Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jessica Marie Konen Date

Dissecting 3-D invasion mechanisms in lung cancer

By

Jessica Marie Konen
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Cancer Biology

Adam Marcus, Ph.D.
Advisor

Lawrence Boise, Ph.D.
Committee Member

Andrew Kowalczyk, Ph.D.
Committee Member

Paula Vertino, Ph.D.
Committee Member

Wei Zhou, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Dissecting 3-D invasion mechanisms in lung cancer

By

Jessica Marie Konen
Bachelor's of Arts, 2010

Advisor:
Adam Marcus, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Graduate Division of Biological and Biomedical Sciences
Cancer Biology
2016

Dissecting 3-D invasion mechanisms in lung cancer

By

Jessica Marie Konen

The process of metastasis is the leading cause of cancer-related deaths, killing about 90% of all cancer patients. An early step in the metastatic cascade is local invasion of tumor cells into the microenvironment, a process that arises due to aberrant cell polarity, reduction of cell-cell contacts, and ultimately increased motility. A known regulator of cell polarity is LKB1, a serine/threonine kinase that is frequently mutated in lung adenocarcinomas. Other known functions of LKB1 involve regulating cell metabolism, growth, and adhesion; thus, the loss of LKB1 in lung adenocarcinomas likely has a vast impact on the tumor cell biology. However, the role of the two known functional domains of LKB1, the kinase domain and C-terminal domain with a membrane-targeting farnesylation motif, in regulating early invasion has not been well studied, especially in the context of physiologically relevant 3-D matrices. Our work illustrates that each functional domain of LKB1 has an independent role. The farnesylation of LKB1 in the C-terminal domain is necessary for maintaining a polarized phenotype during 3-D invasion, and this is dependent on RhoA activity. Conversely, the kinase activity of LKB1 functions to regulate adhesion signaling, via repression of focal adhesion kinase (FAK). When the kinase activity of LKB1 is disrupted, FAK is hyperactive, which in turn increases the ability of the cell to reorganize and align the matrix during invasion. While the study of LKB1 in lung adenocarcinoma progression and metastasis is important, many patients do not harbor mutations in this gene; thus, taking a broader perspective on lung cancer invasion was warranted. In 3-D lung cancer models that utilize collective invasion, there is phenotypic heterogeneity with the collective invasion chain, with highly invasive leader cells that pioneer invasion and cells which immediately attach to and follow the leader cell, termed followers. In order to dissect these two subpopulations within the collective invasion chain, we created an image-guided genomics technique termed SaGA (spatiotemporal genomic and cellular analysis). With this technique, we were able to isolate, analyze, and amplify purified populations of leader and follower cells. We found that leader cells are highly stable, maintaining their invasive capabilities even when cultured as a purified population. Leader cells can also promote the motility and invasion of the noninvasive follower population. Transcriptome and molecular analyses revealed that the formation of the collective unit is dependent upon VEGF-based signaling, and that leader cells seem to hijack this canonical vasculogenic machinery to promote follower invasion. Additionally, leader cells themselves utilize cell-ECM interactions to create force on the microenvironment to promote fibrous highways to facilitate their invasion. Conversely, follower cells are a highly proliferative population, whereas leader cells struggle to grow and survive as a purified population. Follower cells rectify leader cell mitotic defects when grown in co-culture, and this is through an unknown soluble factor. Taken together, our data suggest that leader and follower cells within the lung cancer collective invasion chain possess a symbiotic relationship, wherein leaders provide followers with an escape mechanism while followers provide leaders with increased growth and survival. Collectively, these studies dissect 3-D lung cancer invasion mechanisms, providing novel insights into this early step in the metastatic cascade.

Dissecting 3-D invasion mechanisms in lung cancer

By

Jessica Marie Konen
Bachelor's of Arts, 2010

Advisor:
Adam Marcus, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Graduate Division of Biological and Biomedical Sciences
Cancer Biology
2016

Acknowledgements

Many people deserve my appreciation and acknowledgement for helping me to complete this work and finish my degree.

To my husband and best friend, for so many reasons but I will name just a few. Firstly, it was because of his job transfer to Atlanta that pushed me to apply to the new Cancer Biology program at Emory. While a difficult decision at the time, it was clearly the best one. All of my success is owed to that move and was such a major step in the right direction for my career. Secondly and most importantly, he was the one that had to be my emotional support on a day-to-day basis, through the daily struggles of coursework, research, writing papers, and just the general stress associated with graduate school. When I needed a laugh, he was there with a joke. When I needed someone in my corner, he was ready to take my side no matter what. I truly don't know if I would have made it through all the challenges of the past several years without him. He was always my biggest cheerleader, relentlessly bragging about me to his family and friends. His endless belief in me was always a great motivator to keep striving to succeed every day.

To my family, I owe the deepest gratitude. Moving away from home and my central support system was one of the hardest things I have done. However, even with the distance, they continued to be the rocks that I lean upon. During qualifying exams, my sister was the one who sent the care package full of delicious sweets to power me through those late-night study sessions, and her yearly visits to Atlanta were some of the best times I had during graduate school. My brother, while mostly a quiet presence, always let me know that he was proud of me, and he was there for me without question in one of my greatest times of need. When I needed advice on anything, my dad was always there with wise words and a sarcastic quip to help guide me towards the right decision. And lastly, I want to thank my mom,

who is a reminder of why this career path, while stressful at times, means so much to myself and to the millions of other cancer survivors and their family members. She is a constant source of inspiration, whether intentionally or not. Seeing her bravery and strength motivates me to keep going.

To my mentor, Adam Marcus, for his unwavering belief in my potential. While I don't always have faith and confidence in myself, Adam always seemed to have enough for both of us. I can honestly say that it is because of his constant encouragement that I am pursuing a career in academia and running my own lab. Any of the success that came from the work we developed together, he made sure to share with me, something that I'm not sure all mentors would do for their students. When I had certain struggles come up in my life, he didn't hesitate to offer his support and genuinely cared about my well-being. Without question, Adam was the best mentor I could have asked for to help me not to just make it through graduate school, but to succeed and to continue to strive for more.

To my committee members, Larry Boise, Andrew, Kowalczyk, Paula Vertino, and Wei Zhou, for their endless insights and helpful suggestions. Unlike a lot of graduate students, I was able to enjoy my committee meetings, not just for their intellectual discussions from which I always learned something, but also all the laughs. I felt lucky to have such an intelligent yet supportive group in my corner over the past several years. A special acknowledgement goes to Larry for taking a chance on me when I applied to the program. Even though I missed the application deadline and reached out to him with little hope, Larry brought me to Emory so I owe him thanks for giving me that opportunity.

To my labmates, for your support, help, friendship, comic relief, and so much more. I'll always remember our silly laughs, getting late-night lab credit, the intense games of paper football, and trying to prank Adam (and usually failing!). You were amazing people to share the last several years with,

and your hard work and dedication pushed me to work harder every single day. I wish you all the best of luck in the future, though I know that none of you will need it. I hope that my next lab has people with half as much kindness, humor, and overall character as you do.

Last but certainly not least, I want to acknowledge those that I lost to cancer while pursuing my degree. While I will never forget your daily struggles to fight this disease, I will choose to remember the smiles that remained in spite of those physical and mental struggles. Thank you for all you gave me with your friendship and love, and in return, I will do my best to help others afflicted with cancer to lead better and longer lives.

Table of Contents

List of Figures

List of Abbreviations

In alphabetical order:

AMPK – Adenine monophosphate-activated protein kinase

AMT – Amoeboid to mesenchymal transition

CAF – Cancer-associated fibroblast

CTD – C-terminal domain

DMSO – Dimethyl sulfoxide

ECM – Extracellular matrix

EGFR – Epidermal growth factor receptor

EMT – Epithelial to mesenchymal transition

FAK – Focal adhesion kinase

GEMM – Genetically engineered mouse model

GFP – Green fluorescent protein

LKB1 – Liver kinase B1

MAT – Mesenchymal to amoeboid transition

LOX – Lysyl oxidase

NEDD9 - Neural precursor cell expressed developmentally down-regulated protein 9

NSCLC – Non-small cell lung carcinoma

NTD – N-terminal domain

PAK1 – p21-activated kinase 1

PJS – Peutz-Jeghers Syndrome

SCLC – Small cell lung carcinoma

TAM – Tumor associated macrophage

TCGA – The Cancer Genome Atlas

TKI – tyrosine kinase inhibitor

VEGF – Vascular endothelial growth factor

VEGFR – Vascular endothelial growth factor receptor

WT – Wildtype

Chapter 1. Introduction

1.1 Lung cancer

1.1.1 Lung cancer overview

Lung cancer is the leading cause of cancer-related deaths, killing more people than breast, prostate, and colon cancers combined (Figure 1.1) (1). There are two main histological subtypes of lung cancer, small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC), which can be further divided into either squamous cell, large cell, or adenocarcinoma. NSCLC accounts for the majority of lung cancer cases in the U.S. (approximately 85% of all lung cancer cases) and has a poor overall 5-year survival rate of just 15%. However, this 5-year survival rate varies greatly based upon how extensive the disease progression is at the time of diagnosis. For localized disease, the 5-year survival rate is ~55% and for regional disease, ~28%. However, this rate drops dramatically to only 4% for distant-metastatic disease (2). Additionally, the percentage of cases diagnosed before the cancer metastasizes is only 15%; thus, the majority of patients present with late-stage disease that is more aggressive and harder to treat, and patient prognosis is dramatically decreased in comparison to those with early-stage disease. Further study of the metastatic process would benefit these patients with progressive disease that currently have a bleak disease outlook by identifying novel targets to stop the tumor from metastasizing and instead create a chronic, manageable disease.

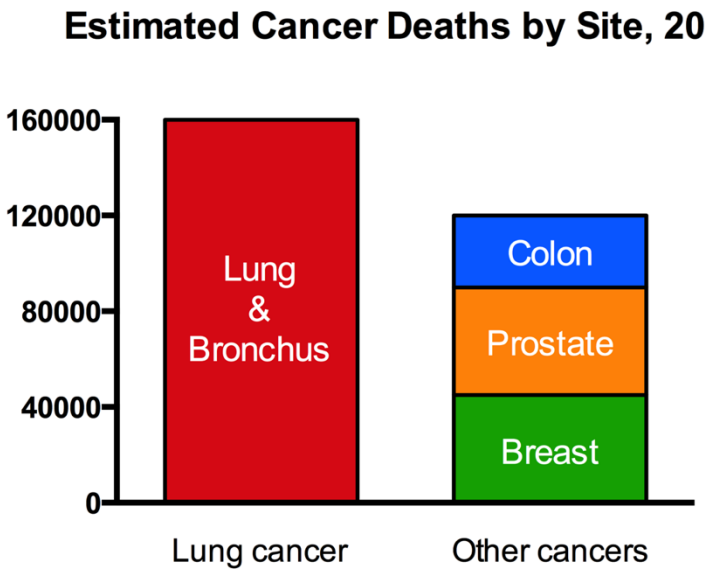

Figure 1.1. Estimated cancer deaths in the U.S. in 2015. Adapted from American Cancer Society (ACS), *Cancer Facts and Figures 2015* (1). Lung cancer was estimated by the ACS to kill more patients in 2015 than the 3 most common cancer types (breast, prostate, and colon cancers) combined. This figure shows the importance of continued study on lung cancer progression and metastasis.

1.1.2 Common lung cancer mutations

Efforts to identify driver mutations in cancer, including The Cancer Genome Atlas (TCGA), have increased our knowledge of the mutational landscape of complex tumors in hopes of creating targeting therapies for patients harboring those mutations. Lung cancer in particular has been well studied and several key driver mutations identified. For example, mutations in *EGFR* are found frequently in lung adenocarcinomas, especially in patients of Asian decent (3). Identification of this oncogene addiction led to the develop of several tyrosine kinase inhibitors (TKIs) such as gefitinib and erlotinib; however, acquired resistance to these agents has limited their efficacy in patients likely through outgrowth of resistant clones (4). Other oncogenic drivers also occur frequently in lung cancer patients, such as *KRAS* or *ALK*, and these mutations, along with EGFR mutations, are generally mutually exclusive (5). Several known tumor suppressors are also commonly disrupted in lung cancer samples, including *TP53, CDKN2A,* and *STK11*, also known as LKB1. *STK11* is the third most commonly mutated gene in lung adenocarcinomas, with about 30% of patients harboring mutations (6, 7). The study of LKB1 in cancer progression, invasion, and metastasis has revealed it to be an important regulator of tumor development and metastasis particularly in lung adenocarcinomas.

1.1.3 Heterogeneity in lung tumors

While the concept of tumor heterogeneity has been around for decades, only recently have we been able to properly understand the breadth of alterations that occur within a single patient with improved sequencing technologies. Multiregion exome and whole genome sequencing efforts revealed vast spatial heterogeneity of genomic alterations (8, 9). While known lung cancer drivers that are therapeutically targetable were present in all tumor regions, putative

driver mutations were generally only present in one or two regions. Sequencing analysis of these regions revealed that these mutations were in fact clonal throughout the region, evidence that they were driver mutations for that subclonal population. Additionally, mutations in metastasized cells often correlated with mutations found in just one particular region of the primary tumor, and a similar discovery found that patients with relapsed disease had larger subclonal fractions within the primary tumor (10). Thus, these studies highlight the possibility that minor subclones within the primary tumor can dictate disease progression and clinical outcome. Taken together, large-scale efforts such as these provide greater insight into the intratumor heterogeneity that patients present with in the clinic. Identification of known lung cancer drivers, such as Kras and EGFR, between patients is the current paradigm for treatment strategy; however, these patients may harbor copious other tumor subpopulations with a different mutational landscape that will ultimately cause treatment to fail and the disease to progress.

1.2 Tumor metastasis overview

Metastasis is a clinically relevant process, as it accounts for 90% of all cancer-related deaths (11, 12). Despite this fact and large research efforts, little is understood about the determinants of metastasis in patients. The metastatic process contains several steps that tumor cells must undergo in order to successfully seed a distant organ site (Figure 1.2). While the majority of patients present with metastatic disease that ultimately kills the patient, the process of metastasis itself is a highly inefficient one, with estimates of only ~0.01% of metastatic cells that enter the blood stream successfully developing into macrometastases (13, 14). This is likely due to the need for the conditions within the primary tumor, the vasculature, and the secondary organ sites to all be favorable in promoting tumor cell growth and survival.

1.2.1 The steps of the metastatic cascade

The first step of the metastatic cascade is expansive growth of the primary tumor, which causes a lack of nutrients and low oxygen available to the cancer cells (14, 15). This lack of oxygen, called hypoxia, causes an upregulation of the production and secretion of vascular endothelial growth factor (VEGF), to which quiescent endothelial cells respond, migrate towards, and begin nascent vascular formation, or neoangiogenesis (14, 16). Metastatic clones that arise from within the heterogeneous primary tumor adhere to and invade through the basement membrane into the surrounding normal microenvironment, towards the newly formed vasculature (14-16). This early invasive process is believed to occur through an epithelial to mesenchymal transition (EMT) in which tumor cells lose their epithelial characteristics, such as E-cadherin expression, and gain mesenchymal ones, such as vimentin and N-cadherin (17, 18).

