represents an average of 12 technical replicates. *P* value was calculated using a student's t-test. k. Representative images of immunofluorescence staining for DAPI (blue), tumor cells (green) and Ki67 (red) of metastatic lung tissue from mice intravenously injected with MDA-MB 231 Scrambled GFP (control) or shPHLDA2 GFP. Scale bar = 50 μm. Quantification was performed for $Ki67^+$ GFP^+ area. See methods for detailed description of quantification. Graph is displayed as mean ± SD where each point represents one lung value obtained by the average of 5-10 20x microscopic fields. *P* value was calculated using a student's t-test.

2.5.4. Identification of cellular programs associated with PHLDA2 overexpression

We investigated potential mechanisms for how *PHLDA2* drives metastasis using RNA sequencing to identify cellular programs that change in response to increased *PHLDA2* expression. Tumors were generated from HCI010 control (n=4) and *PHLDA2*-overexpressing (n=4) cells and dissociated to single cell suspension (**Figure 2-5a**). Human cells were isolated by flow cytometry based on GFP and CD298 expression (**Figure 2-5a**). Libraries were generated using the Illumina TruSeq mRNA stranded protocol and sequenced on the NovaSeq 6000 using paired end 100-base pair read sequencing at 50 million reads per sample. Sequencing reads were aligned to the human reference genome (hg38), and differential gene expression analysis was performed using DESeq2. This revealed 530 significantly upregulated and 288 downregulated genes (adjusted p value > 0.05; log2 fold change ≤ 0.5) (**Figure 2-5b, Figure 2-6, Table 1**) in PHLDA2-overexpressing tumor cells relative to controls. Gene set enrichment analysis (GSEA) using the Gene Ontology Biological Processes (GO BP) database revealed several recurrent cellular programs upregulated in PHLDA2-overexpressing tumors (**Figure 2-5c, Table 1**). These included extracellular matrix (ECM) organization, cell-substrate adhesion, integrin binding, regulation of vascular development, and regulation of cell migration (**Table 1).** Top-ranked genes associated with the ECM included numerous collagens (COL1A1, COL1A2, COL2A1), other ECM genes (VIT, MFAP5, LUM, FBLN1), matrix metalloproteinases (MMP2, MMP11, MMP14), and regulators of ECM restructuring (SPARC) (**Table 2**), suggesting PHLDA2 may promote metastasis through modulating genes that alter the ECM. Top genes associated with cell-substrate adhesion and integrin binding include adhesion molecules such as CEACAM1 and BCAM, metalloproteases like MMP14, and integrins such as ITGA6 and ITGA3, among others (**Table 2**). Genes associated with vascular development included C3, CHI3L1, SPHK1, as well as SPARC, which has also been shown to increase vascular permeability, extravasation, and lung metastasis in melanoma models (**Table 2**)[88–93]. Of note, we found that these top-ranked gene

sets contain many overlapping genes, suggesting that PHLDA2 may regulate a ‘meta-program’ that regulates ECM remodeling, cellular adhesion, and vascular remodeling, and may have pleiotropic effects on promoting metastasis.

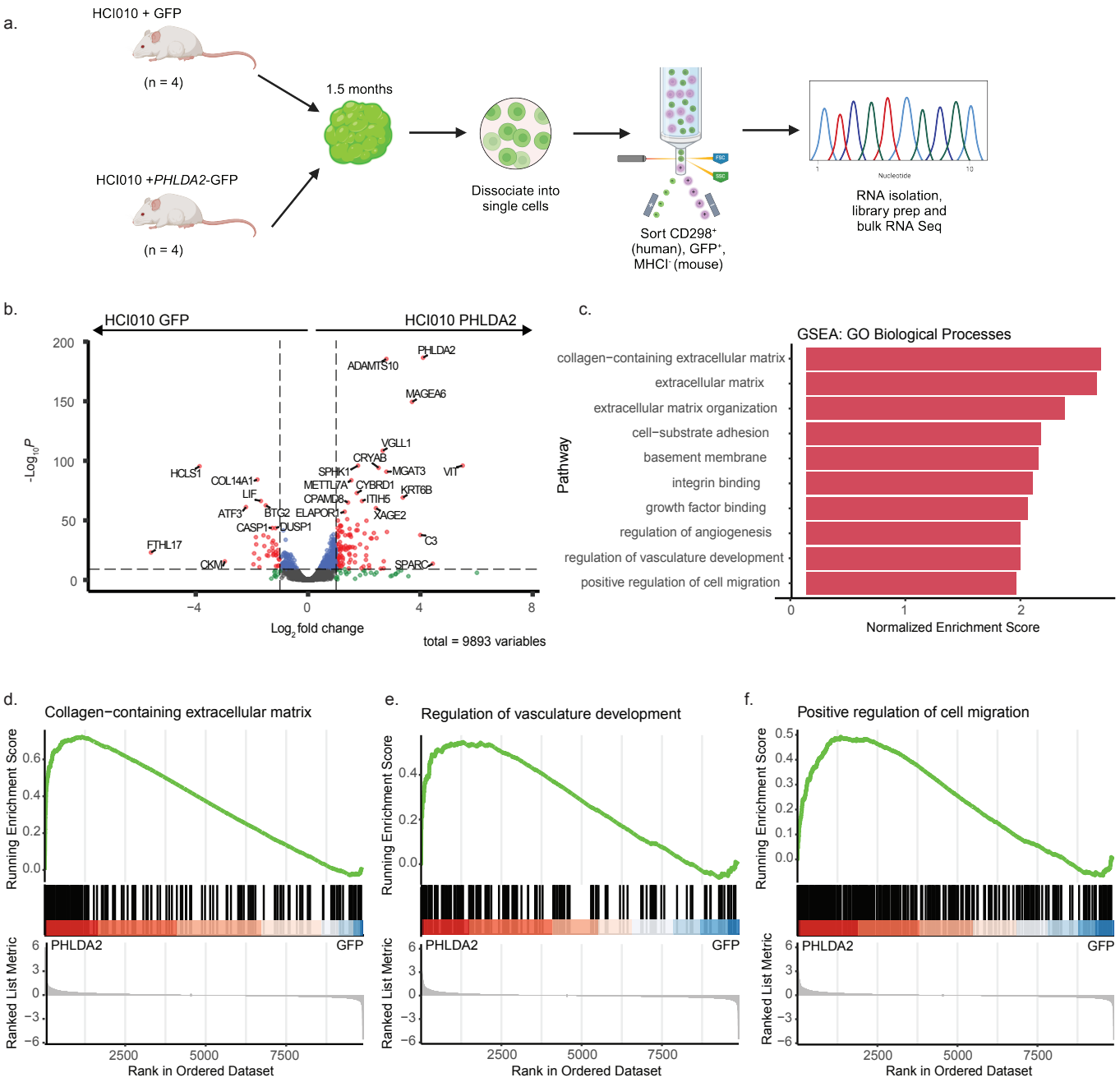


Figure 2-5: Bulk RNA sequencing reveals cellular programs associated with PHLDA2 overexpression

a. Experimental schematic illustrating how bulk sequencing of HCI010 GFP/PHLDA2 primary tumors was performed. HCI010 GFP and HCI010 PHLDA2 primary tumors were collected and tumor cells were sorted using GFP and human-specific CD298 expression. Mouse cells were excluded during sort using a mouse specific MHC-I antibody. RNA was then isolated from tumor cells followed by library preparation and RNA sequencing. b. Volcano plot depicts differentially expressed genes in HCI010 PHLDA2 overexpression versus GFP tumor cells. Plot was generated using the R package EnhancedVolcano. c. Chart depicting top ten gene set enrichment analysis (GSEA) pathways in HCI010 PHLDA2 overexpressing tumor cells. GSEA was performed using R package ClusterProfiler and plot was generated with ggplot2. d. GSEA enrichment plot of collagen-containing extracellular matrix gene term (GO:0062023). Plot was generated using the R package enrichplot. e. GSEA enrichment plot of regulation of vasculature development gene term (GO:1901342). Plot was generated using the R package enrichplot. f. GSEA enrichment plot of positive regulation of cell migration (GO:0030335). Plot was generated using the R package enrichplot.

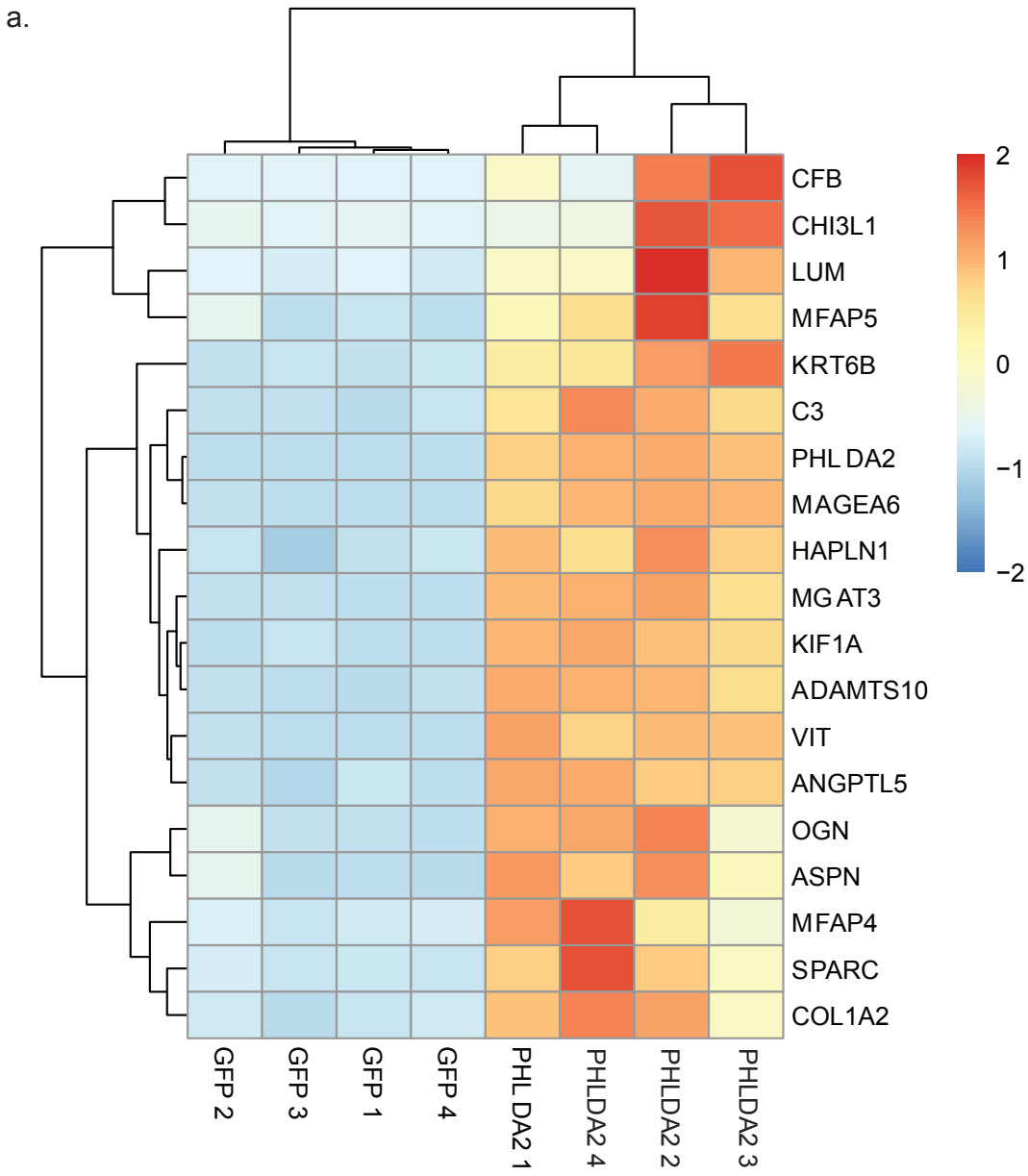


Figure 2-6: Top differentially expressed genes show a conserved upregulation across PHLDA2 overexpressing sample replicates

a. Heatmap of bulk sequencing gene expression results across the four HCI010 GFP samples and four HCI010 PHLDA2 overexpression samples. Top 20 significantly differentially expressed genes are visualized.