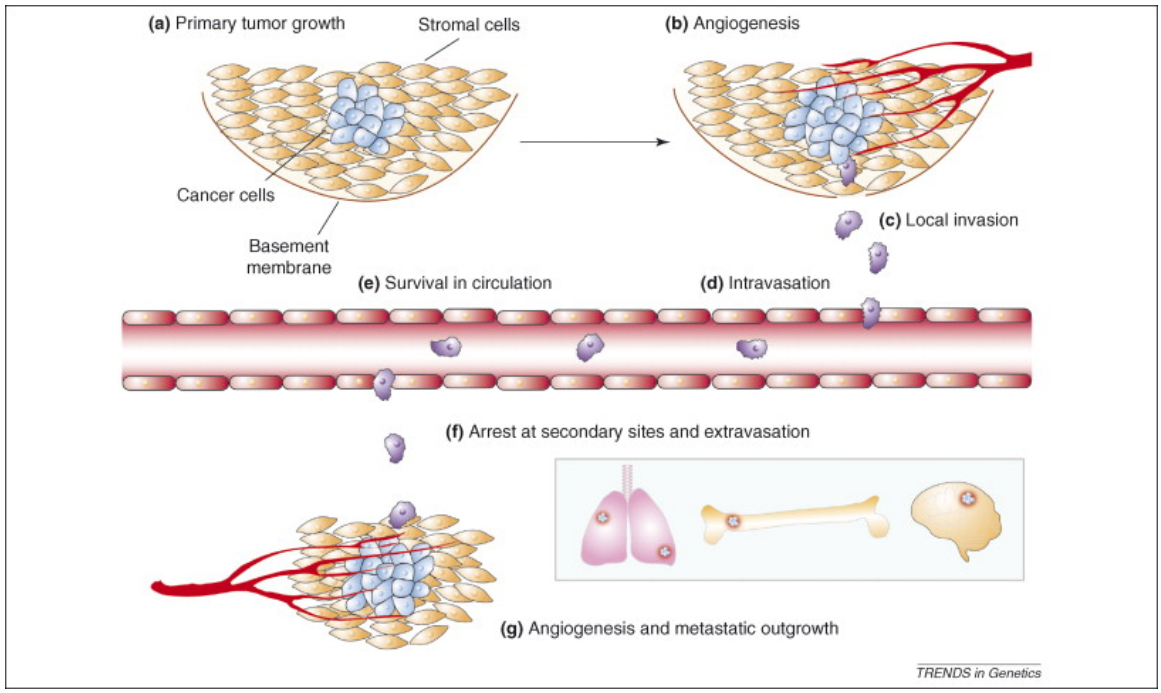

Figure 1.2. The metastatic cascade. Taken from (15). Metastasis is a complex process that involves several steps. Once the primary tumor forms and invasive subclones are activated, the cells must break through the basement membrane that contains the tumor and locally invade into the normal microenvironment. Neovasculature that forms around the tumor provides an escape mechanism to these invasive cells, which intravasate the vasculature and enter the bloodstream. There, the tumor cells interact with platelets to form a protective embolus. The metastatic cells must then adhere to and extravasate the vasculature at a secondary organ site, where they then deposit and grow into macrometastases.

This transition is characterized by epithelial cells losing their apical-basal polarity and subsequently cell-cell contacts, and acquiring a more motile phenotype. While well studied, research in the area of EMT is highly debated, including whether this process occurs in patients and if it is a required step in the metastatic cascade. Recent work has supported the concept of a partial EMT, in which cells maintain some epithelial characteristics during invasion (18-20). Upregulation and secretion of matrix metalloproteinases (MMPs) by the invasive tumor cells is another essential component of early invasion, allowing these cells to breakdown and move through the basement membrane and extracellular matrix (21-23).

Once the metastatic clones have reached the vasculature, they must then intravasate. Because the newly formed vasculature are often highly unorganized and leaky (24), the migrating tumor cells are able to maneuver through loosely connected endothelial cells, and this process can be mediated by cytokines released by tumor associated macrophages or other stromal cells (14, 25). Once in the blood stream, the tumor cells must survive the circulation, which involves interactions with platelets to avoid shear forces (26) and activation of programs to avoid detachment-induced cell death, termed anoikis (27, 28). Once the circulating tumor cells arrest in a distal site, they must extravasate the vasculature and invade into the microenvironment of the secondary organ. This process of extravasation differs from intravasation as the microvessels at the secondary site are likely to be fully functional, unlike the leaky vasculature found near the primary tumor. Thus, tumor cells may rely on MMPs and other factors such as angiopoientin-like-4 and VEGF to disrupt endothelial cell contacts (29-31). These disseminated tumor cells may or may not go on to proliferate and grow into macrometastases. In fact, the conditions within the new metastatic niche can dictate the success of a micrometastasis, and site-specific factors impact outgrowth in various organs (14). However,

if conditions in the secondary site are favorable to the tumor cells, either by pre-metastatic priming through tumor cell release of systemic factors that alter a particular niche or growth factors released by stromal cells, these tumor cells will proliferate and grow into macrometastasis.

1.2.2 Factors that influence metastasis

While a large research effort exists to study each of these important steps in the metastatic cascade, the process is a complicated one and large questions still remain in the field. For example, some work highlights a stepwise progression, from contained primary tumor to the development of a metastatic phenotype, whereas others argue that metastatic cells can shed from the primary tumor even before detectable disease (32). Even less is understood about how the tumor microenvironment contributes to disease progression in patients, though studies have begun to elucidate the complex network of contributing factors.

One important aspect is the interaction between tumor cells and the surrounding physical environment, or the extracellular matrix (ECM). The property of the ECM surrounding a tumor can either promote or hinder progression and invasion. For example, a rigid, organized ECM provides a physical barrier, or can sequester necessary growth factors away from the tumor cells (33). Alternatively, aberrant ECM organization or content has been shown to promote invasion and metastasis by providing tracks of migration or a biomechanical force that activates motility programs within the tumor cells. The ECM can also determine the fate of a migratory cancer cell within a secondary organ post-metastasis by promoting the transition from tumor cell dormancy to metastatic growth (34).

Another factor that can drastically influence the metastatic phenotype of a tumor is the involvement of stromal factors, such as heterotypic interactions with cells surrounding the tumor. For example, cancer associated fibroblasts (CAFs) and macrophages (TAMs) are two populations of stromal cells that have vast effects on tumor cell progression, invasion, and metastasis. For example, CAFs deposit collagen and other ECM proteins (35, 36). Tumor cells can react to collagen deposition and remodel it in order to create pathways in the microenvironment (36). Additionally, as discussed earlier, TAMs can influence the ability of tumor cells to intravaste the neovasculature around the primary tumor (25). In general, CAFs and TAMs can promote invasion indirectly via secretion of growth factors and chemokines (14), as well as take a more active role by participating in heterotypic cell-cell streaming invasive strands (37, 38). Additionally, these factors all play important roles at the distal secondary site, priming the metastatic niche for successful metastasis implantation and survival.

Besides extrinsic factors, the intrinsic features of the tumor cells can impact the metastatic potential of the tumor. Many research studies have attempted to identify novel metastasis suppressors or initiators. While more work is warranted, several possible candidate genes have been identified. For example, caspase-8 was found to be necessary for the survival of neuroblastoma cells particularly those at the tumor-ECM interface, and suppression of caspase-8 was sufficient to inhibit metastasis *in vivo* (39). Other potential metastasis modulating genes are NM23 (40) and KAI1 (41, 42). Research specifically in lung adenocarcinoma reveals the tumor suppressor gene *LKB1* as an important metastatic regulator. *In vivo* studies demonstrate that loss of the LKB1 gene potentiates the metastatic ability of a Kras driven model of lung cancer (43), thus identifying LKB1 as a potential metastasis suppressor.

However, more research is needed, particularly in lung cancer, to identify metastasis-initiating cells and their genomic pattern that potentiates their metastatic capabilities.

1.2 LKB1 background

1.3.1 LKB1 in Peutz-Jeghers Syndrome and cancer

LKB1 is a serine/threonine kinase that was originally identified as a tumor suppressor regulating Peutz-Jeghers Syndrome (PJS), an autosomal dominant disorder. This disease is characterized by development of benign gastrointestinal polyps and an increased risk of developing sporadic intestinal tumors (44). LKB1 germline inactivation occurs in 66-94% of all patients diagnosed with PJS depending on the screening method (45, 46), and somatic inactivation of the unaffected allele is often observed in the polyps and tumors that develop from patients with PJS. After discovering the vital role of LKB1 in PJS, LKB1 was then studied extensively across many different tumor types. Surprisingly, the loss of LKB1 was found to play a role in just a select few cancer types, including melanoma, cervical cancer, and lung cancer (7, 47-49). In lung adenocarcinomas, about 30% of patients are affected by LKB1 mutations (6), and these mutations are generally truncating mutations that cause partial or complete loss of two known functional domains of LKB1 (50). LKB1 has a central kinase domain flanked by a short N-terminal domain and a long C-terminal domain containing a functional farnesylation motif, important for LKB1 insertion into the plasma membrane (51). Thus, the majority of patients with LKB1 truncating mutations have disrupted LKB1 kinase and CTD functions, so determining the independent functions of each of these domains should provide insights into the cellular defects in lung tumors with mutations in LKB1.

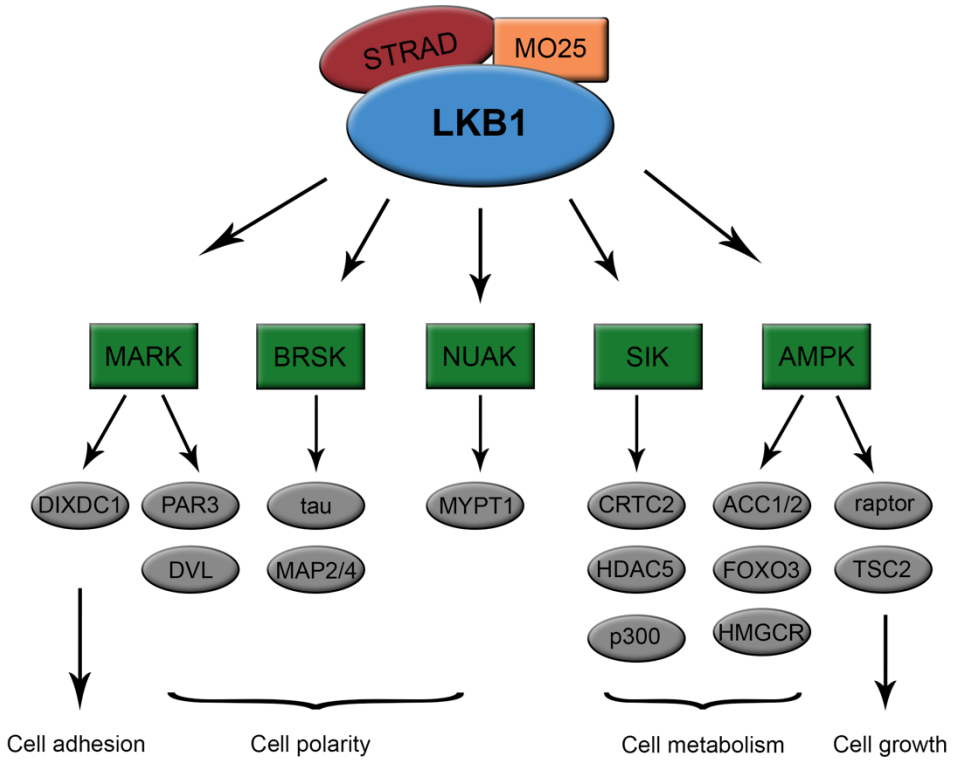

Figure 1.3. Cellular processes regulated by LKB1 kinase activity. Adapted from (52) and (53). LKB1, in complex with regulatory units STRAD and MO25, phosphorylates and activates a family of 14 AMPK-related protein kinases. Several downstream effectors of each kinase are shown that are involved in regulating cell adhesion, polarity, metabolism and growth.

1.3.2 Cellular processes regulated by LKB1

LKB1 regulates a variety of cellular processes by acting upstream of numerous effector proteins. In response to cellular stress, LKB1 is activated and exported from the nucleus with its cofactors STRAD and MO25, and then phosphorylates adenine monophosphate-activated protein kinase (AMPK) (54, 55). AMPK in turn phosphorylates several key players in protein and cholesterol synthesis, most notably TSC2 to regulate the mTOR pathway, thus controlling cellular metabolism and growth in times of nutrient deprivation. Additionally, LKB1 has also been shown to phosphorylate 12 other AMPK family members to impact various downstream cellular processes (56-58), including cell polarity (Figure 1.3). The role of LKB1 in regulating cell polarity has been well studied in epithelial cells as well as in cancer cells during migration. In particular, LKB1 was found to regulate cell apical-basal epithelial polarity in a cell-autonomous manner in intestinal and pancreatic cells, and this was independent of its phosphorylation of AMP-kinase (59). In cancer cells, LKB1 regulates directional migration through the Rho GTPase cdc42 via recruitment of the effector kinase p21-activated kinase 1 (PAK1) (60). LKB1 also regulates directional migration in a 3-D microenvironment via sensing of extracellular matrix (ECM) gradients, also known as haptotaxis (61). In addition to its role in regulating cellular metabolism and polarity, another important function of LKB1 is to regulate adhesion dynamics during cell motility. In particular, LKB1 was found to impact cell adhesion via regulation of focal adhesion kinase (FAK) and neural precursor cell expressed developmentally down-regulated protein 9 (NEDD9) (62-64), while also impacting ECM structure via lysyl oxidase (LOX) (65). Little is known about how LKB1 regulates these polarity and adhesion functions during 3-D invasion, nor which domain of LKB1 seems to be important in its regulation of cell polarity and directional migration.

1.3.3 LKB1 in lung cancer invasion and metastasis

Given the normal role of LKB1 in controlling cell polarity and directional migration, the question remains as to how its loss of function contribute to cancer cell invasion and metastasis. Several research studies implicate LKB1 as an important tumor suppressor gene and evidence suggests that it is a regulator of metastatic potential in tumor cells. In a clinically relevant mouse model, LKB1 loss with concurrent Kras activation results in increased tumor progression and increased metastasis (43). When compared to other common tumor suppressors such as p53 and pTEN that are also inactivated in lung adenocarcinomas, Kras activation with LKB1 inactivation resulted in increased metastasis. This highlights LKB1 as an important modulator of tumor progression and metastasis *in vivo*. Further research suggests that this may be due to LKB1 inhibition of the epithelial to mesenchymal transition (EMT), an important early step in the metastatic cascade. In lung cancer cells, LKB1 regulates EMT through attenuation of ZEB1 transcription factor and when LKB1 is lost or inactivated, these cells acquire a more mesenchymal and motile phenotype (66, 67). However, further research is necessary to understand if these events occur in patients as well, or if LKB1 controls other pathways that could contribute to tumor metastasis.

.

1.4 Phenotypic heterogeneity during invasion

Tumor cells show phenotypic heterogeneity during the initial steps of the metastatic cascade, especially during the process of local invasion through the tumor microenvironment. This heterogeneity ranges from single cell migration to cell-cell streaming to larger collective invasion packs (Figure 1.3) (38, 68-70). Plasticity between each of these phenotypes has also been observed, with invading tumor cells beginning as one morphology but undergoing a series of molecular changes in response to variation within the microenvironment, that result in alteration in their invasive morphology (69, 71-74). While there are studies supporting this ability of invasive tumor cells to transition phenotype *in vitro*, it is still unclear whether these same adaptations occur in patients. Each mode of invasive phenotype is described in detail below.

1.4.1 Single cell invasion

During single cell invasion, cells do not maintain any of their cell-cell contacts, and they can adapt one of two different cell morphologies. Amoeboid cell invasion refers to a rounded cell type that relies on RhoA GTPase signaling to ROCK1 to promote actomyosin contractility, creating membrane blebs that create propulsion force to propel the cell through the extracellular matrix (71, 73, 75-78). These cells do not generally form cell-matrix interactions through canonical adhesion sites and do not perform proteolysis of the matrix; instead, amoeboid cells generally deform the matrix during invasion (75). These characteristics ultimately lead to an increased velocity during movement through the microenvironment as amoeboid cells can squeeze their cell bodies through the matrix fibers instead of mechanically altering them, as is seen with mesenchymal cell invasion (73). Mesenchymal cells are generally

elongated cells that utilize cdc42/Rac1 GTPase signaling to PAK1 to promote actin dynamics at the cellular leading edge (76, 78). Focal adhesions develop at the leading edge, thus resulting in cell-ECM interactions while also promoting protease-dependent matrix degradation to reorganize the microenvironment during migration (38, 69). Because mesenchymal cells are path-generating via proteolysis, they are generally slower moving as compared to amoeboid cells. Transitioning between mesenchymal and amoeboid cell shapes has been observed *in vitro* and is termed MAT (mesenchymal to amoeboid transition) or AMT (amoeboid to mesenchymal transition), and this is generally dependent on differing capabilities of proteolytic cleavage of the ECM.

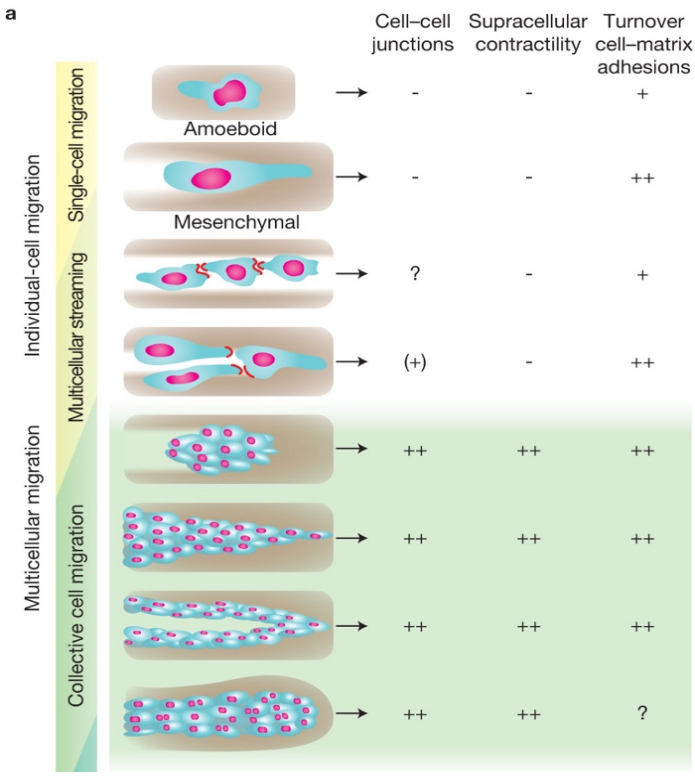

Figure 1.4. Phenotypic heterogeneity in invasive morphologies. Adapted from Friedl et al., 2012 (68). Invasion away from the primary tumor and into the surrounding microenvironment is one of the initial steps in the metastatic cascade. Studies into this process have revealed a variety of invasive phenotypes exist *in vitro* and *in vivo*, and each mode of invasion has its own defining characteristics as well as molecular mechanisms. Single cell invasion refers to cells that have no cell-cell contacts and can vary in morphology from a rounded cell type (amoeboid) to an elongated cell type (mesenchymal). Collective cell invasion also has a variety of modes, from cell clusters to invasive strands, but are all characterized by maintenance of cell-cell contacts.

1.4.2 Collective cell invasion

Collective invasion refers to a cohort of cells that act as a single unit during migration through the ECM. This type of invasion is clinically relevant as patient tissue sections often provide evidence of groups of tumor cells broken from the primary tumor mass, yet invading together within the normal microenvironment (68). Thus, study of the processes that regulate collective invasion could have vast implications across many solid tumor types.

Collective invasion can also be further divided into different categories. Some do not rely on direct cell-cell contacts, such cellular streaming. Here, cells do not directly adhere to their neighbors but instead follow paths and chemoattractant gradients created within the microenvironment by either other tumor cells (homotypic cell streaming) (68) or tumor associated stromal cells such as fibroblasts (heterotypic cell streaming) (37, 38). Other types of collective invasion do rely heavily on cell-cell interactions, such as sheet-like invasion reminiscent of normal wound healing, tumor cell clusters, and collective strands or chains (68). Collective chain invasion has a leading invasive cell, called a tip or leader cell, which has asymmetrical cellular polarity in which cell-cell contacts are lost at the front of the cell but are maintained at the cell rear (68, 79). The cells that are directly attached to the back of the leader are termed follower cells. Recent studies in breast cancer have identified that leader cells are a transient phenotype regulated either by a basal cell type program demarked by cytokeratin-14 expression (80) or by a 7-gene signature including *DOCK10* expression that also correlated with low EpCAM expression (81). However, it is still unclear if leader cells in these models are a transient cell type that shows plasticity upon certain molecular or environmental changes, or whether the leader cell phenotype is hard-wired in the genome of these cells. Additionally,

little research has focused on the phenotype or genomic profile of leader cells in the context of lung cancer collective invasion.

1.5 Dissertation goals

Together, the work presented in this dissertation aims to dissect the mechanisms regulating various modes of lung cancer invasion as well as the cell-cell communication that occurs among invasive cells. When this dissertation work began, cancer invasion and metastasis was named as one of the hallmarks of cancer, and most cancer-related deaths were known to be caused by the metastatic process. Research in this field is varied and extensive, with studies aiming to understand each individual step of metastasis and ways to block it. However, even with the enormous effort to study and understand metastasis, not much progress had been made as far as how to target it in order to increase patient survival. The first part of this dissertation has a focus on the role of LKB1 in lung cancer invasion. While much was already known about LKB1 regulation of normal epithelial cells as well as its impact on cancer cell adhesion through mediation of FAK signaling, little was known about how each domain of LKB1 regulates cell polarity and adhesion within a physiologically relevant 3-D microenvironment. Chapter 2 aims to address this by examining more specifically the role of LKB1 in regulating lung cancer invasion by performing 3-D spheroid invasion assays to analyze the independent functions of both the kinase and C-terminal domains of LKB1 in controlling various aspects of invasion. For the second part of this dissertation, the concept of collective migration was well researched in terms of development and wound closure. In cancer research, many are recognizing the value of learning about collective invasion as it is a clinically relevant process and often observed as a major form of invasion in solid tumor samples. However, the question still remains on how leader cells arise and whether they are a distinct, targetable population. The work presented in Chapter 3 focuses on how cancer cells within the collective invasion pack communicate and cooperate for successful escape. The

chapter describes the development of a novel image-guided genomics technique used to probe phenotypic heterogeneity, in order to better address the mechanisms regulating collective invasion and to probe the underlying biology of the cells within the invasion pack.

Chapter 2: LKB1 kinase-dependent and -independent defects disrupt polarity and adhesion signaling to drive collagen remodeling during invasion

Jessica Konen*[†,‡], Scott Wilkinson*[†,‡], Byoungkoo Lee[§], Haian Fu[†,||], Wei Zhou[†], Yi Jiang[§], Adam I. Marcus[†,¶]

* These authors contributed equally to this work

Affiliations:

[†] Winship Cancer Institute of Emory University, Department of Hematology and Medical Oncology, Atlanta, GA 30322

[‡] Graduate Program in Cancer Biology, Emory University, Atlanta GA, 30322

[§] Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30302

[||] Department of Pharmacology, Emory University, Atlanta, GA 30322

This work was adapted from the work published in Molecular Biology of the Cell. 2016 Apr 1;27(7):1069-84. doi: 10.1091/mbc.E15-08-0569.

Contributions:

The work presented here is the result of a co-first authorship with SW. Each first author contributed equally to the work. JK performed experiments in Figures 2.1, 2.3, 2.5, 2.6-2.8, as well as contributed to the writing and editing of the manuscript.

Abstract

LKB1 is a serine/threonine kinase and a commonly mutated gene in lung adenocarcinoma. The majority of LKB1 mutations are truncations that disrupt its kinase activity and remove its C-terminal domain (CTD). Since LKB1 inactivation drives cancer metastasis in mice and leads to aberrant cell invasion *in vitro*, we sought to determine how compromised LKB1 function impacts lung cancer cell polarity and invasion. Using 3-D models, we show that LKB1 kinase activity is essential for focal adhesion kinase-mediated cell adhesion and subsequent collagen remodeling, but not cell polarity. Instead, cell polarity is overseen by the kinase-independent function of its CTD, and more specifically its farnesylation. This occurs through a mesenchymal-amoeboid morphological switch that signals through the Rho-GTPase RhoA. These data suggest that a combination of kinase-dependent and -independent defects by LKB1 inactivation create a uniquely invasive cell with aberrant polarity and adhesion signaling that drives invasion into the microenvironment.

2.1 Introduction

Liver kinase B1 (LKB1; also known as STK11) is a serine/threonine kinase that was identified as a tumor suppressor in the inherited autosomal-dominant disorder, Peutz-Jeghers Syndrome (PJS). PJS patients have LKB1 loss of heterozygosity resulting in gastrointestinal polyposis and a greater likelihood of developing sporadic tumors in the breast, gastrointestinal tract, and pancreas (44). Somatic inactivation of LKB1 is observed in several cancer types including melanoma, lung, and cervical cancers (7, 47-49). In lung adenocarcinoma, *LKB1* is the third most commonly mutated gene behind *KRAS* and *P53* (6, 50), though how *LKB1* mutations drive lung adenocarcinoma progression remains an area of intense interest.

LKB1 missense and truncating mutations in lung adenocarcinoma primarily occur within its central kinase domain (50). LKB1 kinase activity was first linked to the canonical AMPK energy stress response pathway, where it serves as the upstream kinase of AMPK (5' AMP-activated protein kinase) (54, 55). LKB1 also phosphorylates and activates 14 members of the AMPK family (56-58), including the microtubule affinity-regulating kinases (MARK1-4) (58, 82, 83), NUAK1/2 (84, 85) to control myosin contractility, SIK1 to oversee anoikis (86), and the BRSKs to promote axon differentiation (87, 88).

The LKB1 kinase domain is flanked by a short N-terminal domain and a longer C-terminal domain (CTD) containing a membrane-targeting farnesylation motif (51). Predictably, LKB1 truncating mutations within its kinase domain disrupt kinase function and result in a truncated protein lacking the CTD, or a fully degraded transcript with complete protein loss. This observation raises the intriguing question of how loss of kinase activity compared to loss of

the CTD, impacts LKB1 function and cancer progression. One possible answer could be linked to its role in cell polarity; LKB1 serves as a master regulator of cell polarity across multiple species (53, 89). In *C. elegans*, the LKB1 ortholog, PAR-4, is critical for establishing and maintaining an anterior-posterior (AP) axis during cell division (90). Additionally, in *Drosophila*, LKB1 is essential for establishing the AP axis during oogenesis and for promoting an apical-basal polarity in eye and follicular cells (91, 92). In single human intestinal epithelial cells, LKB1 re-expression leads to a fully polarized cell, even in the absence of cell-cell contacts (59). LKB1 loss in epithelial cells also disrupts apical-basal polarity and basement membrane integrity, while promoting an epithelial to mesenchymal transition (EMT) (66, 93). The role of LKB1 in regulating cell polarity and motility has linked LKB1 to the Rho family of small GTPases (94, 95). Upon activation of cell motility in lung cancer cell lines, LKB1 rapidly translocates to the cellular leading edge where it binds to actin and associates with the small Rho-GTPase, cdc42 (60). Additionally, in normal bronchial epithelial cells LKB1 coordinates with p114RhoGEF to regulate RhoA activity and maintain apical junctions (96, 97).

Complete LKB1 loss in a $Kras^{G12D}$-driven genetically engineered mouse model (GEMM) of lung cancer led to increased tumor burden, shortened survival time, and increased metastasis compared to $Kras^{G12D}$ mutant mice (43). Despite these recent insights into LKB1 function, how LKB1 coordinates its kinase-dependent and -independent functions to regulate cell invasion remains poorly understood. Therefore, we sought to uncouple defects in CTD function from defects in kinase function during cancer cell invasion. Our data show that the combination of defects in LKB1 kinase-dependent and -independent function creates a uniquely invasive cell that is unable to properly polarize and maintains an amoeboid shape; however, unlike classical amoeboid cells, LKB1 compromised cells maintain high FAK activity

and still remodel collagen during 3-D invasion. Our studies show that FAK-driven cell adhesion and collagen remodeling are caused by defective kinase activity, whereas amoeboid cell shape occurs due to RhoA signaling defects caused by a lack of LKB1 CTD farnesylation. Given that LKB1 frequently undergoes truncating mutations in lung adenocarcinomas that predictably affect both farnesylation and kinase activity, this highlights the importance of this potent combination of defects that could impact cancer cell metastasis.

2.2 Materials and Methods

Cell culture and generating stable cells:

H1299, H1792 and H157 human NSCLC cells (ATCC, Manassas, VA) were cultured in Roswell Park Memorial Institute (RPMI-1640) media supplemented with 10% fetal bovine serum and 100 units/mL of penicillin/streptomycin, and maintained at 37°C and 5% CO_2. HeLa human cervical cancer cells (ATCC, Manassas, VA) were cultured in Dulbecco's Minimum Essential media supplemented with 10% fetal bovine serum and 100 units/mL of penicillin/streptomycin, and maintained at 37°C and 5% CO_2. Stable pLKO.1 vector control and LKB1-shRNA (shLKB1) H1299 and H1792 cells were created as previously described (64). Briefly, these cells were created by lentiviral infection using specific shRNA constructs from Open Biosystems (Rockford, IL). Puromycin (2μg/ml, Sigma-Aldrich, St. Louis, MO) was used to select transduced cells and western analysis used to confirm knockdown.

To generate H157 and HeLa cells stably expressing LKB1 and constitutively active RhoA or cdc42, Wildtype LKB1 and the various LKB1 domains and mutations were cloned into a pEGFP-C1 vector. Empty GFP or the GFP-LKB1 constructs were then subcloned from the pEGFP-C1 vector into a pBabe-puro vector. Constitutively active RhoA (Q63L) and cdc42 (Q61L) were subcloned from a pCDNA3 vector into pBabe-Hygro. The pBabe constructs were then transfected into Phoenix-ampho cells with Lipofectamine 2000 and PLUS reagent (Invitrogen, Grand Island, NY). Cells expressing only empty GFP or GFP-LKB1 were selected with 1 μg/ml puromycin, while cells co-expressing the constitutively active RhoA or cdc42 mutants were selected with 1 μg/ml puromycin and 300 μg/ml hygromycin (EMD Millipore, Billerica, MA). Proper expression of GFP-LKB1 was verified using IF and Western

blot to confirm phenotype and molecular weight. Expression of constitutively active RhoA and cdc42 was confirmed using a Rho-GTPase activity assay comparing the constitutively active mutants to their isogenic partner lines.

Antibodies and stains:

Antibodies were used against pFAK^{Y397}, pPaxillin^{Y118}, MARK1, GAPDH, GFP (Cell Signaling, Boston, MA), LKB1 (Santa Cruz, Santa Cruz, CA), RhoA, and cdc42 (Cytoskeleton Inc, Denver, CO) for 3-D immunofluorescence, western blotting, and immunoprecipitation. Alexa Fluor® 555 Phalloidin, Alexa Fluor® 488, and Alexa Fluor® 555 goat-anti-rabbit (Invitrogen) and DAPI (4',6-diamidino-2-phenylindole, Invitrogen) were used for 3-D immunofluorescence. Horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson ImmunoResearch, West Grove, PA) were used for western blotting.

Transient transfection and Western blot:

For LKB1 rescue experiments, H157 cells were transfected with either pCDNA3 empty GFP control vector or pEGFP-C1 LKB1 WT using Lipofectamine 2000 (Invitrogen), per manufacturer's protocol. For silencing experiments, H1299 or H1792 cells were transfected with either a scrambled control siRNA or siRNA targeting MARK1 (Sigma) or LKB1 (Sigma) using Oligofectamine (Invitrogen), or with siRNA targeting FAK (Sigma) using Lipofectamine 2000 (Invitrogen), per manufacturer's protocol. 24 hours later, cells were harvested and lysed with TNES buffer (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA, 1% NP-40) supplemented with Roche Complete Protease Inhibitor Cocktail (Roche, Indianapolis, IN) and a phosphatase inhibitor cocktail (10 mM Sodium Fluoride, 2 mM Sodium Pyrophosphate, 2 mM β-Glycerophosphate, 200 nM Sodium Orthovanadate). Protein concentrations were

determined using the bicinchoninic acid protein (BCA) assay kit (Pierce, Rockford, IL). 30 μg of lysates were boiled in Laemmli sample buffer, loaded onto 10% SDS-PAGE gels, transferred onto polyvinyldifluoride (PVDF) membranes, blocked in 10% milk for 1 hour at room temperature, probed overnight at 4°C with primary antibodies diluted in either 5% BSA or non-fat dried milk, followed by appropriate horseradish peroxidase-conjugated secondary antibody and visualized using HyGlo Chemiluminescent HRP antibody detection (Denville, South Plainfield, NJ).