2.5.5. PHLDA2 promotes ECM remodeling

We further investigated whether PHLDA2 overexpression leads to changes in the ECM *in vivo*. ECM remodeling can promote metastasis in many ways, such as through increased tissue stiffness, alternation of cellular signaling, or providing a scaffold for migration (PMID: 33037194). We first evaluated the expression of several key ECM-associated genes identified by RNA sequencing. qPCR analysis showed increased expression of *SPARC*, *COL1A1, MFAP5* and *LUM* in PHLDA2-overexpressing primary tumors relative to controls (**Figure 2-7a**). Western blot analysis further confirmed increased *SPARC* expression in tumors at the protein level (**Figure 2-7b)**. We also performed *in situ* analyses of protein expression in metastatic tissues. MFAP5 is glycoprotein component of microfibrils in the extracellular matrix that has been associated with breast cancer cell proliferation and migration[94]. Immunofluorescence staining for MFAP5 showed significantly increased expression in metastatic lesions in the lungs of animals transplanted with PHLDA2-overexpressing tumors relative to controls (**Figure 2-7c,d, Figure 2-8a).** We also observed numerous collagens upregulated in PHLDA2-overexpressing tumors by RNA sequencing (**Table 2**). Increased collagen density can increase tumor tissue stiffness and promote metastasis[8,95,96]. Collagen fibers can be quantified in tissues using Masson's Trichrome staining, where aniline blue binds to basic amino acid residues on collagens turning them blue[97]. Masson's Trichrome staining showed 3-7-fold increase in collagen fiber density in metastatic lesions in the lungs of animals transplanted with PHLDA2-overexpressing tumors relative to controls (**Figure 2-7e,f, Figure 2-8b**). Together, these data demonstrate a clear role for PHLDA2 in promoting ECM remodeling, which can have pleiotropic effects on promoting metastasis.

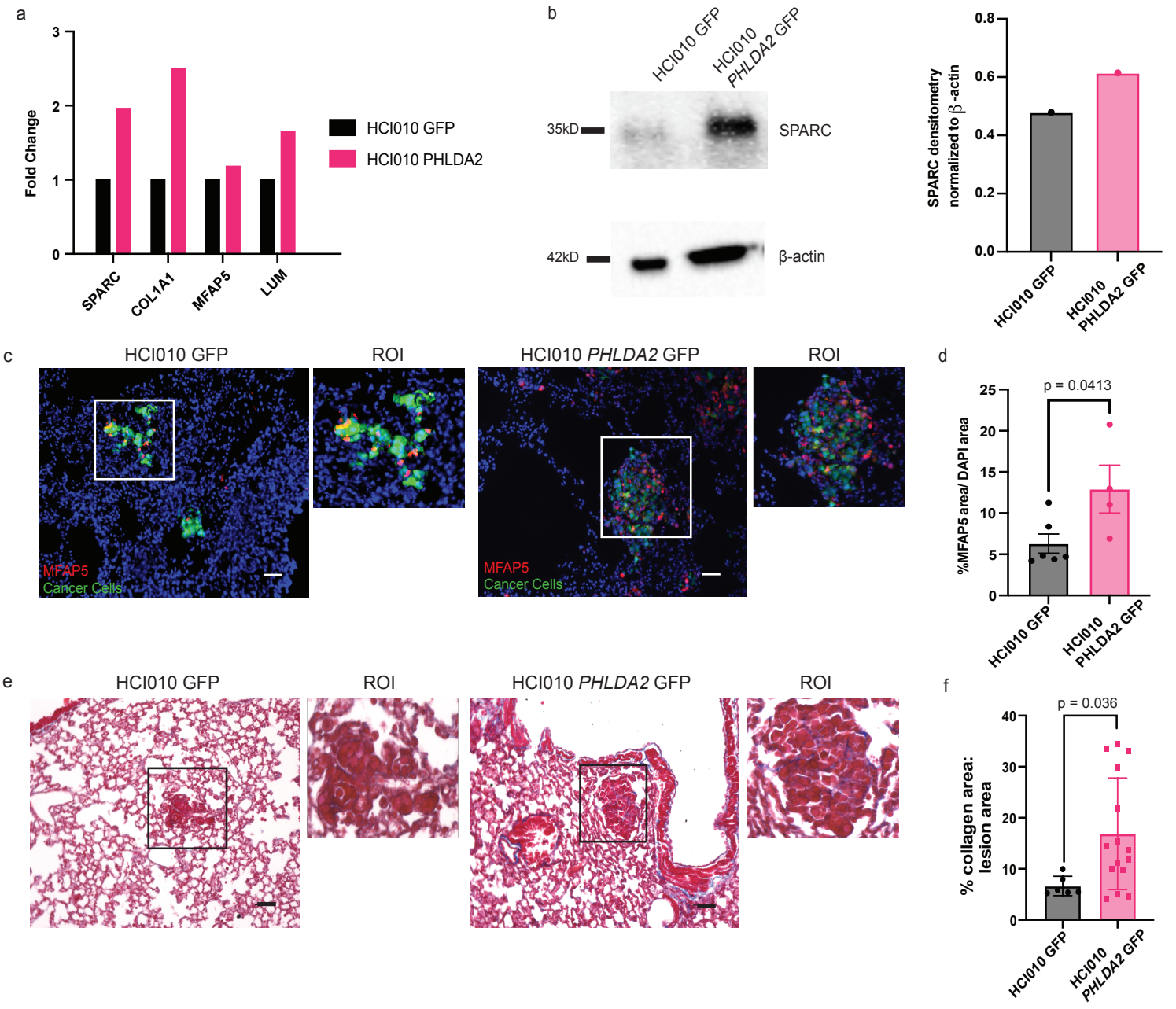


Figure 2-7: PHLDA2 promotes ECM remodeling in metastatic lung lesions

a. qPCR was performed to quantify mRNA expression of ECM related genes in HCI010 primary tumors. Gene expression was normalized to housekeeping gene control, GAPDH, and displayed on graph as fold change relative to HCI010 GFP. b. Western blot for SPARC protein expression in HCI010 GFP and HCI010 PHLDA2 sorted tumor cells. Densitometry quantification for SPARC and β-actin was performed using ImageJ. SPARC densitometry was normalized to β-actin control and graphed. c. Representative images of immunofluorescence staining for DAPI (blue), tumor cells (green), MFAP5 (red) in HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Scale bar = 50 μm. d. Quantification of the percent of $MFAP5^+$ area normalized to $DAPI^+$ area (n= 4-6/group). Graph is displayed as mean ± SD; each point represents one lung tissue value obtained by the average of 5-10 20x objective microscopic fields. Student's t test results are represented on the graph. e. Representative brightfield images of collagen deposition (blue) in HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Collagen was stained for using Masson's Trichrome stain. Scale bar = 100 μm. f. Quantification of percent of $collagen^+$ area normalized to lesion area in HCI010 GFP or HCI010 PHLDA2 metastatic lung tissue (n=6 for HCI010 GFP and n=16 for HCI010 PHLDA2). See methods for detailed description of quantification. Graphs are displayed as mean ± SD; each point represents the value of one metastatic lesion. Student's t test result is represented on the graph.

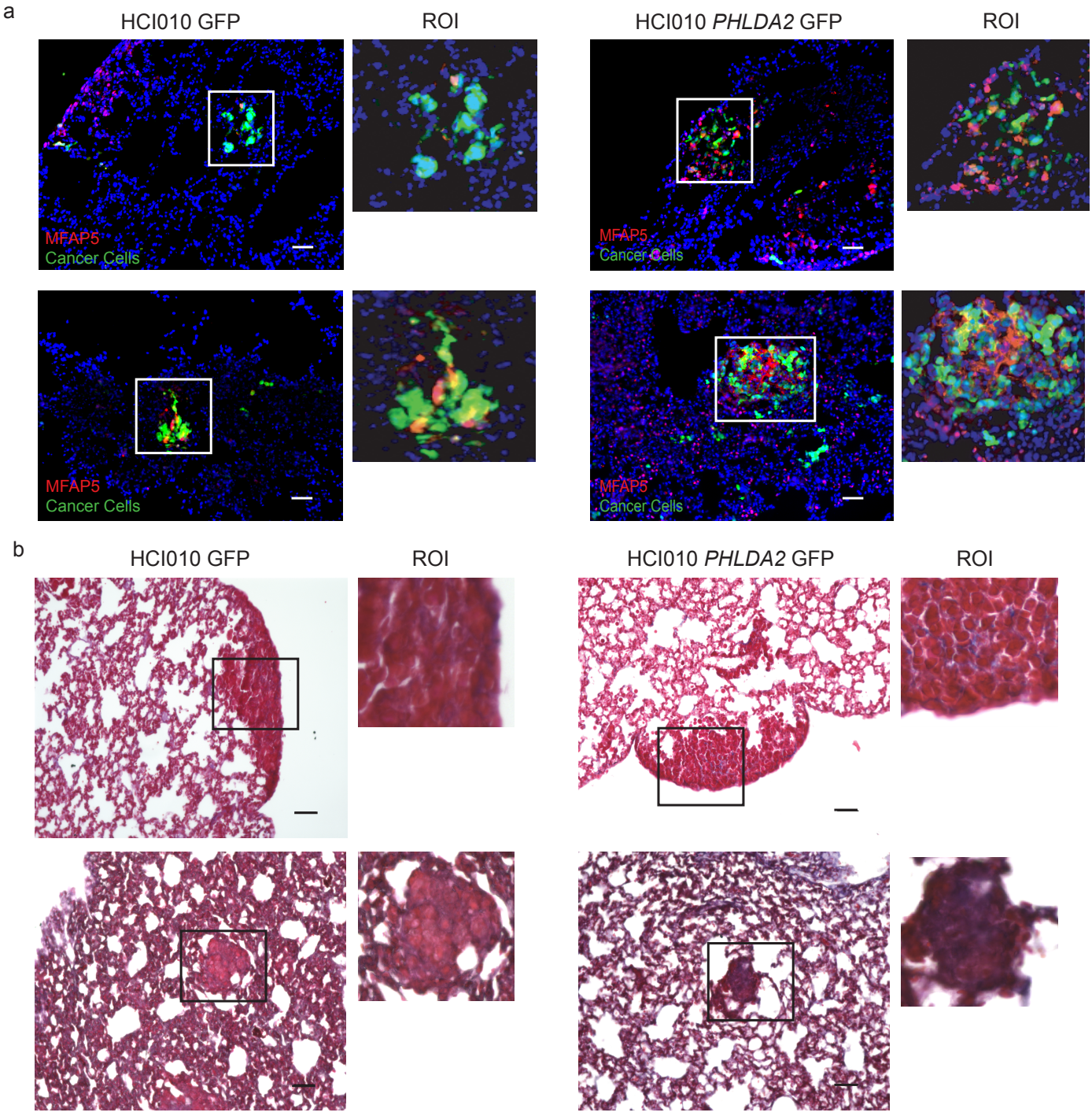


Figure 2-8: PHLDA2 increases MFAP5 expression and collagen deposition in metastatic lung lesions

a. Additional images of immunofluorescence staining of DAPI (blue), tumor cells (green), and MFAP5 (red) in HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Scale bar = 50μm. b. Additional brightfield images of collagen deposition (blue) in HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Collagen was stained for using Masson's Trichrome stain. Scale bar = 100μm.

2.5.6. PHLDA2 promotes vascular permeability

Previous work has demonstrated that increases in tumor density correlated with increased risk of metastasis across cancer types vessel size[98]. We further investigated whether PHLDA2 overexpression leads to changes in tumor vessel morphology vasculature *in vivo*. We performed immunofluorescence staining for CD31 to compare vessels in tumors grown from PHLDA2-overexpressing versus control HCI010 cells. Surprisingly, we found no changes in vessel number, average vessel length, number of elongated vessels, or micro vessel density (**Figure 2-9a,b**).

We next assessed changes in vessel functionality. Interestingly, several genes relating to the 'vascular development' gene set have been reported to specifically promote vessel permeability, including SPHK1, SPARC, C3 and CEACAM1[88,90,91,93,93,99]. The kinase SPHK1 catalyzes the production of Sphingosine 1 Phosphate (S1P) which binds to S1PR2 on endothelial cells, increasing vascular permeability[90]. SPARC is a secreted factor that binds to VCAM1 on activated endothelial cells and disrupts endothelial cell barrier function[88]. C3, a complement component, is part of the C3a/C3a receptor (C3aR) axis that upregulates VCAM1 on endothelial cells, which SPARC binds to for increased vessel permability[88,92]. Finally, CEACAM1 is an adhesion molecule that can anchor cells to blood vessels through homo or heterophilic binding[91,92,100–103]. This molecule facilitates their extravasation and likely alters vascular permeability due to the role of CEACAM family member in regulating barrier function[91]. Bulk sequencing demonstrated these factors were all upregulated in PHLDA2 overexpressing tumors indicating, collectively they may weaken endothelial barrier function and promote metastasis. To investigate whether *PHLDA2* promotes vessel permeability *in vitro*, we first utilized *an in vitro* vascularized micro-tumor (VMT) assay (**Figure 2-10a,b**)[76–81]. The VMT model was created by embedding lung fibroblasts, endothelial cells, and cancer cells (MDA-MB-231 GFP, *PHLDA2*-GFP, Scramble-GFP or sh*PHLDA2*-GFP) in a fibrinogen matrix and polymerizing the mixture in microtissue chambers. Microvessel permeability was analyzed using fluorescence imaging to

quantify leakage of 70kD dextran which was perfused through the microvessels on day 4. We found that PHLDA2 overexpression results in two-fold increase in vessel permeability *in vitro*, while PHLDA2 knockdown significantly reduces vessel permeability (**Figure 2-10c,d**).

We further evaluated the effects of PHLDA2 modulation on vessel permeability *in vivo*. We quantified vessel permeability in primary tumors and metastatic tissues using immunofluorescence staining for albumin. Due to the large size of the protein, albumin cannot leak out of vessels with intact barrier function making albumin leakage into tissues a surrogate marker for vessel permeability[104]. In HCI010 primary tumors, we found a marked but not significant increase in vascular permeability (**Figure 2-9c**). In metastatic lungs from HCI010 *PHLDA2* spontaneous tumors, we observed a pronounced effect where PHLDA2 promotes nearly 2.3-fold higher levels of albumin (**Figure 2-10e)**. This suggests PHLDA2 overexpression increases vascular permeability specifically in lungs. Inversely, albumin signal was 0.4 times the control in lungs containing MDA-MB-231 PHLDA2 knockdown experimental metastases (**Figure 2-10f)**. Together this data demonstrates that PHLDA2 expression in tumor cells promotes vessel permeability in the metastatic lung tissue, while not changing primary tumor vessel permeability.

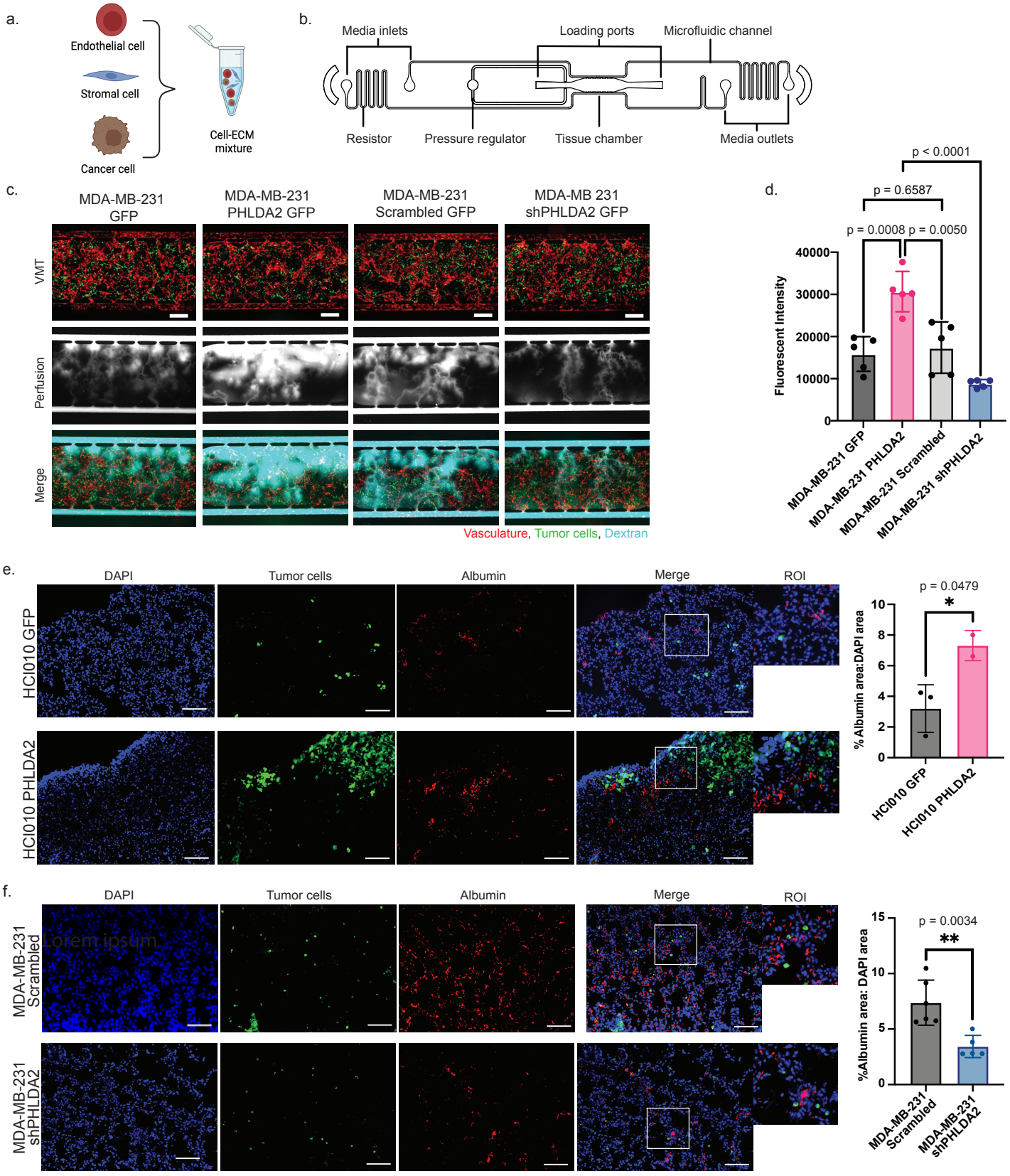


Figure 2-9: PHLDA2 increases vascular permeability in vitro and in vivo

a. Human endothelial cells, stromal cells, and MDA-MB-231 cancer cells were collected from 2D monolayer cultures and combined with extracellular matrix to establish vascularized microtumors (VMT). b. Schematic illustrates our device unit with a single 6 mm-long tissue chamber with medium inlets and outlets. The cell-ECM mixture was loaded into the tissue chamber through two ports. Loading is facilitated by a pressure regulator, and tissues are supported by hydrostatic pressure created across microfluidic channels connecting the media wells. Microfluidic resistors maintain physiological flow rates. c. Representative fluorescent images of VMTs containing MDA-MB-231 tumor cells (green) and vasculature (CD31, red), formed with either MDA-MB-231 GFP, MDA-MB-231 PHLDA2, MDA-MB-231 Scrambled, or MDA-MB-231 shPHLDA2. The greyscale images in the middle panel show 70 kD dextran perfused into the VMTs on day 4. The bottom panel displays a merge of the dextran (cyan) with the VMT images. Scale bar = 500μm. d. Quantification of dextran leakage for each VMT condition (n=5/group). Graph is displayed as mean ± SD; each point represents one VMT value obtained by the average of multiple 10x objective microscopic fields. e. Representative images of immunofluorescence staining for DAPI (blue), tumor cells (green), albumin (red) in HCI010 GFP and HCI010 PHLDA2 metastatic lung tissue. Percent of $albumin^+$ area normalized to $DAPI^+$ area was quantified (n= 2-3/group). Graph is displayed as mean ± SD; each point represents one lung tissue value obtained by the average of 5-10 20x objective microscopic fields. Student's t test results are represented on graph according to the following: $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$. f. Representative images of immunofluorescence staining for DAPI (blue), tumor cells (green), albumin (red) of metastatic lung tissue from mice intravenously injected with MDA-MB 231 Scrambled GFP (control) or shPHLDA2 GFP. Percent of $albumin^+$ area normalized to $DAPI^+$ area was quantified (n= 5-6/group). Graph is displayed as mean ± SD; each point represents one lung tissue value obtained by the average of 5-10 20x objective microscopic fields. Student's t test results are represented on graph according to the following: $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$.

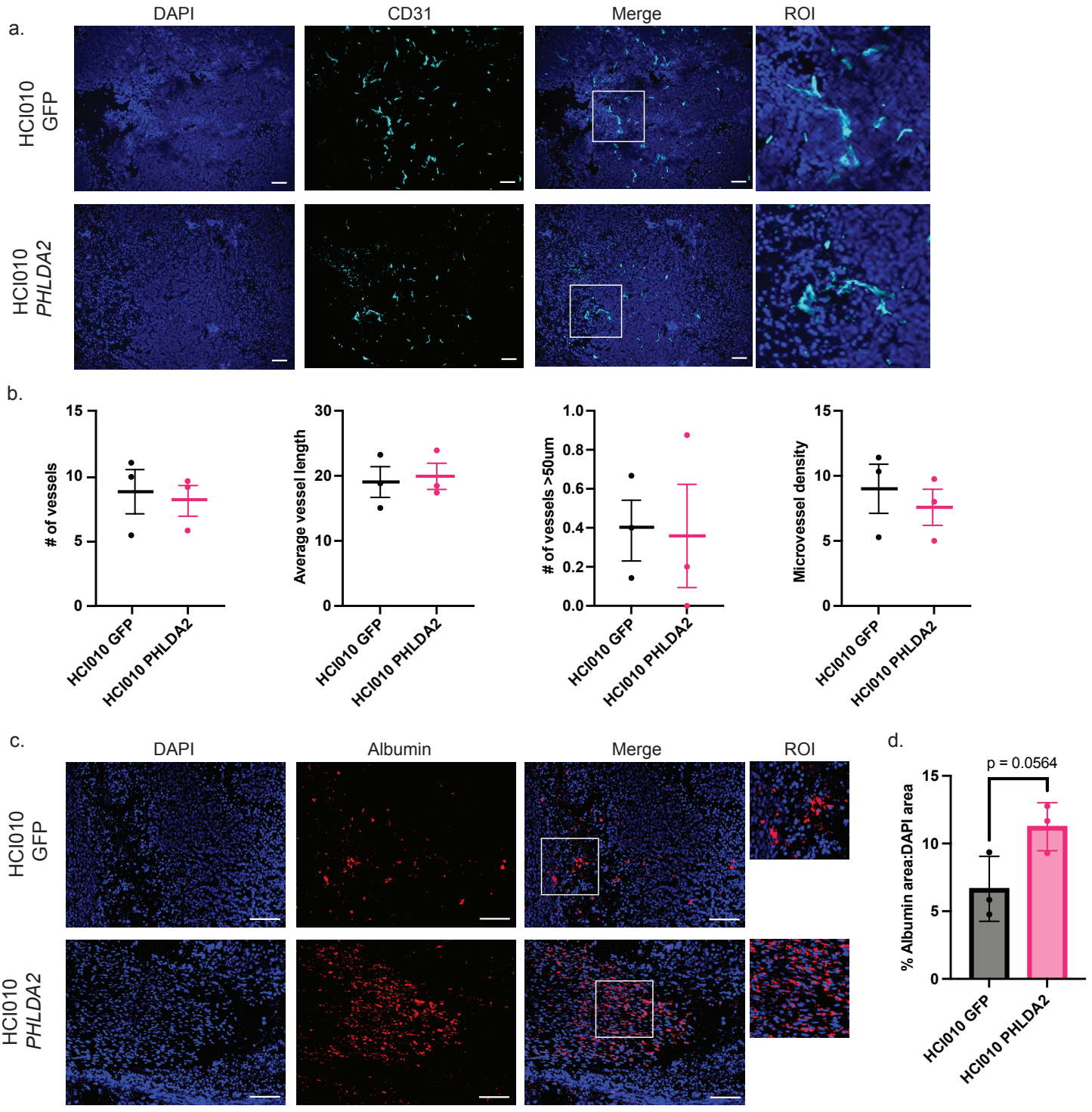


Figure 2-10: Increased PHLDA2 expression does not change vessel morphology but increases vessel permeability in primary tumors

a. Representative merged images of immunofluorescence staining for DAPI (blue) and CD31 (cyan) in HCI010 GFP and HCI010 PHLDA2 primary tumors. Scale bar = 50 μm. b. Quantification of vessel morphology parameters including number of vessels (n=3/group), average vessel length (n=3/group), number of elongated vessels greater than 50μm in length (n=3/group), and microvessel density (n=3/group). See methods for detailed description of quantification. Graph is displayed as mean ± SD; each point represents one tumor obtained by the average of 5-10 random 20x objective microscopic fields. c. Representative merged images of immunofluorescence staining for DAPI (blue) and albumin (red) in HCI010 GFP and HCI010 PHLDA2 primary tumors. Scale bar = 100 μm. d. Quantification of the percent of albumin$^+$ area normalized to DAPI$^+$ area in HCI010 GFP and HCI010 PHLDA2 primary tumors (n=3/group). Graph is displayed as mean ± SD; each point represents one tumor obtained by the average of 5-10 random 20x objective microscopic fields. Student's *t* test result is displayed on the graph.