Generating 3-D tumor spheroids:

H1299 and H1792 cells with a stable or transient LKB1, MARK1, or FAK knockdown and corresponding controls, as well as H157 and HeLa cells stably expressing GFP-LKB1 or GFP-LKB1+constitutively active RhoA or cdc42, were grown to 70% confluency and then trypsinized, neutralized, and resuspended in complete RPMI (Invitrogen). To generate spheroids, 3000 cells in 200 μl (1.5×10^4 cells/ml) were added to a Spheron Nunclon 96 well plate (Thermo Scientific, Waltham, MA). 3 to 4 days later, compacted spheroids were collected and resuspended in 2.0 mg/ml Collagen Type I (Advanced Biomatrix, San Diego, CA), then plated in: a Lab-Tek 8 well borosilicate bottom plate (Thermo Scientific) for immunofluorescence; a 35 mm glass bottom dish (In Vitro Scientific, Sunnyvale, CA) for multiphoton microscopy; or a 6-well plate (Corning Inc, Corning, NY) for Rho-GTPase activation assays. After the collagen solidified, complete RPMI was added to the top of the collagen matrix to provide a chemogradient for the spheroids.

Drug treatments:

Spheroids formed from H1299 shLKB1 or pLKO.1 control cells were generated as described above. While embedding spheroids in the 3-D collagen matrix, the FAK inhibitor PF-573228 (Sigma-Aldrich) was added at 1 μM or GM-6001 (Santa Cruz) was added at 20 μM to both the collagen and complete RPMI media on top of the collagen matrix. Dimethyl sulfoxide (DMSO) was used as the vehicle control. Spheroids were incubated at 37°C and 5% CO_2 for 16-20 hours to allow for invasion.

Rho-GTPase activation assays:

Rho-GTPase activation assays were performed using H1299 and H157 cells, as previously described (60, 98). In brief, H1299 shLKB1 and pLKO.1 control spheroids were embedded in a 3-D collagen matrix (2 mg/ml) as described earlier. At 0, 1, and 5 hours post-embedding, collagen was digested using collagenase (Sigma-Aldrich) at 37°C. Spheroids were centrifuged and supernatant discarded. Remaining cell pellets were lysed using Rho-GTPase activity assay lysis buffer supplemented with 100x protease inhibitor cocktail (Cytoskeleton Inc). H1299 shLKB1 and pLKO.1, as well as H157 cells stably expressing empty GFP control or the various GFP-LKB1 constructs, as well as those co-expressing constitutively active RhoA or cdc42 were grown to 70% confluency. Cells were then trypsinized and 2.0 x 10^6 cells were plated on a 10 cm fibronectin coated plate (40 μg/ml). 24 hours later, cells were lysed using Rho-GTPase activity assay lysis buffer and 100x protease inhibitor (Cytoskeleton Inc). In all cases, total protein quantification was determined using a BCA protein assay kit (Pierce). 300 μg lysate was incubated with either GST-Rhotekin RBD beads (RhoA) or GST-PAK PBD beads (cdc42) for 1 hour at 4°C. Pulldown and input samples (30 μg) were boiled in Laemmli sample buffer, loaded onto either 4-20% gradient (BioRad, Hercules, CA) or 12% SDS-PAGE gels, transferred onto PVDF membranes, blocked in 5% milk for 1 hour at room temperature,

probed overnight at 4°C with either mouse-RhoA (1:250 in TBST) or mouse-cdc42 (1:1000 in 0.1% milk) primary antibodies, followed by goat-anti-mouse horseradish peroxidase-conjugated secondary antibody (1:10,000 in TBST) and visualized using HyGlo Chemiluminescent HRP antibody detection (Denville).

3-D Spheroid Immunofluorescence:

Spheroids generated from either H1299 or H157 stable cells were embedded in a 3-D collagen matrix as previously described. 24 hours later, cells were fixed using 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) for 20 minutes at room temperature and then quenched with 0.1 M glycine in PBS (Sigma-Aldrich) for 10 minutes. Spheroids were then permeabilized with 0.5% Triton-X (Promega, Madison, WI) for 1.5 hours, washed with PBS for 10 minutes, and blocked with 5% Normal Goat Serum (NGS, Jackson ImmunoResearch) for 1.5 hours. H1299 and H157 spheroids were probed with rabbit $pFAK^{Y397}$ (1:200 in PBS with 1% BSA and 1% NGS) overnight at 4°C. Spheroids were then washed with PBS three times for 15 minutes each with vigorous shaking and probed with (H1299) Alexa Fluor® 488 goat-anti-rabbit or (H157) Alexa Fluor® 555 goat-anti-rabbit (1:750 in PBS with 1% NGS) for 1.5 hours with gentle shaking at room temperature. H1299 spheroids were also probed with rabbit $pPaxillin^{Y118}$ (1:200 in PBS with 1% NGS) and Alexa Fluor® 488 goat-anti-rabbit (1:750 in PBS with 1% NGS). After three PBS washes, all spheroids were then stained with 350 nM DAPI for 10 minutes followed by three more PBS washes.

For 40x representative images for cell polarization, H157 and HeLa cells were first fixed with 4% paraformaldehyde then quenched with 0.1 M glycine in PBS. Spheroids were then permeabilized with 0.5% Triton-X for 1.5 hours, washed with PBS for 10 minutes, stained

with Alexa Fluor® 555 Phalloidin (1:100 in PBS) overnight at 4°C, then washed with PBS three times for 15 minutes each with vigorous shaking before imaging.

Microscopy:

Widefield imaging

For cell polarity experiments, still images of H1299, H1792, H157, and HeLa spheroids were acquired at 0 and 24 hours using an Olympus IX51 at 4x (0.13 NA air), 10x (0.30 NA air) and 20x (0.45 NA air) using an Infinity2 CCD camera.

Confocal imaging

To quantify cell meandering and velocity, H1299 spheroids and H157 spheroids re-expressing GFP-LKB1 NTD, NTD-Kinase, Kinase Domain, and Kinase-CTD were imaged using a Perkin Elmer Ultraview spinning disk confocal microscope at 10x (Plan-Neofluar 0.30 NA) mounted onto a Zeiss Axiovert encased at 37°C with 5% CO_2. Transmitted light images were acquired every 10 minutes for 20 hours with 10 um z-stack intervals using a Hamamatsu Orca ER CCD camera with 2X2 binning. H157 spheroids re-expressing Empty GFP control, GFP-LKB1: Wildtype, C430S, K78I, K78I-C430S, CTD, and CTD-C430S were imaged using a Leica TCS SP8 inverted confocal microscope with live cell chamber at 10x (HC Plan Fluotar 0.30 NA), acquiring images every 10 minutes for 24 hours with 5 μm z-stack intervals using a 488 nm argon laser under resonance scanning (8 KHz, 32 averaging). Representative images for H157 cell polarization were acquired using a Leica SP8 inverted confocal microscope at 40x (HP PL APO 1.30 NA oil) using a 514 nm argon laser.

For 3-D immunofluorescence imaging, H1299 spheroids were imaged using the FV1000 inverted confocal mounted on an Olympus IX81 inverted microscope (40x 0.90 NA, Water PlanApo) with 1.3 μm z-stack intervals and sequential scanning (405 nm, 488 nm). H157 spheroids were imaged with the Leica TCS SP8 inverted confocal microscope (40x oil HC PL APO, 1.30 NA) using 1.3 μm z-stack intervals and sequential scanning (405 nm DMOD Flexible, 488 nm argon, 514 nm argon) at 600 hz with 4 averaging.

Multiphoton imaging

Spheroids of H1299 shLKB1 and pLKO.1 or H157 stable cells were dyed using 1μM of Red CellTracker (Invitrogen). The H1299 stable spheroids were imaged at 0, 6, and 24 hours post-invasion, and H157 stable spheroids were imaged at 0 and 24 hours post-invasion, using a standard upright Zeiss Axio Examiner Z1 microscope with 20x water immersion objective (1.0 NA DIC (UV) VIS-IR). The second harmonic generation (SHG) signal was obtained using a bandpass 380-430 nm cube. To image the cells stained with Red CellTracker, a bandpass of 570-610nm cube with a long pass of 550 nm. Images were taken with a Coherent Chameleon Verdi laser at 790 nm wavelength. Z-stack images were taken with a 1μm interval.

Image Analysis:

Cell polarity was calculated using ImageJ/Fiji (NIH, Bethesda, MD), where an invading cell was considered to have a mesenchymal polarity if their length was greater than or equal to 2 times their width (99, 100). Polarity of H1299 shLKB1 was compared to pLKO.1 control, as was polarity of H1299 stable cells in response to FAK inhibition, using the 2-tailed Student's t-test with a p-value of 0.05. Each H157 and HeLa GFP-LKB1 cell line was compared to the respective empty GFP control lines and also to its farnesylation mutant partner (WT vs C430S,

K78I vs K78I-C430S, CTD vs CTD-C430S) using Fisher's exact test with a p-value of 0.05. * ≤0.05; ** ≤0.01; *** ≤0.001; **** ≤0.0001

For H1299 cells, Volocity (Perkin Elmer, Waltham, MA) image analysis software and manual tracking was used to quantify total invasion (FAK inhibitor experiment), cell velocity, and meandering index (displacement/distance) as a means of determining directional persistence. The total number of cells invaded in response to FAK inhibition was compared between H1299 shLKB1 and pLKO.1 control, as was the difference in velocity of mesenchymal and amoeboid cells from H1299 shLKB1 and pLKO.1 control cells upon FAK inhibition. For H157 cells, 30 cells for each condition were tracked using automated tracking through the Spots function with Brownian motion, with a maximum distance of 20 μm and a gap size of 2. Cell velocity and meandering of H1299 shLKB1 was compared to pLKO.1 control using the 2-tailed Student's t-test with a p-value of 0.05. Each H157 GFP-LKB1 cell line was compared to both the H157 empty GFP control line and its respective farnesylation mutant (as described above), while cell lines co-expressing constitutively-active RhoA or cdc42 were compared to their isogenic partner line and empty GFP control cells using the 2-tailed Student's t-test with a p-value of 0.05. * ≤0.05; ** ≤0.01; **** ≤0.0001.

Phospho-FAK and phospho-paxillin levels were analyzed with the vesicle tracking feature in Imaris Cell (Bitplane, South Windsor, CT). Phospho-FAK sites in H1299 pLKO.1 and shLKB1 cells were quantified with quality ranging from 438 to 1201 (16-bit imaging) and the minimum region threshold of 25. In H1299 MARK1 siRNA and siRNA control cells, pFAK sites were quantified with quality ranging from 297 to 1073 (16-bit imaging), and the minimum region threshold of 14.5. H157 cells' pFAK sites were thresholded with the quality ranging

from 54.3 to 255 (8-bit imaging) and the minimum region threshold of 70.12. pPaxillin sites in H1299 pLKO.1 and shLKB1 cells were quantified with thresholding ranging from 333 to 1335 (16-bit imaging) and the minimum region threshold of 33.1. Mean pFAK and pPaxillin intensity and number of sites/cell of H1299 shLKB1 was compared to pLKO.1 control, while each H157 GFP-LKB1 cell line was compared to H157 empty GFP control cells using the 2-tailed Student's t-test with a p-value of 0.05. * ≤0.05; **** ≤0.0001

Quantification of collagen alignment:

A novel local alignment coefficient was used to quantify the heterogeneous alignment patterns. The collagen fibers in microscopy images were extracted using the CT-FIRE (curvelet transform fiber extraction) software (101). All fibers were quantized with a 5-pixel length. For every pixel in X and Y axis of all Z-stack images, the local alignment coefficient was measured (Supplemental figure explains definition of the local alignment coefficient and the choice of optimal measurement parameters). A local alignment value of 0 means that the fiber angular distribution is isotropic with no bias in any orientation, or the number of fibers in a local circular bin is below a threshold and considered too few to count. A local alignment value of 1 means all the fibers are perfectly aligned. Histograms of local alignment coefficients, surface plots, and contour plots were generated. When comparing between different time points, we normalized to the 0 hour measurements, and plotted the difference in histograms to show the changes in the alignment distribution.

2.3 Results

LKB1 loss induces a morphological switch during 3-D invasion to create a unique amoeboid cell population

To probe the role of LKB1 in regulating 3-D invasion, LKB1 was stably depleted in H1299 non-small cell lung cancer cells (LKB1 wildtype) and compared to isogenic parental vector control cell pLKO.1 (Figure 2.1). Stable knockdown of LKB1 resulted in cells switching from a mesenchymal morphology to an amoeboid morphology. In the pLKO.1 cells, 47% of the invasive cells showed an amoeboid morphology, as compared to 73% in the shLKB1 cells. Mesenchymal cells were defined as those cells that had a length greater than two times its width as previously described (99, 100) (Figure 2.1). A similar transition was observed in H1792 cells (LKB1 wild-type NSCLC), with stable LKB1 depletion (Figure 2.1). Similarly, a second LKB1 targeted shRNA in H1299 cells, as well as transient knockdown of LKB1 in H1299s and H1792s using siRNA, also had similar effects, resulting in amoeboid-like morphology during invasion (Supplementary Figure S2.1). To confirm this was an LKB1-dependent effect on cell invasion, the reverse experiment was performed and H157 non-small cell lung cancer cells (LKB1-null) expressing either GFP-tagged full length (WT) LKB1 or empty GFP control vector were used to analyze invasive morphology. Re-expression of WT LKB1 caused invaded cells to switch from an amoeboid-like morphology to an elongated mesenchymal-like morphology (Figure 2.1). Of the invasive cells expressing LKB1, 61% of the cells were mesenchymal in morphology, as compared to 27% in the LKB1-null parental cells.

Live cell imaging of H1299 pLKO.1 control and shLKB1 spheroids was performed to determine the percentage of amoeboid cells present in the total invasive population over time. These data confirm that LKB1 loss induces a switch to amoeboid morphology compared to control cells, and this switch was stable across all time points measured (Figure 2.1). Single cell track plots show that LKB1-depleted amoeboid cells move greater distances from their point of origin compared to mesenchymal cells found in the LKB1-depleted population and even other amoeboid cells found in pLKO.1 control cells (Figure 2.1). While no difference in cell directionality was seen with LKB1 loss as measured by meandering index, LKB1-depleted amoeboid cells show a significantly increased velocity compared to all other cell types (Figure 2.1), including amoeboid cells found in LKB1 wild-type pLKO.1 controls. These data suggest that amoeboid cell morphology alone cannot solely explain the increase in velocity and distance from origin observed in the LKB1-depleted amoeboid cells.