Chapter 2.6. Discussion

To address the clinical challenges of treating breast cancer metastasis, it is necessary to discover molecular drivers and define their role in metastatic disease. Understanding the mechanisms of metastatic drivers will reveal novel drug targets which can greatly improve therapeutic strategy and patient survival. Our study addresses a significant gap of knowledge investigating the role of *PHLDA2*, a gene largely known for regulating placental development, in cancer. Specifically, we investigate the function of *PHLDA2* in breast cancer metastasis demonstrating that it is a pro-metastatic gene in TNBC models.

PHLDA2 is a small 17kD protein containing, most notably, a Pleckstrin Homology (PH) domain. PH domains interact with phosphoinositides (PIPs) in the plasma membrane and can function in potentiating signal transduction or act as scaffolds[45]. PHLDA family proteins have been demonstrated to bind to PIPs, inhibiting AKT signaling through competitive binding[33,38,44]. Additionally, PHLDA2 has been shown to bind to phosphatidic acid (PA) and phosphatidylserine (PS)[52,53]. PA is a secondary messenger that facilitates dynamics required for cell signaling through direct activation of kinases such as MAPK and PDK1, leading to cellular pathway regulation[105]. PS is presented on the surface of cancer cells, assisting in immune cell evasion[105]. Our bulk sequencing revealed hundreds of upregulated genes in response to *PHLDA2* overexpression; thus, it is plausible that PHLDA2 interacts with lipids on the cell membrane to promote signal transduction resulting in the transcription of a pro-metastatic meta-program.

Pathway analysis revealed genes upregulated in *PHLDA2* overexpressing cells were largely associated with ECM and blood vessel function. ECM components play critical roles in promoting cancer metastasis, including supporting circulating tumor cells, enhancing extravasation into the metastatic site, and promoting the formation a suitable metastatic niche[7]. Our data demonstrates that expression of PHLDA2 and ECM related genes such as COL1A1, SPARC and MFAP5 are positively correlated. Interestingly, *PHLDA2* increased collagen

deposition in metastatic lungs, but no change was detected in primary tumor tissue suggesting a differential regulation of collagen in the metastatic niche. In tumors, increased deposition of structural proteins including collagen, promote growth and survival of cancer cells[7]. This is consistent with a significant increase of Ki67+ cells and overall number of lesions observed in our *PHLDA2* overexpressing model, suggesting an enhancement in proliferation and seeding capacity. However, further studies are required to establish a causal relationship between ECM proteins and metastatic seeding and outgrowth.

Loss of endothelial cell barrier integrity resulting in vascular permeability enhances metastasis[106]. Consistent with this, we found overexpression of *PHLDA2* increases vascular permeability while knockdown decreases permeability in the metastatic lung. In our spontaneous model, *PHLDA2* overexpression also increased vascular permeability in primary tumor vasculature, illustrating that *PHLDA2* influences vessel leak regardless of tissue site. Bulk sequencing revealed upregulation of several genes associated with secreted factors which are reported to induce vascular permeability. Additionally, genes associated with cell-cell interaction were upregulated suggesting an enhancement in the attachment of tumor cells to endothelium to facilitate their extravasation. Cell transmigration and microvessel permeability are related though the mechanisms by which are still not fully understood. We hypothesize that collectively secreted factors and adhesion molecules contribute to vessel permeability and promote metastasis in *PHLDA2* overexpressing cells. Additionally, increased ECM stiffness reduces vascular barrier integrity; thus, ECM deposition may promote vessel permeability in the metastatic niche[95]. Others have linked PHLDA2 expression to tumor angiogenesis; however, our study is the first to our knowledge to demonstrate PHLDA2 expression alters vessel permeability[54].

Consistent with its pro-tumorigenic roles discussed above, *PHLDA2* is more highly expressed in primary breast tissue compared to normal. However, the mechanisms underlying its increased expression in breast tumor tissue remains elusive. Given that *PHLDA2* is an

imprinted gene, we hypothesized that its expression levels are modulated through demethylation. Consistent with this, we uncovered that loci cg04720330 and cg05167973 were significantly demethylated in primary tumor compared to normal tissue. Methylation levels at these loci negatively correlated with *PHLDA2* gene expression in normal breast and breast tumor tissue. Hypomethylation driven *PHLDA2* expression has been similarly reported in the context of clear cell renal carcinoma[86]. Notably, these two significant loci are located within the gene body, rather than in the promoter region. Thus, methylation at these loci likely cause chromatin densification making the gene less accessible to transcription machinery and resulting in less of gene expression, which we observe in normal tissue[107].

Overall, these results suggest a role for epigenetic regulators like demethylases in the regulation of *PHLDA2* expression in breast cancer. TNBC is considered broadly hypomethylated based on its low levels of methylation compared to other subtypes and normal breast tissue[23]. Interestingly, demethylase TET1 is upregulated in about 40% of TNBC patients and has been shown to increase oncogenic signaling through demethylation[23]. This illustrates a potential mechanism by which *PHLDA2* loci become demethylated in cancer to drive increased expression. Further investigation will be performed to define the relationships between demethylases, *PHLDA2*, and breast cancer metastasis.

Chapter 2.7. Conclusion

In summary, our work is the first evidence that expression of the metastatic gene, *PHLDA2*, is potentially regulated by a demethylation event in tumor cells. We further demonstrate that *PHLDA2* enhances cancer cell migration in addition to increasing ECM deposition and vascular permeability. Taken together, our study identified a marker of metastatic breast cancer that could be targeted to improve survival for those with metastatic disease, particularly in the TNBC subtype.

Chapter 3: Developing a Novel Methodology for Tissue Digestion and Flow Cytometry of Cancer Cells for Metabolomics

Chapter 3.1. Overview.

Cancer cell migration to secondary organs through metastasis is terminal and drives breast cancer patient fatality[13]. A better understanding of the metabolic drivers of metastasis can lead to targeted treatments which will affect patient outcome. Our lab performed single cell sequencing on primary tumors and metastatic lungs and lymph nodes from immunocompromised NOD scid gamma (NSG) mice carrying patient derived primary tumors (PDX)[17]. In this study we discovered that breast cancer cells switch their metabolism from glycolysis to oxidative phosphorylation (OXPHOS) during lung metastasis[17]. We further demonstrated that pharmacologic inhibition of OXPHOS attenuates metastatic spread *in vivo*.[17] This provided evidence that OXPHOS is critical for metastasis[17]. However, it remains unclear what nutrient source allows for metastatic cancer cells to make this metabolic switch during the metastatic cascade. Understanding what nutrient sources promote this invasive cancer cell behavior could lead to treatment of metastasis through cellular nutrient starvation mechanisms or macroscopic dietary intervention.

To determine what drives this metabolic switch in metastatic cancer cells, we have adapted and optimized a novel methodology for LC-MS using solid lung tissues sorted by florescence activated cell sorting (FACS). We have demonstrated that this method preserves the native metabolic signatures of tissues and allows for obtention of high-quality LC-MS metabolomics data. Using this novel method, steady state metabolomics can be performed to compare the profile of primary tumor and metastatic cells in the lungs in robust models of breast cancer metastasis. We found using a lung tropic MDA-MB 231 cell line, that metastatic cells in the lungs have higher levels of citrate metabolite than matched primary tumor cells. This finding gives us insights into what drives the metastatic capacity of breast cancer cells.

Chapter 3.2. Significance

To support increased proliferation, migration, and adaptation to new tissue environments, cancer cells frequently reprogram their metabolism to meet the high energy demands and biosynthetic requirements of metastasis[65]. Therefore, it is crucial to identify the specific metabolic changes that occur in metastatic cancer cells to develop innovative strategies for preventing metastasis by targeting key nutrient pathways. This project establishes a foundation for uncovering the metabolic profiles of both primary tumor cells and lung metastatic cancer cells through the development of a novel single-cell metabolomics methodology, which allows for analysis of cellular metabolism of rare cell populations. Furthermore, this work has identified a preliminary metabolite that appears to play a critical role in enhancing the metastatic potential of breast cancer cells, providing a promising target for future therapeutic interventions aimed at limiting cancer progression and improving patient outcomes.

Chapter 3.3. Innovation

The inability to assess the metabolic profile of cells that are in minority in a tissue is a significant limitation in the field of metastatic cancer metabolomics. To identify and characterize metastatic cancer cells, cell sorting methodology such as fluorescence activated cell sorting (FACS) is necessary to isolate cancer cells from a tissue. Metabolomics is typically performed on whole tissues which are preserved by flash freezing in liquid nitrogen. When studying cancer metastasis, the population of cancer cells of interest are in a vast minority in a tissue. Therefore, in a breast cancer lung metastasis model, performing metabolomics on whole lung tissue fails to distinguish the metabolic signatures of metastatic breast cancer cells from those of native lung cells.

Typically, to isolate metastatic breast cancer cells, lung tissues are enzymatically digested at 37°C to obtain single cell suspensions suitable for FACS sorting. However, it has been established that digestion at physiological temperatures does not preserve metabolomic

signatures[64]. A group has published a protocol that employs kidney tissue digestion using a serine protease optimal at cold temperatures to reduce single cell sequencing artifacts[65]. Additionally, another group has established a flow sorting protocol to recover rare hematopoietic stem cells for metabolomics[66]. By optimizing and integrating these published protocols for cold tissue digestion and flow sorting, we've developed a novel method to extract high yields of rare metastatic breast cancer cells from murine lungs while preserving their metabolic data. This protocol will allow for improved understanding of metabolomic profiles of cancer cells in their metastatic sites as well as comparison of profiles from the corresponding primary tumor.

Chapter 3.4. Materials and Methods

3.4.1. Generation of MDA-MB 231 Lung Tropic (LM2) Cell Line

250,000 MDA-MB 231 parental cells were intravenously (IV) injected into the tail vein of 4-6 week old NSG mice. After a 30-day incubation period, mice were sacrificed using asphyxiation with CO_2 followed by cervical dislocation and perfusion with 10 mM EDTA in PBS. Using sterile tools and a sterile work environment to prevent cell contamination, murine lungs were dissected and digested using Collagenase IV solution at 37°C. Cells were filtered and red blood cells were lysed using ACK Lysis Buffer. Cells were resuspended in DMEM media with 10% FBS and plated into a cell culture plate. The following day, media was changed to fresh DMEM with 10% FBS. Two days post dissociation, DMEM without FBS was added to eliminate fibroblasts. Media was changed to DMEM with 10% FBS and 1% Penicillin-Streptomycin antibiotic the following day. Once cells grew out, this was repeated to generate LM1 line. Once LM1 cells were growing well *in vitro*, cells were injected via IV as above. After 23 days post injection, mice were sacrificed as described prior to generate LM2 cell line. After generation of line, MDA-MB 231 LM2 cells were transfected with eGFly virus to generate a line with a traceable label by flow cytometry and *in vivo*

imaging capacity. This transduction reduced time in tissue digestion by eliminating staining step for flow cytometry as all cancer cells are labeled with GFP.

3.4.2. Orthotopic Mammary Fat Pad Implantation

All animal experiments were reviewed and approved by the University of California, Irvine Institutional Animal Care and Use Committee (IACUC). For bilateral orthotopic tumor injections, mice were placed under isoflurane anesthesia (1.5–2.0%). Mammary fat pad area was prepared by shaving and cleaning with alcohol swabs. 500,000 or 250,000 (study dependent) MDA-MB 231 Parental or LM2/LM2-eGFly cells were resuspended in 50 µL of phosphate buffered saline (PBS). Cells were mixed in a 1:1 ratio with Corning Matrigel Matrix and placed on ice until injection. Using sterilized tools, a small incision was made into the skin of the mouse in the center of the lower abdomen. Cell suspension (100 µL) was injected in the fourth mammary fat pad of 10-12 week old NSG mice on each side. Skin was secured using staples and veterinary glue bond. Pain killer was administered as directed. Tumors were allowed to grow and metastasize over a 30-day period. The study was compliant with all relevant ethical regulations regarding animal research.