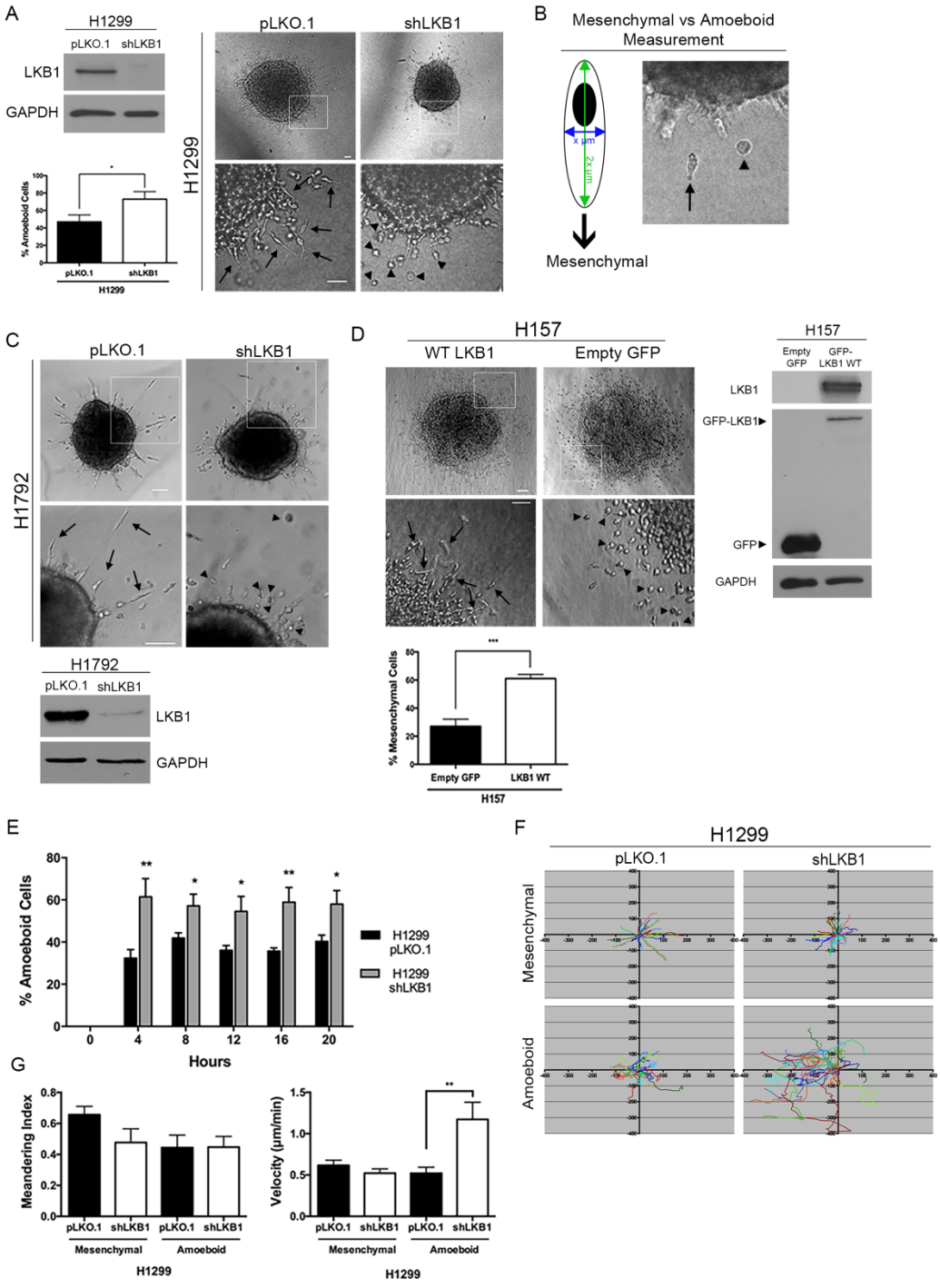

Figure 2.1. LKB1 induces a mesenchymal-amoeboid switch in 3-D invasive morphology. (A) Western blot (left) showing H1299 cells were stably depleted of LKB1 using a targeting shRNA lentivirus. Empty pLKO.1 vector was used as a control. Spheroids from H1299 pLKO.1 and shLKB1 cells were embedded in a collagen type I matrix and imaged at 24 hours post-embedding (right). Zoomed images are shown below. Amoeboid and mesenchymal cell morphologies were quantified as a percentage back to the total number of cells invaded in each spheroid. n = 3 spheroids. Scale = 50 µm. Arrows = mesenchymal cells,

arrowheads = amoeboid cells. (B) Schematic illustrating how mesenchymal and amoeboid cells were quantified. Any cell whose length was greater than or equal to twice its width was considered a mesenchymal cell. Image on right shows an example of each cell morphology. Arrow = mesenchymal cell, arrowhead = amoeboid cell. (C) Experiment from (A) was repeated in H1792 NSCLC cells. Western blot confirming LKB1 knockdown is shown below. (D) H157 cells (LKB1-null) were transiently transfected with either empty pcDNA3-GFP or pEGFP-C1 LKB1 WT vector. Spheroids were analyzed at 24 hours for invasive phenotypes. Western blot confirms expression of GFP-LKB1 (right). Bar graph showing the percentage of mesenchymal cells in empty GFP control and LKB1 transfected H157 cells is shown below. (E-G) H1299 pLKO.1 and shLKB1 spheroids were embedded in a collagen matrix and imaged using live cell confocal. (E) Bar graph showing the percentage of amoeboid cells calculated over time in H1299 pLKO.1 and shLKB1 cells. N = 8 spheroids. (F) Cell tracks were plotted from a single point of origin in H1299 pLKO.1 and shLKB1 invasive amoeboid and mesenchymal cells. (G) Bar graph showing meandering index and velocity of H1299 shLKB1 and pLKO.1 subtypes. n = 11 cells. * = $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

The LKB1 C-terminal domain, and specifically its farnesylation, regulate cellular polarity and directional persistence.

Since the majority of LKB1 mutations in lung cancer patients are truncations (50) (Figure 2.2), we made a series of stable cells re-expressing GFP-tagged LKB1 mutants and domain truncates (Figure 2.2) to determine if they could induce mesenchymal invasion in both H157 LKB1-null human lung cancer cells and HeLa (LKB1-null cervical cancer) cells. Using 3-D invasion assays of spheroids embedded in collagen, a full length wildtype LKB1 induced mesenchymal polarization during invasion as compared to empty GFP control (Figure 2.2, Supplementary Figure S2.2), confirming the data seen with the transient transfections. Similarly, H157 cells re-expressing an LKB1 K78I kinase dead mutant (Supplementary Figure S2.3) also exhibit a mesenchymal polarity, indicating that kinase activity is not required for promoting mesenchymal polarization. In contrast, a C430S farnesylation mutant or a K78I and C430S double mutant were unable to significantly restore mesenchymal polarization over empty GFP control, highlighting the role of LKB1 farnesylation in promoting mesenchymal polarization during invasion in a kinase-independent manner.

We then tested the hypothesis that the C-terminal domain of LKB1 alone can restore mesenchymal polarization during invasion, since this contains the C430 farnesylation site. The LKB1 CTD alone is sufficient to promote mesenchymal polarization, highlighting a kinase-independent promotion of mesenchymal polarity. Furthermore, mutation of the C430 site abolishes the ability of the CTD to promote mesenchymal polarity, in both H157 and HeLa cells (Fig 2.2, Supplementary Figure S2.2). Additionally, the LKB1 N-terminal domain (NTD) alone and kinase domain alone both are unable to promote this mesenchymal polarization during invasion (Supplementary Figure S2.4). Overall, these data suggest that the LKB1

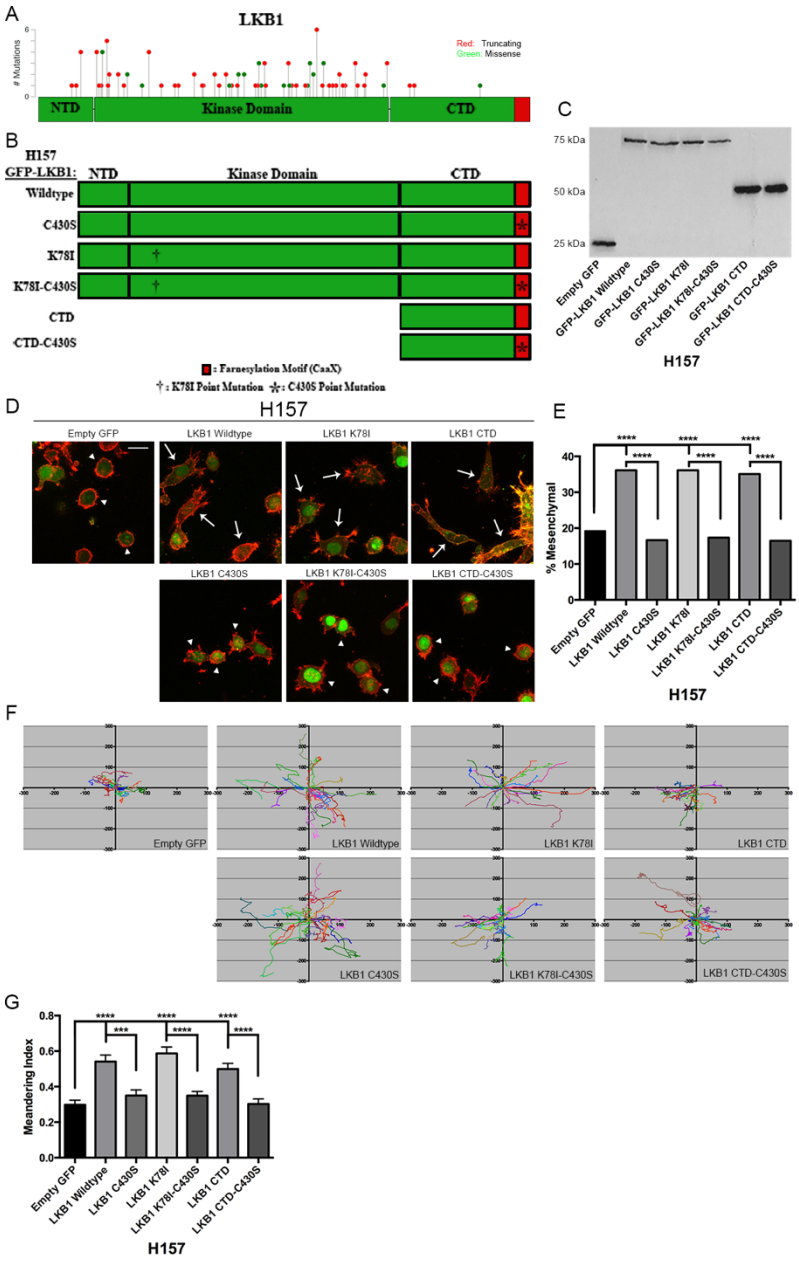

Figure 2.2. LKB1 regulates cellular polarization through its C-Terminal Domain in a farnesylation-dependent manner. (A) LKB1 consists of a central kinase domain with a C-terminal farnesylation motif. Schematic of LKB1 mutations in lung adenocarcinoma patients; data adapted from cBioPortal. Red are truncating mutations, green are missense. (B) Schematic showing H157 (NSCLC, LKB1-null) cells that were generated stably expressing GFP-tagged: wildtype LKB1, a C430S mutation to disrupt farnesylation, a K78I kinase dead mutation, a double mutation with both K78I and C430S, the C-terminal domain (CTD) alone, or the CTD alone with a C430S mutation. (C) Western blot probed with a GFP antibody verifying expression of the H157 stable cells. (D) Immunofluorescence of H157 spheroids embedded in collagen and stained with phalloidin. Amoeboid and mesenchymal morphologies (described in Figure 1) were quantified as a percentage back to the total number of cells invaded in each spheroid. n=4 spheroids. Scale=20 μm. Arrows=mesenchymal cells, arrowheads=amoeboid

cells. (E) The percentage of mesenchymal cells was quantified for each cell line at 24 hours post-embedding. (F) Each cell line was tracked over time. Cell tracks were plotted from a single point of origin. (G) Meandering index was calculated using the cell tracks from F. Meandering index is defined as the linear distance divided by the total path length. n=30 cells ***=$p \leq 0.001$, ****=$p \leq 0.0001$.

C-terminal domain and specifically its farnesylation promote mesenchymal polarity during invasion and, importantly, this occurs independently of LKB1 kinase activity.

Our data show that the CTD and farnesylation, promote mesenchymal polarization during invasion, therefore we examined the role of LKB1 in regulating directional migration. Full length wildtype LKB1, the LKB1 K78I kinase dead mutant, and the LKB1 CTD alone significantly restore directional persistence over empty GFP control (Figure 2.2). Upon mutation of the C430 farnesylation site in wildtype LKB1, LKB1 K78I, and the LKB1 CTD there is an abrogation of directional persistence as compared to their respective wildtype farnesylation construct. Taken together, these data highlight the importance of LKB1 farnesylation, independent of its kinase activity, and specifically the CTD in regulating directed cell invasion.

LKB1 differentially regulates Rho-GTPases through its CTD and farnesylation

We next sought to understand the mechanism by which LKB1 regulates this amoeboid-mesenchymal invasion switch. Since amoeboid invasive motility is driven through a balance of RhoA and cdc42 activity (78, 102, 103), we probed the activity of these GTPases in 3-D spheroids. A time course of H1299 (wildtype LKB1) pLKO.1 control cells shows a robust activation of both RhoA and cdc42 at the 5 hour timepoint (Figure 2.3); however, the isogenic shLKB1 cells have reduced active cdc42 and RhoA. To enrich for motile cells, GTPase activation assays were performed in 2-D. The data show that H1299 cells activate both RhoA and cdc42 but this activation is severely attenuated upon LKB1 loss (Fig 2.3). These data are consistent with previous reports showing that LKB1 depletion reduces cdc42 and RhoA

activity in motile cells (60, 96, 97), and suggests that LKB1-depleted cells do not rely on canonical Rho-GTPase activity.

Since the LKB1 CTD and specifically its farnesylation are critical for promoting mesenchymal motility, we examined the role of these domains in regulating Rho-GTPase activity. RhoA activity is enhanced upon re-expression of wildtype LKB1 relative to the empty GFP control (Figure 2.3), confirming that LKB1 is responsible for promoting RhoA activity. Similarly, the LKB1 CTD alone is sufficient for promoting RhoA activity, although inhibiting LKB1 farnesylation within the CTD completely abrogates RhoA activation (Figure 2.3). Taken together, these data suggest that LKB1 regulates RhoA activity through its CTD and specifically through its farnesylation.

The role of LKB1 in regulating cdc42 activity was then probed further, since we have previously shown that depletion of LKB1 impacts cdc42 in human lung cancer cell lines in 2-D (60). We show that wildtype LKB1 promotes cdc42 activity, and that this occurs through the LKB1 CTD (Figure 2.3). Interestingly, LKB1 farnesylation has no impact on the activity of cdc42, as the full-length farnesylation mutant is able to activate cdc42 at levels similar to wildtype LKB1. These data suggest that while LKB1 regulates cdc42 activity through its C-terminal domain, this activation is farnesylation-independent, unlike the farnesylation-dependent activation of RhoA (Figure 2.3).

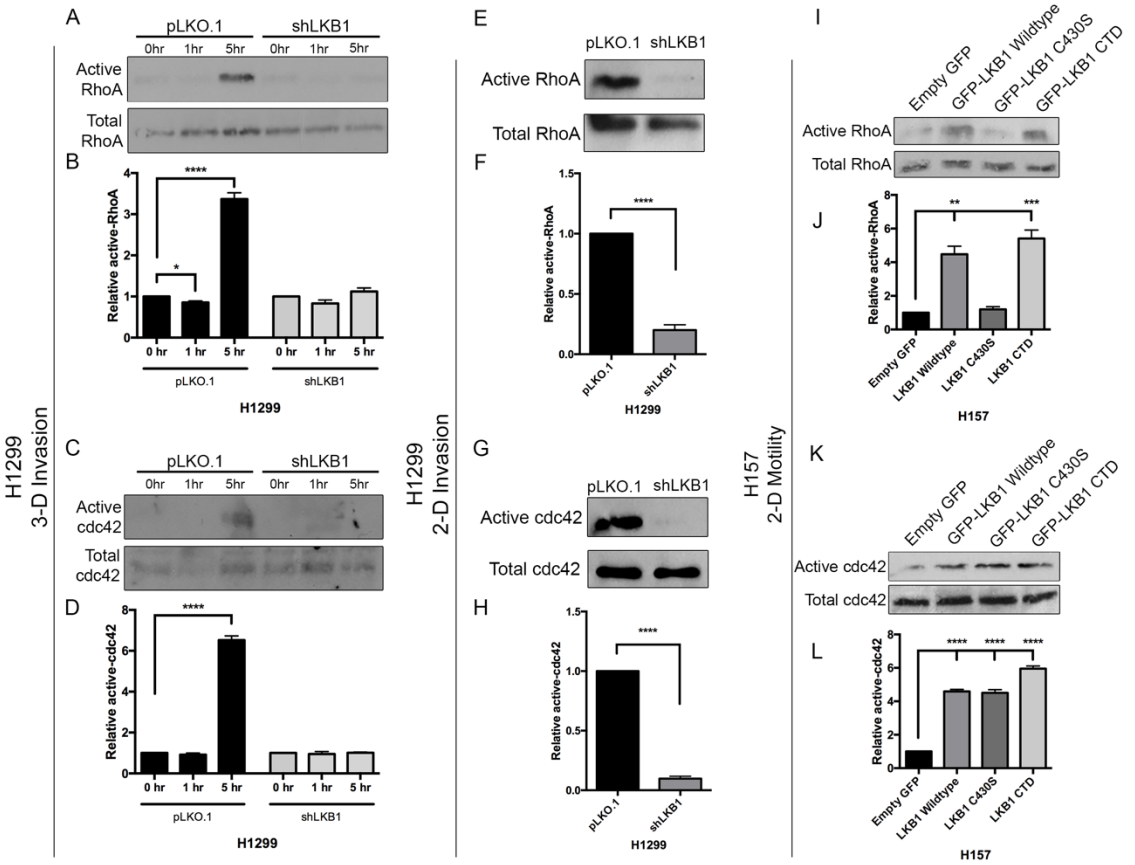

Figure 2.3. LKB1 differentially regulates RhoA and cdc42. (A) Representative western blot showing RhoA activity assay over time in pLKO.1 and shLKB1 H1299 cells embedded in a collagen type I matrix. (B) Densitometry of western blot from (A) normalized to total RhoA levels. (C) Representative western blot of a cdc42 activity assay over time in pLKO.1 and shLKB1 H1299 cells embedded in a collagen type I matrix. (D) Densitometry of western blot from (C), normalized to total cdc42 levels. (E) Representative western blot showing RhoA activity assay in 2-D in pLKO.1 and shLKB1 H1299 cells. (F) Densitometry of western blot from (E), normalized to total RhoA levels. (G) Representative western blot showing cdc42 activity assay in 2-D in pLKO.1 and shLKB1 H1299 cells. (H) Densitometry of western blot from (G), normalized to total cdc42 levels. (I) Representative RhoA activity assay of H157 cells stably expressing either empty GFP control or GFP-tagged: wildtype LKB1, LKB1 C430S, or the LKB1 C-terminal domain. (J) Densitometry of western blot from (I), normalized to total RhoA levels. (K) Representative western blot showing a cdc42 activity assay of H157 cells stably expressing either empty GFP control or GFP-tagged: wildtype LKB1, LKB1 C430S, or the LKB1 C-terminal domain. (L) Densitometry of western blot from (K), normalized to total cdc42. n=3 experiments. *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****=$p \leq 0.0001$.

LKB1 farnesylation activates RhoA to promote mesenchymal polarity

The current data show the LKB1 CTD, and specifically its farnesylation, promote both mesenchymal cell polarity and RhoA activity, we next sought to determine whether LKB1 promotes mesenchymal polarization through RhoA signaling. We created double stable cells expressing either empty GFP or GFP-LKB1 wildtype or a C430S farnesylation mutant, and the constitutively active form of either RhoA (Q63L) or cdc42 (Q61L) (Figure 2.4). Cells re-expressing wildtype LKB1 and either constitutively active RhoA or cdc42 maintain a mesenchymal polarization similar to cells re-expressing wildtype LKB1 alone (Figure 2.4). Similarly, rescuing cdc42 activity in farnesylation-mutant cells results in a predominantly amoeboid phenotype, similar to LKB1 farnesylation-mutant cells alone (Figure 2.4), suggesting that LKB1 does not signal to cdc42 to promote mesenchymal polarization. However, upon rescuing RhoA activity in LKB1 farnesylation-mutant cells, cells re-acquire a mesenchymal polarization during 3-D invasion (Figure 2.4). Importantly, this result is consistent with mesenchymal polarization in cells re-expressing wildtype LKB1. Restoring either RhoA or cdc42 activity in empty GFP control cells fails to restore this mesenchymal polarization (Figure 2.4). Given that our previous data highlight the role of LKB1 farnesylation in promoting mesenchymal polarity, these data suggest this occurs through LKB1 signaling to RhoA, as rescuing RhoA activity in farnesylation-compromised cells restores mesenchymal polarity.

This data was expanded by examining directional migration in these cells expressing constitutively active RhoA or cdc42. Similar to cells re-expressing wildtype LKB1, cells re-expressing constitutively active RhoA or cdc42 maintain a strong directional persistence (Figure 2.4). Additionally, cells re-expressing constitutively active cdc42 in LKB1 farnesylation-mutant cells have poor directional persistence similar to farnesylation-mutant

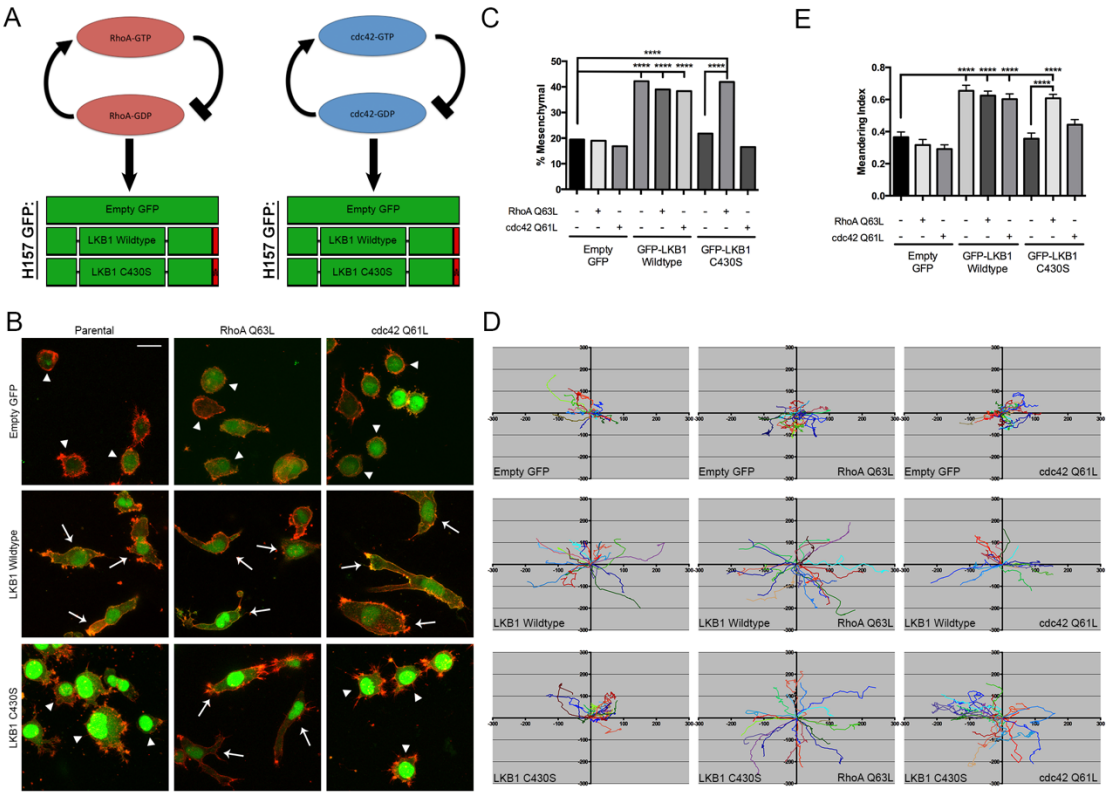

Figure 2.4. Constitutively active RhoA restores mesenchymal polarity in LKB1 farnesylation-mutant cells. (A) Schematic of double-stable cell lines expressing either empty GFP control or GFP-tagged: wildtype LKB1 or LKB1 C430S, with constitutively active RhoA (Q63L) or cdc42 (Q61L). (B) Spheroids of H157 cells expressing either empty GFP control or GFP-tagged: wildtype LKB1 or LKB1 C430S, and spheroids of these cells also expressing constitutively active RhoA or cdc42, were embedded in a collagen type I matrix. 24 hours post-embedding cells were fixed and stained with phalloidin. Amoeboid and mesenchymal morphologies (described in Figure 1) were quantified as a percentage back to the total number of cells invaded in each spheroid. n=4 spheroids. Scale=20 μm. Arrows=mesenchymal cells, arrowheads=amoeboid cells. (C) The percentage of mesenchymal cells was quantified for each cell line at 24 hours post-embedding. (D) Each cell line was tracked over time. Cell tracks were plotted from a single point of origin. (E) Meandering index was calculated using the cell tracks from (D). Meandering index is defined as the linear distance divided by the total path length. n=30 cells. ****=p≤0.0001.

cells alone, suggesting that cdc42 signaling is not responsible for promoting directionality. However, upon rescuing RhoA activity in these farnesylation-compromised cells, cells restore their directional persistence similar to cells re-expressing wildtype LKB1 (Figure 2.4). Similar to mesenchymal polarization, restoring RhoA or cdc42 activity in empty GFP control cells fails to restore directional persistence. Together, these data highlight a LKB1 farnesylation-RhoA pathway to promote mesenchymal polarity and strong directional persistence during 3-D invasion.

The regulation of FAK activity by LKB1 is kinase dependent.

Our current data show that the LKB1 CTD is responsible for controlling this amoeboid-mesenchymal switch during 3-D invasion. We and others have also shown that LKB1 loss results in an increase in adhesion signaling, most notably through hyperactive focal adhesion kinase (FAK) signaling (43, 62-64). Thus, we sought to determine if this amoeboid-mesenchymal switch is related to adhesion signaling by using our panel of H157 stable cells to assess $pFAK^{397}$ activity. Compared to empty GFP control cells, H157 cells expressing wildtype LKB1 showed repression of the total number of $pFAK^{397}$ sites using immunofluorescence (Figure 2.5). This result is consistent with previous findings that LKB1 is a pFAK repressor in lung cancer cells (63, 64). Interestingly, this repression was not dependent on LKB1 farnesylation, as mutation of the LKB1 C430 farnesylation motif had no impact on the ability of LKB1 to repress $pFAK^{397}$ (Figure 2.5). In contrast, when re-expressing either the K78I kinase dead mutant or the K78I-C430S double mutant LKB1, $pFAK^{397}$ expression was not repressed and remained at similar levels as empty GFP control cells (Figure 2.5). While LKB1 served to repress total number of $pFAK^{397}$ sites/cell, the mean intensity of each pFAK site was similar across all cell lines (Supplementary Figure S2.5). These data suggest that the kinase

activity of LKB1, but not farnesylation, is necessary for LKB1 to repress FAK during 3-D invasion.

To further probe the mechanism by which LKB1 regulates FAK, we analyzed the role of the downstream target of LKB1 kinase activity, MARK1, which represses FAK through an LKB1-MARK1 pathway (63). Immunofluorescence staining of $pFAK^{Y397}$ in MARK1 siRNA-depleted cells show that MARK1 loss increases pFAK expression when compared to scrambled siRNA control (Figure 2.5), thus phenocopying the loss of LKB1. Quantification of $pFAK^{Y397}$ staining confirms that MARK1 loss increases the number of pFAK sites per cell (Figure 2.5), with a slight increase in the intensity of each pFAK site (Supplementary Figure S2.5). Western blot analysis of three LKB1 wildtype lung cancer lines (H1792, H1299, and H157 + GFP-LKB1) also shows an increase in $pFAK^{Y397}$ with MARK1 knockdown (Figure 2.5); however, MARK1 knockdown had no significant impact on cell morphology (Supplementary Figure S2.6), which is consistent with our data that cell polarity during 3-D invasion is independent of its kinase function.

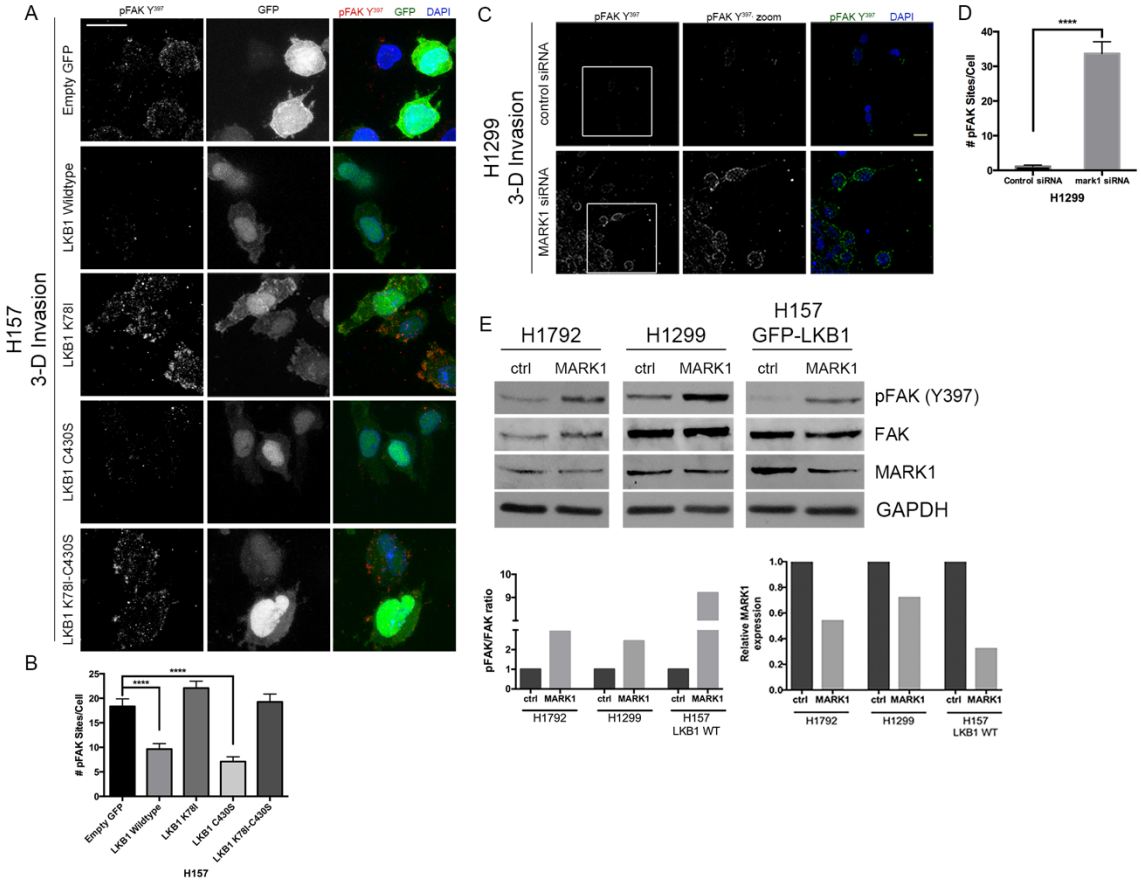

Figure 2.5. LKB1 regulates pFAK activity through its kinase domain. (A) Spheroids of H157 cells stably expressing either empty GFP control or GFP-tagged LKB1 constructs embedded in a collagen type I matrix. After 24 hours, cells were fixed and stained by immunofluorescence for pFAKY397, GFP, and DAPI. (B) Total number of pFAKY397 sites for each experimental group in (A) were quantified. (C) Expression of pFAKY397 was examined by immunofluorescence of spheroids after 24hrs of invasion. DAPI was used to stain the nuclei of the cells. (D) The total number of pFAKY397 sites per cell was quantified from the images obtained in (C). ****=p≤0.0001. (E) Western blot showing pFAKY397 expression after MARK1 siRNA depletion in H1792, H1299, and H157 LKB1 WT cells compared to scrambled control siRNA. GAPDH was used as a loading control. Densitometry analysis of phospho to total FAK ratio (left) and relative MARK1 expression (right) in control siRNA and MARK1 siRNA treated cells are shown below.