3.4.3. Cold Protease Tissue Digestion

Mice were fasted from food 6 hours before sacrifice to eliminate variability of chow consumption on the metabolome. Mice were sacrificed by cervical dislocation as asphyxiation with CO_2 alters the metabolic profile of the tissue. Lungs and tumors were harvested onto ice-cold PBS on ice without perfusion which disrupts the metabolomic profile. 20-50 µL of blood was collected from the chest cavity and placed in a tube on ice to clot for 20-30 minutes to serve as a control. Blood was spun down at max speed for 10 minutes. Blood supernatant was collected and frozen at -80°C until ready for analysis. Enzyme mix was made on ice consisting of: 10 mg protease/1 mL tissue (Protease from Bacillus licheniformis Type VIII, Sigma Aldrich, Cat. No:

P5380), 5 mM $CaCl_2$, DNAse I (125 U/mL) (Worthington Biochemical, Cat. No. LS002139), and DPBS with Ca^{2+} and Mg^{2+} up to 1 mL volume. Tissues were minced on chilled petri dishes placed on ice with a razor blade. Tissue and enzyme mix (2 mL of enzyme mix /organ) were left on in a 15 mL conical tube ice for 7 minutes while pipetting 15 strokes using a 1 mL pipet (with tip cut off) every 2 minutes set to 700 µL. After 7 minutes, digest mix and tissue were transferred to a Miltenyi-C tube placed on ice. Program lung_02 was ran once for lungs and tumor sample. MACS tubes were gently spun at 4°C. Digestion of cells continued in tube on ice for 8 minutes while titrating every 2 minutes, 15 strokes using a 1 mL pipet. 6 mL ice cold 10% FBS/PBS was added to digest mix in C-tube to inhibit protease. The quenched mix was transferred to a 15 mL conical tube and spun at 600 x G for 5 minutes at 4°C. Supernatant was discarded and lung and tumor pellets were resuspended in 2 mL ice-cold PBS. Resuspended cells were filtered through a 40 µM filter into a sterile 50 mL conical on ice. Filter was rinsed with 5 mL ice cold PBS. To tumor cells, 45 mL of ice-cold PBS was added. Filtered tumor cells were spun at 100 x G for 5 minutes at 4°C to eliminate debris. Supernatant was aspirated. Pellet was resuspended in 1 mL DMEM media. Conical tube containing lung filtered cells was spun at 600 x G for 5 minutes at 4°C, before discarding the supernatant and resuspending the pellet in 1 mL DMEM media. During this procedure, always keep samples on ice or at 4°C. Mice should also be sacrificed individually so to minimize delays in sorting samples and loosing critical metabolites. The study is compliant with all relevant ethical regulations regarding animal research.

3.4.4. Flow Cytometry Set up for Sorting of Cold Digested Tissues

Flow cytometer was set up as detailed in previously published literature[66]. Sheath fluid that was used for these experiments was either 1X or 0.5X PBS. Sorted cells were sorted into cooled tube holders.

3.4.5. Labeling Cancer Cells for Flow Cytometry

For unlabeled human cell lines (MDA-MB 231 Parental, MDA-MB 231 LM2), human-specific antibody CD298, diluted 1:100 (APC, BioLegend, Cat. No. 341706), and the mouse-specific antibody MHC-I, diluted 1:100 (PE/Cy7, BioLegend, Cat. No. 114717) were used to label human cancer cells. Cell viability was determined by negative staining for SYTOX Blue, diluted 1:1000 (ThermoFisher Scientific Cat. No. S34857). Unlabeled cells were sorted for $CD298^{+}$, $MHCI^{-}$, $Sytox^{-}$. eGFly labeled cells were sorted for $Sytox^{-}$, GFP^{+}.

3.4.6. Metabolomics Protocol and Analysis—Collaboration with Dr. Cholsoon Jang and Johnny Le

A quadrupole orbitrap mass spectrometer (Q Exactive; ThermoFisher Scientific) was operated in negative ion mode with electrospray ionization and used to scan from m/z 70 to 1,000 at 2 Hz, with a 140,000 resolution. MS was coupled to a Vanquish UHPLC system (ThermoFisher Scientific) with autosampler temperature set at 5°C and injection volume 3ul. LC separation used a XBridge BEH Amide column (2.1 × 150 mm2, 2.5um particle size, 130 Å pore size; Waters Corporation) and run on a gradient of solvent A (95:5 water: acetonitrile with 20 mM of ammonium acetate and 20 mM of ammonium hydroxide, pH 9.45) and solvent B (acetonitrile). Flow rate was 150ul/min. The LC gradient was: 0 min, 85% B; 2 min, 85% B; 3 min, 80% B; 5 min, 80% B; 6 min, 75% B; 7 min, 75% B; 8 min, 70% B; 9 min, 70% B; 10 min, 50% B; 12 min, 50% B; 13 min, 25% B; 16 min, 25% B; 18 min, 0% B; 23 min, 0% B; 24 min, 85% B; and 30 min, 85% B. Preliminary data informed targeted analysis using MAVEN software. Metabolites were assessed that are related to glycolysis and the TCA cycle. Analysis was conducted utilizing R. Statistical testing was conducted using a t-test.

Chapter 3.5. Results

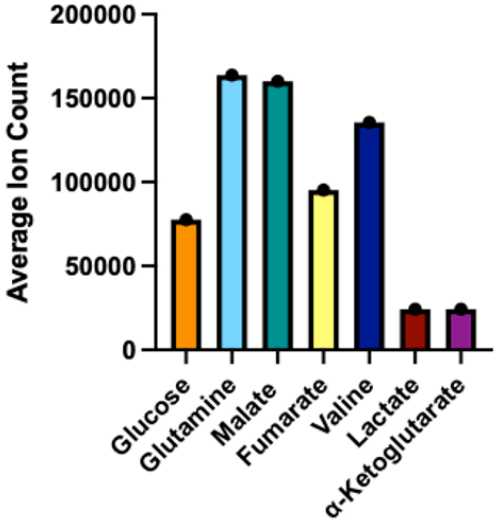
3.5.1. Novel Single Cell Metabolomics Method Recovers a Metabolomic Profile from Live Cancer Cells

To be able to identify metabolite profiles from metastatic lung cancer cells, it was necessary to optimize effective digestion of both lung and tumor tissues. In our experimentation, we have optimized the amount of cold protease enzyme necessary for effective digestion of lung and tumor tissues. The previously published protocol employs a mechanical tissue homogenizer for rapid and consistent tissue digestion of liver tissue[65]. Therefore, we refined the appropriate settings for lung and tumor tissue dissociation that can sufficiently digest these tissues while suitably recovering live and viable cells using the Countess Cell Counter **(Figure 3-1a)**. Additionally, we have confirmed that with cold protease digestion, live, GFP positive cancer metastatic cells can be recovered from the tumor and lung by flow cytometry. Finally, we have demonstrated that with cold protease digestion, significant metabolites can be effectively recovered from cold protease digested lung tissue using LC-MS analysis **(Figure 3-1b)**. Together, this approach enabled us to obtain a metabolic profile from live cancer cells using the optimized protocol. (**Figure 3-1c**).

a.

Average	Live Cell Count	Viability
Standard Digestion Lung	6.7 x 10^6 cells/mL	98%
Novel Digestion Lung	1.2 x 10^6 cells/mL	90%
Standard Digestion Tumor	1.4 x 10^6 cells/mL	67%
Novel Digestion Tumor	1.1 x 10^6 cells/mL	60%

b.



c.

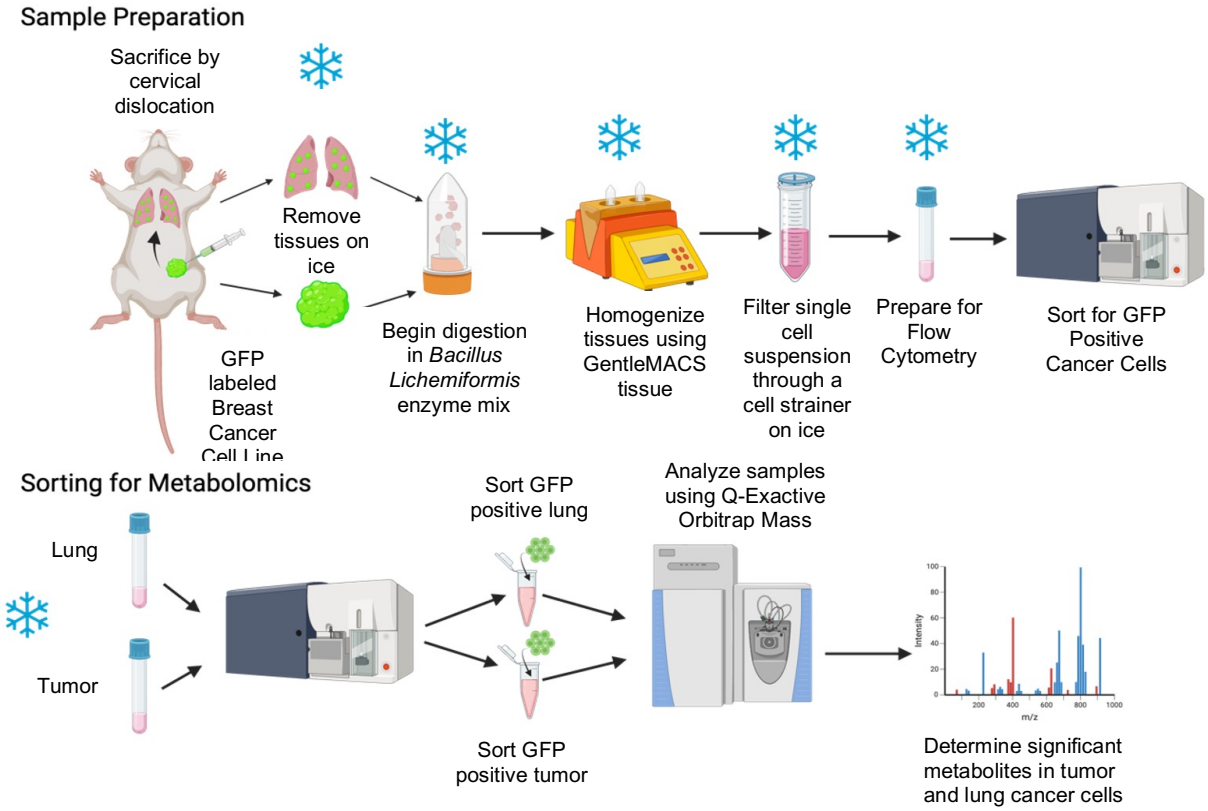


Figure 3-1: Cold Protease Tissue Digestion Recovers a Metabolomic Profile from Live Cancer Cells.

a. Table depicts average live cell count and viability of lung and tumor tissues digested with standard (37°C) protocol and novel digestion (cold protease) protocol. Countess was used to determine cell count and viability. b. Bar graph demonstrates average ion counts of select metabolites detected in high abundance in lung samples digested with the novel digestion (cold protease) methodology. These metabolite ion counts were normalized to the corresponding blood sample controls. c. Experimental schematic of single cell metabolomics digestion protocol sample prep and sorting.

3.5.2. At least 300,000 cells Must Be Recovered to Generate a Significant Metabolite Signature

To ensure we had enough cancer cells to obtain a significant metabolomic signature, we performed metabolomics on sorted quantities of cold digested and flow sorted live native lung cells. We sorted 10,000, 100,000 and 300,000 cells to determine the smallest ppm window for analysis and detection. In analyzing these samples, we sought to reduce background and visualize a true metabolite peak. We determined that 300,000 cells were minimally necessary to generate a significant metabolomics signature (**Figure 3-2**).

Metabolite	10K	100K	300K
Glycolysis			
Glucose	o	o	o
Fructose-6-phosphate	x	x	x
glyceraldehyde-3-phosphate	x	x	x
3-phosphoglycerate	x	x	x
pyruvate	x	x	o
lactate	x	x	x
TCA			
citrate	x	x	x
succinate	x	x	o
malate	x	x	o
fumarate	x	x	x
amino acids			
alanine	x	x	o
aspartate	x	o	o
glutamine	x	x	o
glutamate	x	x	o
valine	x	x	o
leucine	x	x	o
isoleucine	x	x	o
methionine	x	x	x
glycine	x	x	x
serine	x	x	o
histidine	x	x	x
lysine	x	x	o
arginine	x	x	o
tyrosine	x	x	o
tryptophan	x	x	o
phenylalanine	x	x	o
asparagine	x	x	x
cysteine	x	x	x
proline	x	x	o
threonine	x	x	o
fatty acids			
myristate	o	o	o
palmitate	o	o	o
oleic acid	o	o	o
linoleic acid	o	o	o
arachidonic acid	x	x	x

Figure 3-2: 300,000 Cells are Necessary to Recover a Significant Metabolite Profile.

Table demonstrates detected metabolites at various cell numbers. 10,000, 100,000 and 300,000 live lung cells were sorted using SytoxBlue Live/Dead stain. Green boxes with an o denote metabolite is detected with appropriate cell dose. X denotes peaks in samples that were less than the blank (sheath) or undetected. Ppm window denotes reduction of analysis window in order to reduce background and visualize metabolite peak in the sample.

3.5.3. A Breast Cancer Lung Tropic Cell Line Can Generate a Higher Yield of Lung Metastatic Cells than other Models

To recover 300,000 lung cancer cells, it was necessary to generate a new murine model to increase the yield of metastatic cancer cells in the lungs. MDA-MB 231 cells were passaged twice through the tail vein of mice, to select for cells with a lung metastatic capability to generate MDA-MB 231 LM2 lung tropic cells[108]. Generation of the lung tropic cell line assisted with increasing metastatic cancer cell yield in the lungs (**Table 3**). This cell line was later labeled with the eGFly fluorescent tag to eliminate the tissue staining step necessary with prior flow cytometry. Additionally, to compare metabolic profiles between primary tumor and metastatic cells more fairly, it is necessary to obtain an approximately equal amount of primary tumor and lung metastatic cells.

Table 1: Cell Counts of Sorted Primary Tumor and Metastatic Lung Cells from Breast Cancer Models

Cell Line	Injection Type	Cell # Injected	Duration	Tumor Cells Sorted at Sacrifice	Metastatic Cells Sorted from Lung at Sacrifice
MDA-MB 231 Parental	Bilateral orthotopic MFP injection	500,000	30 days	Mouse 1: 86,000[#] Mouse 2: 86,000*	Mouse 1: 1 Mouse 2: 414
MDA-MB 231 Parental	Bilateral orthotopic MFP injection	250,000	30 days	Mouse 1: 86,000*	Mouse 1: 604
MDA-MB 231 Parental	Bilateral orthotopic MFP injection	250,000	35 days	Mouse 1: 100,000 Mouse 2: Did not run	Mouse 1: 2 Mouse 2: 3
MDA-MB 231 LM2	Bilateral orthotopic MFP injection	500,000	24 days	Mouse 1: 45,926 Mouse 2: 35,792 Mouse 3: 37,088	Mouse 1: 221.286 Mouse 2: 52,273 Mouse 3: 600,000
MDA-MB 231 LM2 eGFly	Bilateral orthotopic MFP injection	500,000	22-24 days	Mouse 1: 450** Mouse 2: 64,662[#] Mouse 3: 33,147*** Mouse 4: 36,441****	Mouse 1: 450** Mouse 2: 155,004 Mouse 3: 33,147***

				Mouse 5: 32,354[#] Mouse 6: 170,153[#] Mouse 7: 93,393[#] Mouse 8: 257,587[#]	Mouse 4: 36,441**** Mouse 5: 88,620 Mouse 6: 300,000[&] Mouse 7: 300,147[&] Mouse 8: 292,593

Note: * denotes matched tumor/lung counts, [#] indicates if a primary tumor was sorted to sample completion, [&] denotes that sample was sorted to 300,000 cells and sorting was stopped)

3.5.4. Breast Cancer Lung Metastatic Cells have higher significantly expression of Citrate than Primary Tumor Cells

Through combination of MDA- MB 231 LM2 and MDA-MB 231 eGFly cohorts as described above with removal of low signal samples and scaling samples by their means for each metabolite, we found a significant presence of citrate in lung metastatic cells that was absent in primary tumor cells (**Figure 3-3**). There were non-significant changes in metabolites associated with Glycolysis (Glucose, Pyruvate and Lactate), Essential Amino Acids (Lysine and Arginine), Non-essential Amino Acids (L-Alanine and Glutamine) and Branched Chain Amino Acids (Isoleucine, Leucine and Valine) (**Figure 3-3**). This indicates that metastatic breast cancer cells in the lung have more citrate which might promote their metastatic capacity and survival in the metastatic niche.

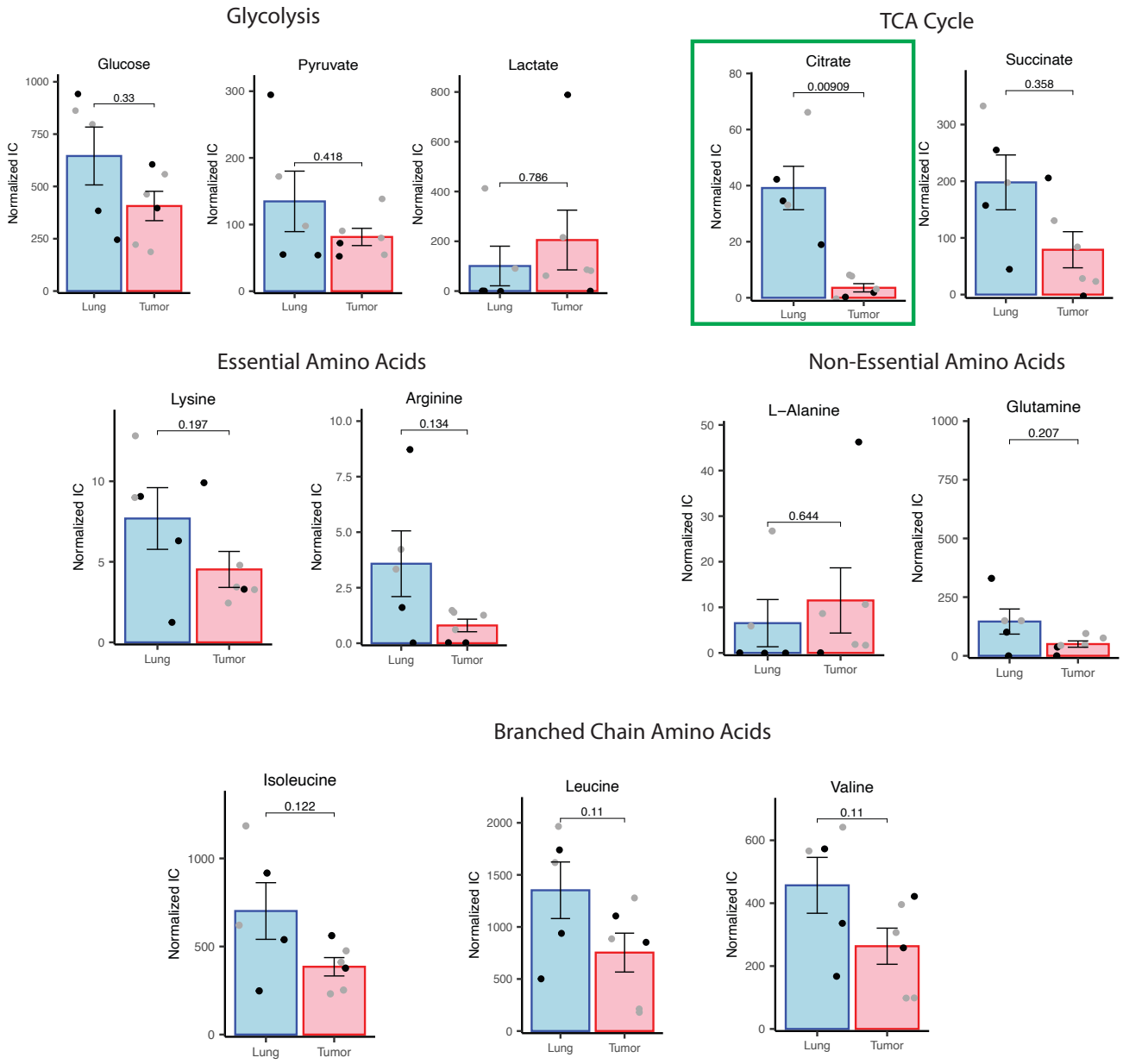


Figure 3-3: Significant Increase in Citrate was Observed in Metastatic Breast Cancer Cells as compared to Primary tumor cells

Bar plots compare steady state metabolite levels in breast cancer metastatic lung versus corresponding primary tumor in MDA-MB 231 LM2 a spontaneous tumor metastasis model. Samples were normalized to corresponding blood samples. Additionally, 2 cohorts of mice (MDA-MB 231 LM2-black points and MDA-MB 231 LM2 eGFly-grey points) were pooled together for analysis through normalization by scaling mean metabolite levels. MDA-MB 231 LM2 cohort was sorted using 1X PBS sheath fluid whereas MDA-MB 231 LM2 eGFly cohort was sorted into 0.5X PBS sheath fluid. The sheath fluid marginally improved metabolite yield (data not shown). Statistical significance was determined by a t-test

Chapter 3.6. Discussion

Assessing the metabolic differences between primary tumor and metastatic cells has been greatly limited by sample preparation constraints. Metabolomics techniques rely on whole tissue samples, which prevents recovery of minority metastatic cells from a tissue using flow cytometry. Additionally, tissues are often digested under 37°C conditions for flow cytometry which has been demonstrated to alter the metabolomic profile of tissues[64]. To address these challenges, we have developed a novel method for recovering significant metabolites from live breast cancer tumor and lung metastatic cells (**Figure 3-1**). While a published protocol existed for cold protease digestion of kidney tissue for single-cell sequencing, we optimized the digestion protocol for lung and tumor tissue[65]. This optimization addressed the architectural differences of kidney tissue from the adapted protocol to our tissues of interest[65]. Additionally, despite the presence of a flow sorting protocol for rare stem cells for metabolomics applications, we tailored the sorting conditions specifically for our cancer cell types of interest[66].

Alongside development of the digestion method, it was necessary to develop a model suitable for this protocol. Metastatic breast cancer cells are a largely rare cell population within the lung, so it is critical to be able to recover a high yield of viable cells with tissue digestion. Although metastatic breast cancer cells are in vast minority in the lung, these cells can significantly impact animal survival. Therefore, it was necessary to develop a breast cancer lung metastatic model that would yield the required cell number input necessary to recover significant metabolomic signatures (**Figure 3-2**). Through experimentation, we determined that due to the presence of the immune system in a 4T1 breast cancer model, it was difficult to obtain a significant number of lung metastasis while primary tumor size was within IACUC limitations. Additionally, our HCI010 PDX breast cancer model utilized in our prior work yields a small number of metastatic cells as the model is limited in its metastatic aggression. Using a more aggressive breast cancer cell line model, MDA-MB 231, that were passaged twice through the lungs of NSG mice to

generate the MDA-MB 231 LM2 line we were able to yield the most cancer cells in the lung without distress to the animal. While this model still faces challenges due to the significant mouse-to-mouse variability, using the MDA-MB 231 LM2 line we were able to recover the metastatic cell number necessary to perform metabolomics. The successful establishment of the MDA-MB 231 LM2 model allowed us to perform a pilot study examining the differences between metabolic profiles of primary tumors and metastatic breast cancer cells in a few independent animals.

Using the two spontaneous metastasis cohorts where we had recovered cell numbers sufficient for steady state metabolomics, MDA-MB 231 LM2 and MDA-MB 231 LM2 eGFly, we were able to observe differences in metabolites between primary tumors and their matched metastatic lungs (**Table 2**). We found that while there were no significant changes in amino acids or metabolites associated with glycolysis, we observe a significant change in citrate, a metabolite produced in the first step TCA cycle (**Figure 3-3**). We observe a higher content of citrate in metastatic breast cancer cells in the lung than their matching primary tumor (**Figure 3-3**). Using single cell sequencing, we have previously defined in a triple negative breast cancer PDX model that cancer cells make a metabolic switch from glycolysis to rely on oxidative phosphorylation (OXPHOS) from primary tumor to metastatic site[17]. As the TCA cycle is a critical part of OXPHOS, this finding aligns with our prior understanding that metastatic breast cancer cells rely on the TCA Cycle/OXPHOS for their metastatic potential[17]. Additionally, citrate has been found to play a role in metabolic regulation of cancer cells[109]. A shortcoming of steady state metabolomics is that presence of a metabolite does not discern if cells are using less of a highly detected metabolite or excreting more of this metabolite into their environment. To determine how citrate might be benefitting metastatic cancer cells, an isotope tracing study with radio-labeled citrate should be employed to understand how cancer cells are interacting with citrate during the metastatic cascade.