LKB1-depleted amoeboid cells require FAK activation to navigate through a collagen matrix

We next wanted to determine the impact of LKB1 loss specifically on the amoeboid population as these cells showed a significant increase in velocity during invasion, even more so than amoeboid cells in the wildtype population. We first assessed if FAK was hyperactive in LKB1-depleted cells. Western blot confirms an increase in $pFAK^{Y397}$ in H1299 and H1792 LKB1-depleted cells as compared to controls (Figure 2.6); however, this only provided information on the whole population. Therefore, we then analyzed single invasive cells for their pFAK status based upon morphology in H1299 pLKO.1 and shLKB1 spheroids via immunofluorescence. We confirmed that LKB1 loss resulted in $pFAK^{Y397}$ hyperactivation compared to pLKO.1 control cells (Figure 2.6). The mean intensity of each individual pFAK site significantly increased from 748 in pLKO.1 cells to 1751 in shLKB1 cells and the total number of pFAK sites per cell showed a significant increase from 3 to 72 sites per cell with LKB1 depletion (Figure 2.6). This increase in $pFAK^{Y397}$ also resulted in increased downstream adhesion signaling, since both the mean phospho-paxillin ($pPax^{Y118}$) site intensity and total number of $pPax^{Y118}$ sites per cell significantly increases (Figure 2.6).

We next examined $pFAK^{Y397}$ expression specifically in mesenchymal and amoeboid cells of pLKO.1 and shLKB1 cells to determine if shLKB1 amoeboid cells have unique FAK activity compared to wild-type LKB1 amoeboid cells. These data show that LKB1 loss increases $pFAK^{Y397}$ expression in both mesenchymal and amoeboid cells that lack LKB1 (Fig 2.6); however, the amoeboid populations within shLKB1 cells express significantly higher levels of pFAK when compared to amoeboid cells in pLKO.1 control cells (Figure 2.6), showing that this pFAK activity does not solely arise from amoeboid morphology.

To further probe this finding, we tested the hypothesis that inhibition of FAK activity will specifically inhibit shLKB1 amoeboid cell invasion due to the aberrantly active $pFAK^{Y397}$ levels. We first transiently knocked down FAK via targeted siRNA and found that FAK knockdown completely abolished invasion in the H1299 shLKB1 cells (Supplementary Figure S2.6). To address how FAK activity alters cell morphology and motility specifically in LKB1-depleted cells, we pharmacologically inhibited FAK. H1299 pLKO.1 and shLKB1 spheroids were exposed to 1μM of FAK inhibitor PF-573228 and assayed for invasion over time using live cell imaging. Inhibition of FAK did result in a significant decrease in the total shLKB1 cells invaded when compared to DMSO control, decreasing from about 65 cells/spheroid to 35 cells/spheroid after 16h of invasion (Figure 2.6). In contrast, the FAK inhibitor had no significant impact on the total cells invaded in pLKO.1 spheroids (Figure 2.6), suggesting these cells are less dependent on adhesion during invasion. Importantly, using 1μM FAK inhibitor did not completely abolish invasion in the shLKB1 cells as siRNA treatment did, thus allowing for further analysis of cell motility features in the presence of inhibitor. Interestingly, it was specifically the velocity of LKB1-depleted amoeboid cells that was significantly inhibited by exposure to FAK inhibitor (Figure 2.6). The velocity of shLKB1 amoeboid cells in the presence of PF-573228 was equivalent to the velocity of pLKO.1 amoeboid cells, suggesting that increased FAK activity provides LKB1-depleted amoeboid cells an advantage of faster motility during invasion. Additionally, while FAK inhibition specifically impacts the velocity of shLKB1 amoeboid cells, it had no significant impact on cell shape (Figure 2.6). These data taken together suggest that amoeboid cells lacking LKB1 expression represent an atypical population of rounded cells that utilize adhesion signaling for invasion through the collagen gel.

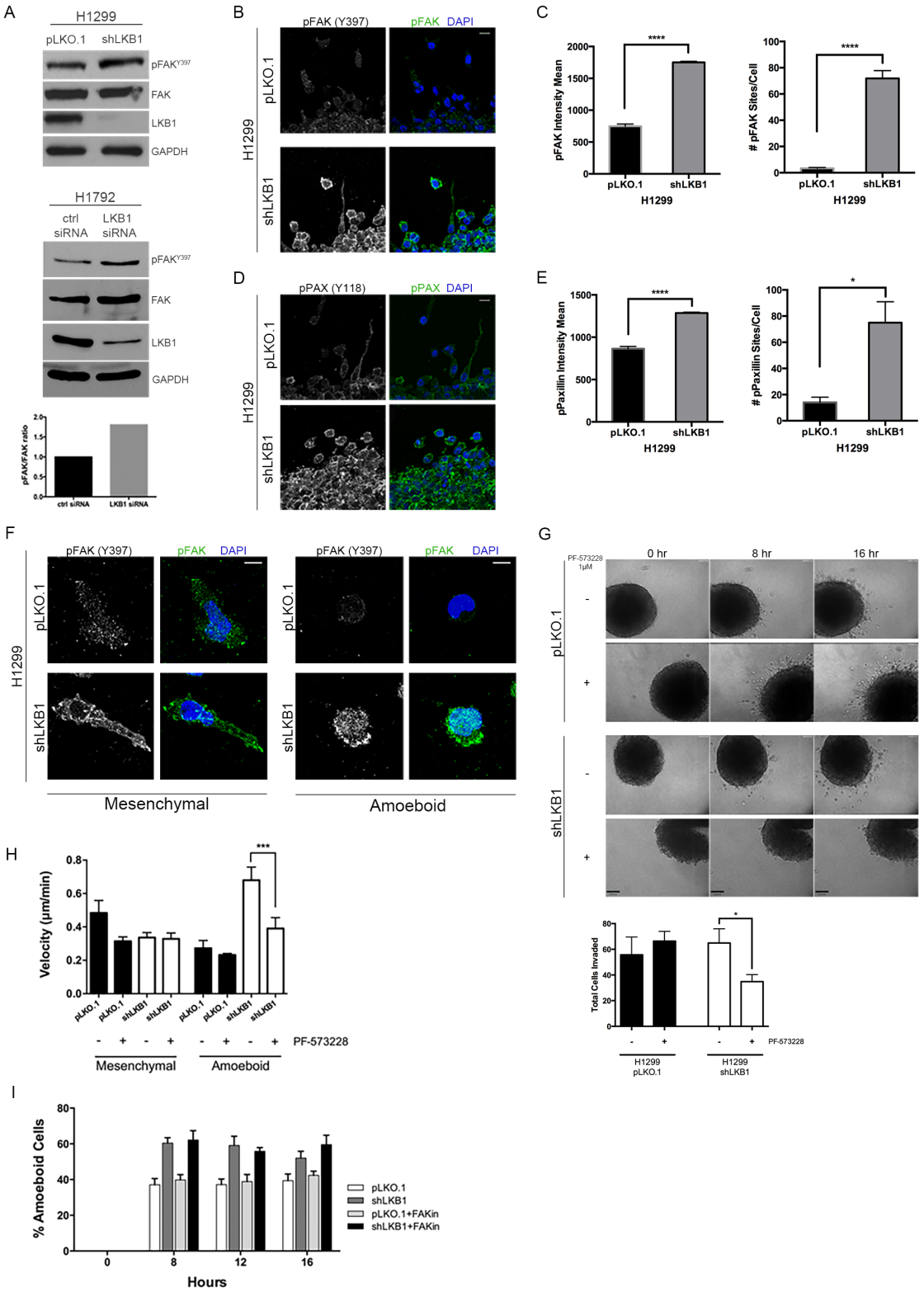

Figure 2.6. LKB1-depleted amoeboid cells are dependent on pFAK during invasion. (A) Western analysis of pFAK^{Y397} expression in H1299 pLKO.1 and shLKB1 cells (top). A similar experiment was performed in H1792 LKB1 siRNA treated cells (below). Bar graph

shows densitometry of the phospho to total FAK ratio in the H1792 LKB1 knockdown cells as compared to control siRNA. (B) H1299 pLKO.1 and shLKB1 spheroids were analyzed for expression of $pFAK^{Y397}$ using immunofluorescence imaging. Scale=20 μm. (C) Quantification of $pFAK^{Y397}$ site intensity (left) and total number of $pFAK^{Y397}$ sites per cell (right). (D) Activity of downstream FAK signaling was analyzed using immunofluorescence imaging of $pPAX^{Y118}$ in H1299 pLKO.1 and shLKB1 spheroids. Scale=20 μm. (E) Quantification of $pPAX^{Y118}$ site intensity (left) and total number of $pPAX^{Y118}$ sites per cell (right). (F) Zoomed images of pFAK immunofluorescence imaging (described in B). Mesenchymal and amoeboid cell types in H1299 pLKO.1 and shLKB1 are shown. Scale=10 μm. (G-I) H1299 pLKO.1 and shLKB1 spheroids were exposed to either 1μM of PF-573228 FAK inhibitor or vehicle control and imaged over time. (G) Still images of the live cell imaging experiment were taken every 8hrs. Quantification of total cells invaded shown below shows significantly decreased cells invaded in shLKB1 cells treated with FAK inhibitor. Scale=100 μm (H) Cell tracks from the FAK inhibitor experiment showed that LKB1-depleted amoeboid cells have a significantly decreased velocity during invasion as compared to vehicle control. n=8-15 cells. (I) The percentage of amoeboid cells in the live cell FAK inhibitor experiment was assessed at 0, 8, 12, and 16 hours. n=5-6 spheroids. *=$p \leq 0.05$, ***=$p \leq 0.001$, ****=$p \leq 0.0001$.

LKB1 loss causes an increase in collagen remodeling during 3-D invasion

Since we show that LKB1 loss results in a unique amoeboid cell population, we wanted to determine if this provides an invasive advantage while navigating the microenvironment. To do this we performed multiphoton imaging on H1299 pLKO.1 and shLKB1 spheroids to visualize collagen remodeling and its relationship to cell type and invasive potential. LKB1 loss resulted in an increase in collagen alignment at 6 and 21 hours (Figure 2.7). We used a novel local alignment coefficient to quantify the heterogeneous alignment patterns. CT-FIRE (curvelet transform fiber extraction) software was utilized to extract collagen fibers (Figure 2.7). All fibers were quantized with a 5-pixel length. Then for every pixel, we measured the local alignment coefficient parameter for every pixel by selecting all fiber segments within a circular neighborhood of 20 pixels (Figure 2.7) to generate the alignment field (Supplementary Figure 2.7 explains optimization of local alignment coefficient calculation). Using this parameter, histograms of local alignment coefficients, surface plots, and contour plots are generated to quantify alignment. Using this quantification of the local alignment coefficient and comparing back to the 0 hour baseline value, we found that at 6 hours, shLKB1 cells show an increase in collagen alignment, which is further accentuated at 21 hours; on the other hand, pLKO.1 control cells result in a decrease in the number of aligned fibers over time (Figure 2.7), suggesting that LKB1 may actually negatively regulate remodeling during invasion. Thus, LKB1-depleted cells are more efficient at re-aligning collagen fibers as they invade. Interestingly, this realignment of the collagen matrix seems to be occurring via a matrix metalloproteinase (MMP)-independent mechanism. When treated with the pan-MMP inhibitor GM6001, shLKB1 show no significant change in invasion when compared to vehicle control (Supplementary Figure S2.8). These data suggest that LKB1 loss promotes collagen remodeling in an MMP-independent manner.

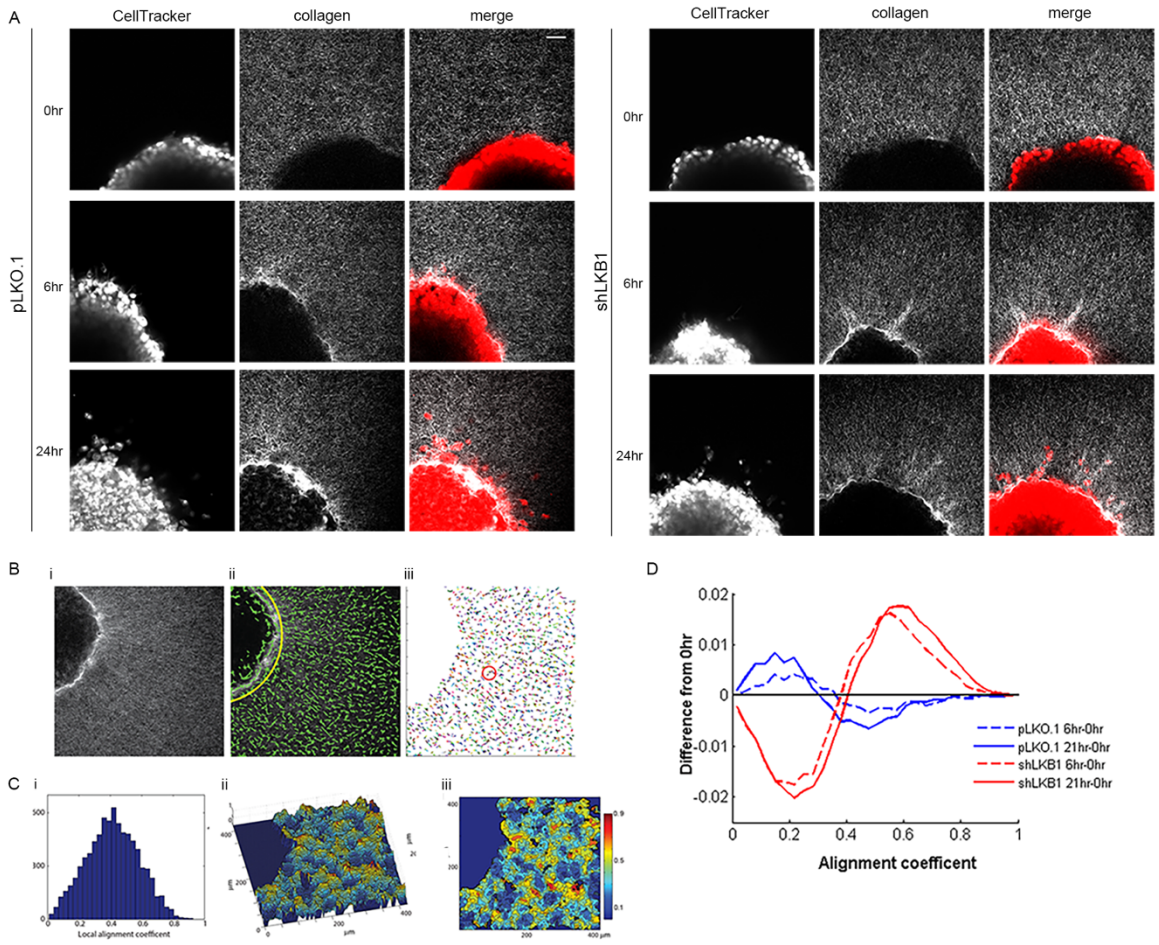

Figure 2.7. Loss of LKB1 results in increased collagen remodeling during invasion. (A) H1299 pLKO.1 and shLKB1 spheroids and the collagen matrix were imaged using second harmonic generation microscopy. Spheroids were dyed using CellTracker Red in order to visualize cells during invasion. Images were obtained at 0, 6, and 21 hours post-embedding. Scale=50 μm. (B) Images from A were quantified using collagen alignment analysis. A single z-stack image (i) is used in CT-FIRE software to extract collagen fibers (green, ii). The software automatically determines various fiber lengths in the image, represented as different line colors (iii). Yellow line represents the manually selected tumor boundary. (C) i) Example histogram generated via CT-FIRE analysis of collagen alignment coefficients. ii) Surface and iii) contour plots of local alignment show topography of alignment patterns. (D) Alignment analysis was performed as described in B,C for H1299 pLKO.1 (blue) and shLKB1 (red) spheroids at 6 and 21 hours, with the 0 hour baseline alignment subtracted to remove any initial bias.