Chapter 3.7 Conclusion

To address the technical challenges of observing metabolic differences in primary tumor cells versus cancer cells in metastatic organs, we have established a novel single cell metabolomics method for primary tumor and metastatic lung cells. This method is compatible with flow cytometry sorting of live, rare cancer cells. Additionally, we have developed a best fit model to accommodate cell yield requirements for obtaining significant metabolite profile. Through a paired single cell metabolomics experiment with primary tumor and metastatic MDA-MB 231 LM2 cells, we found that metastatic lung cells have significantly higher levels of citrate than their primary tumors. This finding is aligned with our previous finding that breast cancer metastatic cells switch their metabolism to OXPHOS during the metastatic cascade as citrate is produced in the TCA cycle[17].

To better understand if citrate is consumed or generated by metastatic breast cancer cells, further experimentation with a radio-labeled citrate compound must be used. Additionally, citrate production can be reduced by targeting citrate synthase which is responsible for generating citrate[109]. The effect of reduced citrate on cancer cell metastasis can be observed using a knock-down of citrate synthase in the MDA-MB 231 LM2 cell line. This finding has potential to reduce metastatic cancer cells through cellular citrate nutrient restriction or citrate reduction by targeting its formation in the TCA Cycle in metastatic cancer cells. Overall, this method allows us to better study metabolic changes in rare cell populations in a tissue, such as metastatic cells.

Chapter 4: Conclusions and Final Remarks

Metastatic disease is a critical determinant of survival for breast cancer patients. The aggressive nature of metastatic cancer cells, combined with a lack of targeted treatment options, leads to significantly reduced survival rates as these cells spread to distant organs. My thesis work focuses on investigating novel genetic and metabolic drivers of breast cancer metastasis. A deeper understanding of the factors that drive cancer cells to metastasize not only illuminates the metastatic cascade but also creates opportunities for the development of innovative drug targets and biomarkers for the disease.

Chapter 2 of this thesis delves into *PHLDA2* as a genetic driver of breast cancer metastasis. Our prior research indicated that *PHLDA2* expression is elevated exclusively in metastatic cells compared to primary tumor cells, prompting us to investigate its role in metastatic burden and the mechanisms by which it facilitates metastasis. Notably, *PHLDA2* expression correlates with poor prognosis and decreased relapse-free survival in human patients. Additionally, *PHLDA2* expression is significantly elevated in tumor tissue relative to normal breast tissue. Given that *PHLDA2* is an imprinted gene, we aimed to elucidate the regulatory mechanisms governing its expression in cancer. Our findings suggest that methylation likely drives overexpression of this gene in cancer.

We demonstrated *in vivo* that *PHLDA2* promotes spontaneous metastasis in patient-derived xenograft (PDX) models. In experiments involving MDA-MB-231 cells, the loss of *PHLDA2* resulted in a notable reduction in metastatic burden. To further explore the role of *PHLDA2* in metastasis, we employed bulk sequencing to identify its downstream effects. Our analysis revealed that *PHLDA2 e*xpression enhances pathways related to extracellular matrix (ECM) remodeling, and vascular development. We validated these pathways, demonstrating that *PHLDA2* promotes collagen deposition and increased expression of ECM markers *SPARC* and *MFAP5*. Additionally, *PHLDA2* expression strongly influenced vessel permeability *in vitro*

using a vascularized organ-on-a-chip system. *In vivo,* we observed enhanced permeability in primary tumors with *PHLDA2* overexpression, which was even more pronounced in metastatic lungs. Together, these findings suggest that *PHLDA2* activates specific programs that facilitate metastasis. However, questions remain regarding the precise role of *PHLDA2* within the metastatic cascade and whether it directly regulates protein activity related to ECM dynamics and vascular permeability. Addressing these questions could provide valuable avenues for future exploration.

Metabolomics has limitations that make it challenging to study metastatic cells. Since cancer cells comprise a minority of the cellular composition in metastatic tissues, standard whole tissue metabolomics is ill-equipped to address these questions. Additionally, traditional tissue digestion methods can be disruptive to metabolomic analyses. In Chapter 3, I developed a protocol for flow sorting cancer cells that preserves their metabolic profiles. I also generated a lung tropic MDA-MB 231 model of spontaneous metastasis that can obtain a sufficient number of metastatic cells, 300,000, to facilitate significant metabolomics analysis. Through this work, I determined that citrate levels differ significantly between metastatic and primary tumor cells. However, this finding does not clarify whether the consumption of citrate drives this process or if citrate uptake does. To elucidate this, we would employ radiolabeled isotope tracing. For a spontaneous metastasis model, this would involve tracking the incorporation of the radiolabeled citrate into metabolic pathways, allowing us to determine the dynamics of citrate metabolism in both metastatic and primary tumor cells.

In conclusion, this thesis provides valuable insights into the genetic and metabolic factors driving breast cancer metastasis, a critical determinant of patient survival. Chapter 2 highlights the role of *PHLDA2* as a potential genetic driver, revealing its association with poor prognosis and its impact on metastatic processes through mechanisms such as ECM remodeling and vascular permeability. Finally, Chapter 3 introduces a novel approach to studying the metabolomics of metastatic cancer cells, paving the way for a deeper

understanding of metabolic alterations associated with metastasis. By elucidating the dynamics of citrate metabolism and its potential implications, this work opens avenues for the development of targeted therapies and biomarkers that could improve outcomes for breast cancer patients. Moving forward, it is imperative that we integrate our findings on genetic and metabolic drivers to unravel the complexities of the metastatic cascade, ultimately leading to the identification of transformative therapeutic strategies that can significantly enhance patient survival.

References

1. Erratum: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **70**, 313–313 (2020).
2. Chaffer, C. L. & Weinberg, R. A. A perspective on cancer cell metastasis. *Science* **331**, 1559–1564 (2011).
3. Wu, Q. *et al.* Breast cancer subtypes predict the preferential site of distant metastases: a SEER based study. *Oncotarget* **8**, 27990–27996 (2017).
4. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31–46 (2022).
5. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
6. Huang, J. *et al.* Extracellular matrix and its therapeutic potential for cancer treatment. *Signal Transduct. Target. Ther.* **6**, 1–24 (2021).
7. Sleeboom, J. J. F. *et al.* The extracellular matrix as hallmark of cancer and metastasis: From biomechanics to therapeutic targets. *Sci. Transl. Med.* **16**, eadg3840 (2024).
8. Zhang, Q., An, Z.-Y., Jiang, W., Jin, W.-L. & He, X.-Y. Collagen code in tumor microenvironment: Functions, molecular mechanisms, and therapeutic implications. *Biomed. Pharmacother.* **166**, 115390 (2023).
9. Moore, K. H., Murphy, H. A. & George, E. M. The glycocalyx: a central regulator of vascular function. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **320**, R508–R518 (2021).
10. Liu, Z.-L., Chen, H.-H., Zheng, L.-L., Sun, L.-P. & Shi, L. Angiogenic signaling pathways and anti-angiogenic therapy for cancer. *Signal Transduct. Target. Ther.* **8**, 1–39 (2023).
11. Acerbi, I. *et al.* Human Breast Cancer Invasion and Aggression Correlates with ECM Stiffening and Immune Cell Infiltration. *Integr. Biol. Quant. Biosci. Nano Macro* **7**, 1120–1134 (2015).
12. Cho, N. Imaging features of breast cancer molecular subtypes: state of the art. *J. Pathol. Transl. Med.* **55**, 16–25 (2021).
13. Howlader, N., Cronin, K. A., Kurian, A. W. & Andridge, R. Differences in Breast Cancer Survival by Molecular Subtypes in the United States. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **27**, 619–626 (2018).

14. Masoud, V. & Pagès, G. Targeted therapies in breast cancer: New challenges to fight against resistance. *World J. Clin. Oncol.* **8**, 120–134 (2017).
15. Senkus, E. & Łacko, A. Over-treatment in metastatic breast cancer. *The Breast* **31**, 309–317 (2017).
16. DeRose, Y. S. *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).
17. Davis, R. T. *et al.* Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. *Nat. Cell Biol.* **22**, 310–320 (2020).
18. Peters, J. The role of genomic imprinting in biology and disease: an expanding view. *Nat. Rev. Genet.* **15**, 517–530 (2014).
19. Jensen, A. B., Tunster, S. J. & John, R. M. The significance of elevated placental *PHLDA2* in human growth restricted pregnancies. *Placenta* **35**, 528–532 (2014).
20. Holm, T. M. *et al.* Global loss of imprinting leads to widespread tumorigenesis in adult mice. *Cancer Cell* **8**, 275–285 (2005).
21. Murrell, A. Genomic Imprinting and Cancer: From Primordial Germ Cells to Somatic Cells. *Sci. World J.* **6**, 1888–1910 (2006).
22. Lim, D. H. K. & Maher, E. R. Genomic imprinting syndromes and cancer. *Adv. Genet.* **70**, 145–175 (2010).
23. Panjarian, S. & Issa, J.-P. J. The Roles of DNA Demethylases in Triple-Negative Breast Cancer. *Pharmaceuticals* **14**, 628 (2021).
24. Fang, F. *et al.* Breast Cancer Methylomes Establish an Epigenomic Foundation for Metastasis. *Sci. Transl. Med.* **3**, 75ra25 (2011).
25. Good, C. R. *et al.* A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. *Nucleic Acids Res.* **45**, 8269–8281 (2017).
26. Salas, M. *et al.* Placental growth retardation due to loss of imprinting of *Phlda2*. *Mech. Dev.* **121**, 1199–1210 (2004).
27. Simmons, D. G. & Cross, J. C. Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. *Dev. Biol.* **284**, 12–24 (2005).

28. Tunster, S. J., Creeth, H. D. J. & John, R. M. The imprinted Phlda2 gene modulates a major endocrine compartment of the placenta to regulate placental demands for maternal resources. *Dev. Biol.* **409**, 251–260 (2016).
29. Singh, V. B. *et al.* Blocked transcription through KvDMR1 results in absence of methylation and gene silencing resembling Beckwith-Wiedemann syndrome. *Dev. Camb. Engl.* **144**, 1820–1830 (2017).
30. Wu, Y. *et al.* A spatiotemporal transcriptomic atlas of mouse placentation. *Cell Discov.* **10**, 1–17 (2024).
31. Ladurner, A. G. & Fersht, A. R. Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J. Mol. Biol.* **273**, 330–337 (1997).
32. Chang, S.-H., Chang, W.-L., Lu, C.-C. & Tarn, W.-Y. Alanine repeats influence protein localization in splicing speckles and paraspeckles. *Nucleic Acids Res.* **42**, 13788–13798 (2014).
33. Fuselier, T. T. & Lu, H. PHLD Class Proteins: A Family of New Players in the p53 Network. *Int. J. Mol. Sci.* **21**, 3543 (2020).
34. Nagai, M. A. Pleckstrin homology-like domain, family A, member 1 (PHLDA1) and cancer. *Biomed. Rep.* **4**, 275–281 (2016).
35. Park, C. G., Lee, S. Y., Kandala, G., Lee, S. Y. & Choi, Y. A Novel Gene Product that Couples TCR Signaling to Fas(CD95) Expression in Activation-Induced Cell Death. *Immunity* **4**, 583–591 (1996).
36. Ren, L. *et al.* Characterization of the metastatic phenotype of a panel of established osteosarcoma cells. *Oncotarget* **6**, 29469–29481 (2015).
37. Totzeck, F., Andrade-Navarro, M. A. & Mier, P. The Protein Structure Context of PolyQ Regions. *PLoS ONE* **12**, e0170801 (2017).
38. Chen, Y. *et al.* PHLDA1, another PHLDA family protein that inhibits Akt. *Cancer Sci.* **109**, 3532–3542 (2018).
39. Coutinho-Camillo, C. M. *et al.* Expression of PAR-4 and PHLDA1 is prognostic for overall and disease-free survival in oral squamous cell carcinomas. *Virchows Arch.* **463**, 31–39 (2013).
40. Zhao, P., Lu, Y. & Liu, L. Correlation of decreased expression of PHLDA1 protein with malignant phenotype of gastric adenocarcinoma. *Int. J. Clin. Exp. Pathol.* **8**, 5230–5235 (2015).

41. ZHAO, P., LI, X., LU, Y. & LIU, L. Downregulated expression of PHLDA1 protein is associated with a malignant phenotype of cholangiocarcinoma. *Oncol. Lett.* **10**, 895–900 (2015).
42. Johnson, E. O. *et al.* PHLDA1 is a crucial negative regulator and effector of Aurora A kinase in breast cancer. *J. Cell Sci.* **124**, 2711–2722 (2011).
43. Hossain, G. S. *et al.* TDAG51 Is Induced by Homocysteine, Promotes Detachment-mediated Programmed Cell Death, and Contributes to the Development of Atherosclerosis in Hyperhomocysteinemia*. *J. Biol. Chem.* **278**, 30317–30327 (2003).
44. Kawase, T. *et al.* PH domain-only protein PHLDA3 is a p53-regulated repressor of Akt. *Cell* **136**, 535–550 (2009).
45. Singh, N. *et al.* Redefining the specificity of phosphoinositide-binding by human PH domain-containing proteins. *Nat. Commun.* **12**, 4339 (2021).
46. Dannemann, N., Hart, J. R., Ueno, L. & Vogt, P. K. Phosphatidylinositol 4,5-bisphosphate (PIP2)-specific AKT1 is oncogenic. *Int. J. Cancer J. Int. Cancer* **127**, 239–244 (2010).
47. Frank, D. *et al.* A novel pleckstrin homology-related gene family defined by lpl/Tssc3, TDAG51, and Tih1: Tissue-specific expression, chromosomal location, and parental imprinting. *Mamm. Genome* **10**, 1150–1159 (1999).
48. Ohki, R. *et al.* PHLDA3 is a novel tumor suppressor of pancreatic neuroendocrine tumors. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E2404-2413 (2014).
49. Takikawa, M. & Ohki, R. A vicious partnership between AKT and PHLDA3 to facilitate neuroendocrine tumors. *Cancer Sci.* **108**, 1101–1108 (2017).
50. Takao, T., Asanoma, K., Tsunematsu, R., Kato, K. & Wake, N. The Maternally Expressed Gene Tssc3 Regulates the Expression of MASH2 Transcription Factor in Mouse Trophoblast Stem Cells through the AKT-Sp1 Signaling Pathway *. *J. Biol. Chem.* **287**, 42685–42694 (2012).
51. Zhao, G. *et al.* TSSC3 promotes autophagy via inactivating the Src-mediated PI3K/Akt/mTOR pathway to suppress tumorigenesis and metastasis in osteosarcoma, and predicts a favorable prognosis. *J. Exp. Clin. Cancer Res.* **37**, 188 (2018).
52. Yang, X. *et al.* PHLDA2-mediated phosphatidic acid peroxidation triggers a distinct ferroptotic response during tumor suppression. *Cell Metab.* **36**, 762-777.e9 (2024).

53. Wang, X. *et al.* PHLDA2 is a key oncogene-induced negative feedback inhibitor of EGFR/ErbB2 signaling via interference with AKT signaling. *Oncotarget* **9**, 24914–24926 (2018).
54. Feng, K., Peng, H., Lv, Q. & Zhang, Y. PHLDA2 reshapes the immune microenvironment and induces drug resistance in hepatocellular carcinoma. *Oncol. Res.* **32**, 1063 (2024).
55. Wei, Q., Qian, Y., Yu, J. & Wong, C. C. Metabolic rewiring in the promotion of cancer metastasis: mechanisms and therapeutic implications. *Oncogene* **39**, 6139–6156 (2020).
56. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).
57. Koppenol, W. H., Bounds, P. L. & Dang, C. V. Otto Warburg's contributions to current concepts of cancer metabolism. *Nat. Rev. Cancer* **11**, 325–337 (2011).
58. Camarda, R. *et al.* Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **22**, 427–432 (2016).
59. Dornier, E. *et al.* Glutaminolysis drives membrane trafficking to promote invasiveness of breast cancer cells. *Nat. Commun.* **8**, 2255 (2017).
60. Yang, L. *et al.* Metabolic shifts toward glutamine regulate tumor growth, invasion and bioenergetics in ovarian cancer. *Mol. Syst. Biol.* **10**, 728 (2014).
61. Christen, S. *et al.* Breast Cancer-Derived Lung Metastases Show Increased Pyruvate Carboxylase-Dependent Anaplerosis. *Cell Rep.* **17**, 837–848 (2016).
62. Elia, I. *et al.* Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nat. Commun.* **8**, 15267 (2017).
63. Antalis, C. J., Uchida, A., Buhman, K. K. & Siddiqui, R. A. Migration of MDA-MB-231 breast cancer cells depends on the availability of exogenous lipids and cholesterol esterification. *Clin. Exp. Metastasis* **28**, 733–741 (2011).
64. Fang, M. *et al.* Thermal Degradation of Small Molecules: A Global Metabolomic Investigation. *Anal. Chem.* **87**, 10935–10941 (2015).
65. Adam, M., Potter, A. S. & Potter, S. S. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Dev. Camb. Engl.* **144**, 3625–3632 (2017).

66. DeVilbiss, A. W. *et al.* Metabolomic profiling of rare cell populations isolated by flow cytometry from tissues. *eLife* **10**, e61980 (2021).
67. Tasdogan, A. *et al.* Metabolic heterogeneity confers differences in melanoma metastatic potential. *Nature* **577**, 115–120 (2020).
68. Ferraro, G. B. *et al.* Fatty acid synthesis is required for breast cancer brain metastasis. *Nat. Cancer* **2**, 414–428 (2021).
69. Frank, D. *et al.* Placental overgrowth in mice lacking the imprinted gene Ipl. *Proc. Natl. Acad. Sci.* **99**, 7490–7495 (2002).
70. Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
71. Savage, H. *et al.* Aerobic Exercise Alters the Melanoma Microenvironment and Modulates ERK5 S496 Phosphorylation. *Cancer Immunol. Res.* **11**, 1168–1183 (2023).
72. Blighe, Kevin. EnhancedVolcano.
73. Kolde, Ravio. pheatmap.
74. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS J. Integr. Biol.* **16**, 284–287 (2012).
75. Carlson, Marc. org.Hs.eg.db.
76. Sobrino, A. *et al.* 3D microtumors in vitro supported by perfused vascular networks. *Sci. Rep.* **6**, 31589 (2016).
77. Hachey, S. J. *et al.* An in vitro vascularized micro-tumor model of human colorectal cancer recapitulates in vivo responses to standard-of-care therapy. *Lab. Chip* **21**, 1333–1351 (2021).
78. Hachey, S. J. *et al.* A human vascularized microtumor model of patient-derived colorectal cancer recapitulates clinical disease. *Transl. Res. J. Lab. Clin. Med.* **255**, 97–108 (2023).
79. Hachey, S. J., Gaebler, D. & Hughes, C. C. W. Establishing a Physiologic Human Vascularized Micro-Tumor Model for Cancer Research. *J. Vis. Exp. JoVE* (2023) doi:10.3791/65865.
80. Hachey, S. J. *et al.* Targeting tumor–stromal interactions in triple-negative breast cancer using a human vascularized micro-tumor model. *Breast Cancer Res.* **26**, 5 (2024).

81. Phan, D. T. T. *et al.* A vascularized and perfused organ-on-a-chip platform for large-scale drug screening applications. *Lab. Chip* **17**, 511–520 (2017).
82. Chandrashekar, D. S. *et al.* UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia N. Y. N* **19**, 649–658 (2017).
83. Chandrashekar, D. S. *et al.* UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia N. Y. N* **25**, 18–27 (2022).
84. Győrffy, B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput. Struct. Biotechnol. J.* **19**, 4101–4109 (2021).
85. Györffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123**, 725–731 (2010).
86. Zhao, J. *et al.* Epigenetic modification of PHLDA2 is associated with tumor microenvironment and unfavorable outcome of immune checkpoint inhibitor-based therapies in clear cell renal cell carcinoma. *Eur. J. Med. Res.* **29**, 378 (2024).
87. Ma, D. *et al.* Patient-derived xenograft culture-transplant system for investigation of human breast cancer metastasis. *Commun. Biol.* **4**, 1–15 (2021).
88. Tichet, M. *et al.* Tumour-derived SPARC drives vascular permeability and extravasation through endothelial VCAM1 signalling to promote metastasis. *Nat. Commun.* **6**, 6993 (2015).
89. Yu, J. E. *et al.* Significance of chitinase-3-like protein 1 in the pathogenesis of inflammatory diseases and cancer. *Exp. Mol. Med.* **56**, 1–18 (2024).
90. Sanchez, T. *et al.* Induction of vascular permeability by the sphingosine-1-phosphate receptor-2 (S1P2R) and its downstream effectors ROCK and PTEN. *Arterioscler. Thromb. Vasc. Biol.* **27**, 1312–1318 (2007).
91. Nouvion, A.-L. *et al.* CEACAM1: a key regulator of vascular permeability. *J. Cell Sci.* **123**, 4221–4230 (2010).
92. Propson, N. E., Roy, E. R., Litvinchuk, A., Köhl, J. & Zheng, H. Endothelial C3a receptor mediates vascular inflammation and blood-brain barrier permeability during aging. *J. Clin. Invest.* **131**, e140966 (2021).

93. Tauseef, M. *et al.* Activation of Sphingosine Kinase-1 Reverses the Increase in Lung Vascular Permeability Through Sphingosine-1-Phosphate Receptor Signaling in Endothelial Cells. *Circ. Res.* **103**, 1164–1172 (2008).

94. Wu, Z. *et al.* MFAP5 promotes tumor progression and bone metastasis by regulating ERK/MMP signaling pathways in breast cancer. *Biochem. Biophys. Res. Commun.* **498**, 495–501 (2018).

95. Karki, P. & Birukova, A. A. Substrate stiffness-dependent exacerbation of endothelial permeability and inflammation: mechanisms and potential implications in ALI and PH (2017 Grover Conference Series). *Pulm. Circ.* **8**, 2045894018773044 (2018).

96. Xu, S. *et al.* The role of collagen in cancer: from bench to bedside. *J. Transl. Med.* **17**, 309 (2019).

97. PUCHTLER, H. & ISLER, H. THE EFFECT OF PHOSPHOMOLYBDIC ACID ON THE STAINABILITY OF CONNECTIVE TISSUES BY VARIOUS DYES. *J. Histochem. Cytochem.* **6**, 265–270 (1958).

98. Bielenberg, D. R. & Zetter, B. R. The Contribution of Angiogenesis to the Process of Metastasis. *Cancer J. Sudbury Mass* **21**, 267–273 (2015).

99. Arnold, S. A. & Brekken, R. A. SPARC: a matricellular regulator of tumorigenesis. *J. Cell Commun. Signal.* **3**, 255–273 (2009).

100. Skubitz, K. M. & Skubitz, A. P. Interdependency of CEACAM-1, -3, -6, and -8 induced human neutrophil adhesion to endothelial cells. *J. Transl. Med.* **6**, 78 (2008).

101. Götz, L., Rueckschloss, U., Najjar, S. M., Ergün, S. & Kleefeldt, F. Carcinoembryonic antigen-related cell adhesion molecule 1 in cancer: Blessing or curse? *Eur. J. Clin. Invest.* e14337 (2024) doi:10.1111/eci.14337.

102. Ergün, S. *et al.* CEA-Related Cell Adhesion Molecule 1: A Potent Angiogenic Factor and a Major Effector of Vascular Endothelial Growth Factor. *Mol. Cell* **5**, 311–320 (2000).

103. Ghavampour, S. *et al.* Endothelial barrier function is differentially regulated by CEACAM1-mediated signaling. *FASEB J.* **32**, 5612–5625 (2018).

104. Vandoorne, K., Addadi, Y. & Neeman, M. Visualizing vascular permeability and lymphatic drainage using labeled serum albumin. *Angiogenesis* **13**, 75–85 (2010).

105. Szlasa, W., Zendran, I., Zalesińska, A., Tarek, M. & Kulbacka, J. Lipid composition of the cancer cell membrane. *J. Bioenerg. Biomembr.* **52**, 321–342 (2020).

106. Tomita, T., Kato, M. & Hiratsuka, S. Regulation of vascular permeability in cancer metastasis. *Cancer Sci.* **112**, 2966–2974 (2021).

107. Wang, Q. *et al.* Gene body methylation in cancer: molecular mechanisms and clinical applications. *Clin. Epigenetics* **14**, 154 (2022).

108. Minn, A. J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).

109. Icard, P. *et al.* Understanding the Central Role of Citrate in the Metabolism of Cancer Cells and Tumors: An Update. *Int. J. Mol. Sci.* **22**, 6587 (2021).