LKB1 kinase activity represses collagen remodeling through regulation of FAK activity

We next wanted to test the hypothesis that LKB1 kinase activity represses collagen remodeling, therefore we assessed collagen remodeling and invasion in the H157 cells stably expressing various LKB1 domains or mutants. At 24 hours, H157 empty GFP invading cells show epicenters of significant collagen alignment (Figure 2.8), whereas cells expressing either LKB1 wildtype or LKB1 C430S show significantly less alignment (Figure 2.8). However, cells expressing either the K78I kinase dead LKB1 or the K78I-C430S double mutant LKB1 fail to repress alignment (Figure 2.8), suggesting that kinase activity is required for collagen remodeling. Quantification of alignment shows that empty GFP, LKB1 K78I, and LKB1 K78I-C430S cells have increased alignment coefficients when compared to those cells expressing LKB1 wildtype or C430S (Figure 2.8).

MARK1 was then transiently knocked down via siRNA in H157 LKB1 wild-type cells to determine if LKB1 kinase activity signals through MARK1 to repress FAK activity. Cells lacking MARK1 phenocopy LKB1-depleted cells, showing increased remodeling when compared to scrambled siRNA control (Figure 2.8). This remodeling can be inhibited by treating H1299 shLKB1 cells with PF-573228 FAK inhibitor (Supplementary Figure S2.8), supporting the concept that FAK activity is required for the increase in manipulation of the collagen gel. These data point to the LKB1-MARK1-FAK pathway for regulating collagen remodeling as cells invade.

We next reasoned that since the LKB1-depleted amoeboid cells have high levels of FAK activity, even compared to wild-type amoeboid cells, perhaps these cells also can remodel collagen. This would be unlike typical amoeboid cells, which in general do not remodel

collagen (73, 77). To test this, we acquired images of single amoeboid cells within the H157 cells expressing different LKB1 domains and mutants. Surprisingly, H157 empty GFP invading amoeboid cells were associated with high collagen alignment (Figure 2.8). In contrast, H157 amoeboid cells with wild-type LKB1 do not show local collagen alignment. The LKB1 K78I mutant was similar to the GFP control, indicating that LKB1 kinase activity is required to suppress local collagen alignment. Re-expression of the LKB1 C430S mutant was similar to LKB1 wildtype cells and showed no local alignment in amoeboid cells, again indicating that LKB1 kinase activity but not farnesylation is important for collagen remodeling. These data show that amoeboid cells lacking LKB1 kinase activity boast the invasive advantage of remodeling a 3-D collagen matrix, which does not typically occur in amoeboid motile cells.

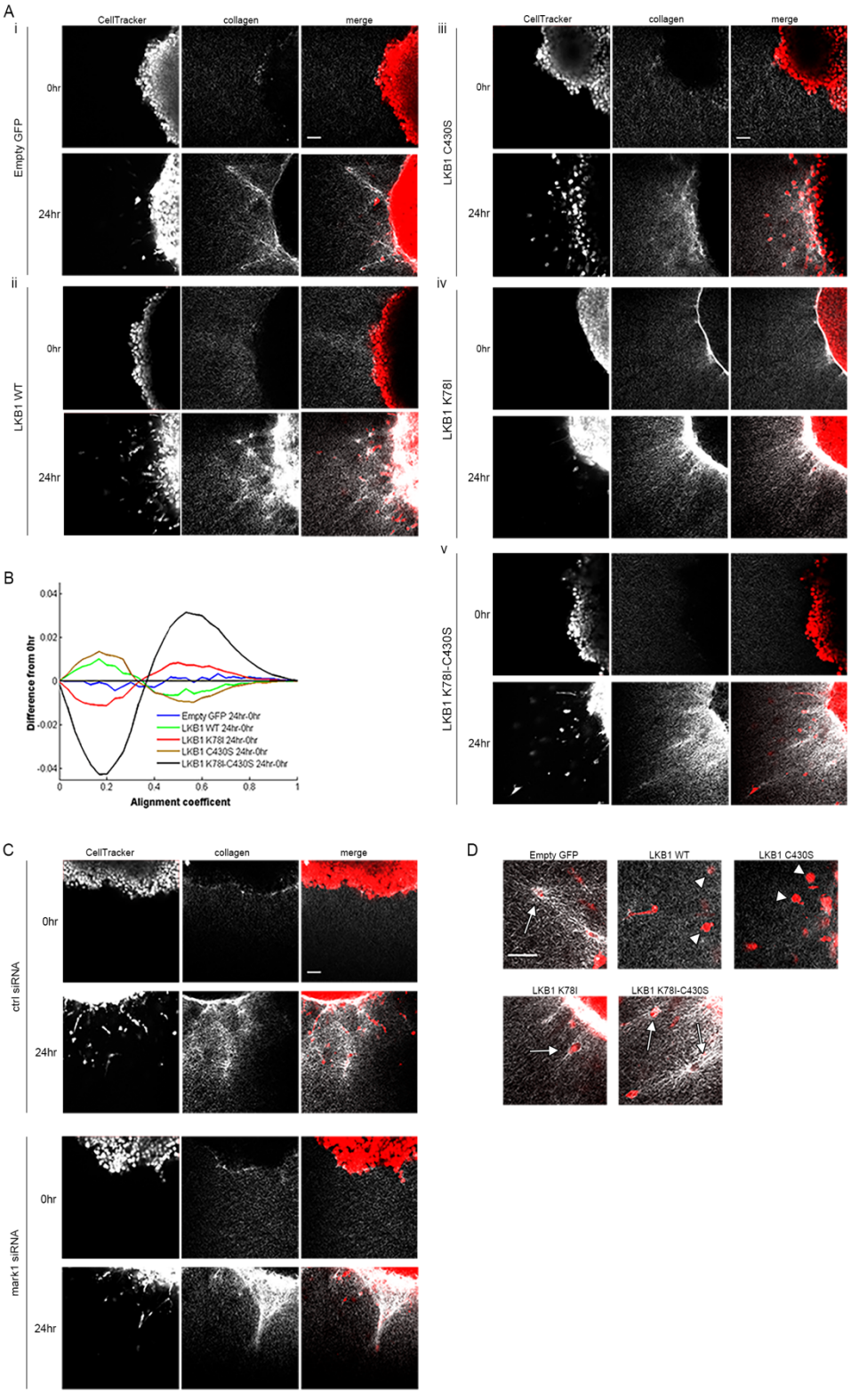

Figure 2.8. Kinase activity of LKB1 represses collagen remodeling. (A) Multiphoton imaging was performed at 0 and 24 hours to visualize collagen using H157 stables with the

following LKB1 constructs: Empty GFP control, LKB1 WT, LKB1 C430S (farnesylation mutant), LKB1 K78I (kinase dead), and LKB1 K78I-C430S. (B) Images obtained in A were quantified for collagen alignment using CT-FIRE image analysis. (C) Collagen SHG imaging of MARK1 siRNA depleted H157 LKB1 wildtype stable cells compared to scrambled siRNA. (D) Zoomed images showing collagen and single invading cells. Arrow = amoeboid cell that remodels collagen, arrowhead = amoeboid cell that does not remodel collagen. Scale = 50 μm.

2.4 Discussion

The functional diversity of LKB1 has been attributed to it phosphorylating 14 members of the AMPK family of proteins, which, when activated, go on to regulate a diverse set of biological processes (56-58). However, multiple reports show a kinase-independent function of LKB1 that is linked to cell polarity (89, 104). Since a majority of the truncating mutations (50) (~72% of *LKB1* mutations in lung adenocarcinoma patients, Figure 2.2) would predictably disrupt its kinase activity and remove its C-terminal domain, we sought to uncouple defects in CTD function from defects in kinase function in the context of cancer cell invasion. These results support an overall model whereby the LKB1 CTD regulates cell polarization through a mesenchymal-amoeboid cell switch, while the kinase domain regulates FAK-based cell adhesion during invasion. These results would suggest that when both kinase activity and CTD function are compromised, both aberrant cell polarity and adhesion programs would ensue. Our data indicates that this is indeed the case; complete LKB1 depletion creates a uniquely invasive, amoeboid shaped cell that, in contrast to typical amoeboid cells (38, 105, 106), maintains a hyperactive FAK-based cell adhesion program and remodels collagen. We speculate that truncating mutations in *LKB1* mutant patients may create a similar scenario (Figure 2.9), where both CTD and kinase function is disrupted, leading to increased invasiveness by creating an agile cell that can be amoeboid, adherent, and able to navigate the tumor microenvironment.

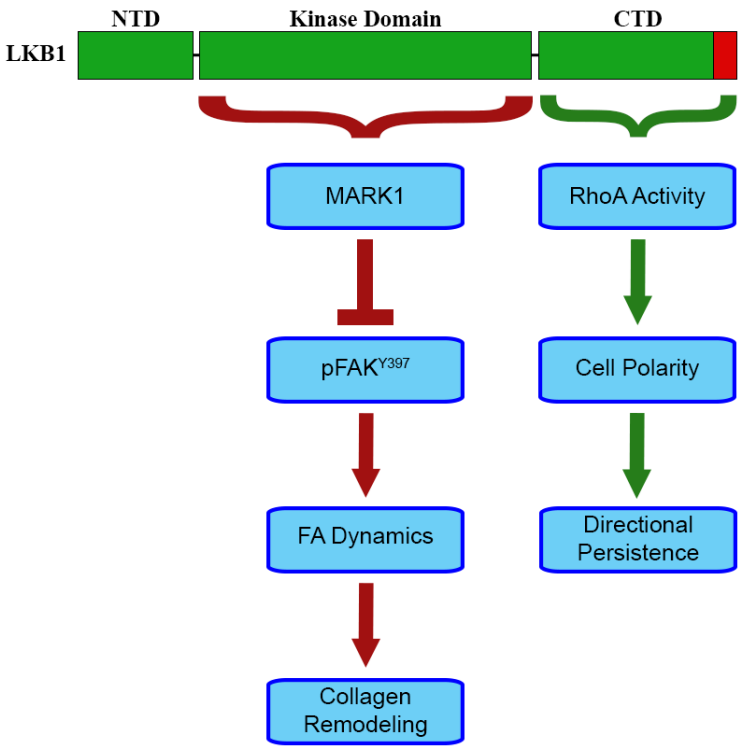

Figure 2.9. Model figure. LKB1 provides kinase-dependent and -independent mechanisms of regulating cell polarity during invasion. Through the LKB1 C-terminal domain (CTD) and its farnesylation, LKB1 activates the Rho-GTPase RhoA to promote mesenchymal polarization and strong directional persistence during invasion. Independent of its CTD, LKB1 kinase activity phosphorylates MARK1 to repress the active form of focal adhesion kinase (pFAK), leading to regulation of focal adhesion dynamics and collagen remodeling during 3-D invasion.

Our results show that specifically LKB1 CTD farnesylation is required for proper polarization during invasion, such that when farnesylation is intact, cells are more mesenchymal, while cells revert to an amoeboid morphology upon loss (Figure 2.2). This suggests that LKB1 membrane localization drives proper cell polarization and controls a mesenchymal-amoeboid morphological switch. Since LKB1 loss leads to inactivation of the small Rho-GTPases RhoA and cdc42 (60, 96, 97) (Figure 2.3), we probed this observation and show that while the LKB1 CTD alone is capable of activating both RhoA and cdc42, LKB1 farnesylation is critical only for the activation of RhoA (Figure 2.3), suggesting a prenylation-independent regulation of cdc42. Similarly, a constitutively active RhoA, but not cdc42, can rescue mesenchymal polarization upon LKB1 farnesylation defects (Figure 2.4), again supporting a role for RhoA but not cdc42 in promoting mesenchymal polarization. Interestingly, it appears a region of LKB1 is necessary to rescue polarity defects in these cells, as restoring RhoA activity in empty GFP control cells fails to restore mesenchymal polarity. Previous studies show that LKB1 regulates p114RhoGEF to promote RhoA activity (Xu et al., 2013), and thus we propose that LKB1 CTD farnesylation anchors LKB1 into the membrane to promote this activity. It has always remained unclear as to why LKB1 loss would inactivate RhoA and cdc42 signaling (60, 96, 97), since it would be expected that LKB1 loss, which causes increased invasion (61, 107), would instead hyperactivate RhoA and cdc42. However, recent studies show RhoA loss-of-function driver mutations in gastric cancer (108, 109), suggesting that inactivation of RhoA, and perhaps cdc42, could in fact be drivers of tumor progression.

We show that LKB1 is a repressor of adhesion signaling and collagen remodeling where intact kinase activity is required to maintain normal FAK levels (Figures 2.5-2.8). Restoration of LKB1 kinase activity in LKB1-deficient cells is sufficient to repress FAK-positive adhesion

sites and collagen remodeling (Figure 2.5). Depletion of the LKB1 kinase target, MARK1, is sufficient to phenocopy this finding, and these data are consistent with previous findings that highlight LKB1 as a FAK repressor (62, 64) through its kinase-dependent activation of MARK1 (63); though our previous data also suggest that the LKB1 NTD alone can also repress FAK, suggesting a potential dual mechanism for interacting with and controlling FAK activity. Our data also builds on the LKB1-MARK1 signaling pathway by showing that it regulates collagen remodeling during invasion in a FAK-dependent manner (Figure 2.8). LKB1 regulates lysyl oxidase (LOX), a collagen crosslinking enzyme, through the mTOR/HIF-1α signaling pathway (65), where LKB1 loss leads to increased LOX expression and collagen deposition (65, 110). Potentially, these two pathways are linked or related through FAK, since LOX is responsible for hypoxic human cancer cell invasion through FAK activity and cell-matrix adhesion (111).

In contrast to classical amoeboid cell motility, LKB1-depleted amoeboid cells lack Rho-GTPase activity, have high FAK activity, and still remodel collagen (Figures 2.3, 2.5-2.8). Classical amoeboid cells do not generally make sustained contacts with the extracellular matrix and do not remodel collagen fibers as they move; instead, their motility is dependent upon squeezing through and deforming the matrix during invasion (38, 73, 105, 106, 112). Additionally, RhoA is frequently implicated in amoeboid invasion by promoting ROCK-mediated myosin light chain (MLC) phosphorylation to promote actomyosin contractility required for cell blebbing during invasion (76, 113-116). In contrast, amoeboid cells lacking LKB1 activity have reduced RhoA activity, and instead, restoring RhoA in LKB1-depleted cells promotes a mesenchymal, polarized cell (Figure 2.4). Therefore, we propose that LKB1 inactivation in both its kinase domain and CTD disrupts cell polarity and adhesion signaling,

resulting in a uniquely invasive cell that is adhesive, amoeboid in shape, and remodels collagen, but may only represent an amoeboid cell in shape rather than the underlying molecular biology.

In vivo studies highlight the importance of LKB1 as a metastasis suppressor. In a seminal publication, LKB1 function was assessed using a *Kras*^{G12D}-driven genetically engineered mouse model (GEMM) of lung cancer (43). In this model, *Lkb1* inactivation in mutant *Kras* tumors led to increased tumor burden, shortened survival time, and increased metastasis compared to *Kras* mutant-only mice. Furthermore, these mice have hyperactive FAK (62), similar to that described in our 3-D model here. It remains difficult to assess polarity during invasion in an *in vivo* model to determine if these cells have polarity defects as well; however, future imaging of living lung tissue slices could shed light on this. Additionally, given that LKB1 serves as the upstream regulator of AMPK (5' AMP-activated protein kinase) in the energy stress response pathway (55), the interplay between defects in metabolic signaling, adhesion, and cell polarity remains unclear in the context of lung cancer metastasis.

Taken together, our data support a model whereby LKB1 kinase-dependent and -independent functions have separate roles in regulating various cellular processes during cancer cell invasion that when lost, synergize to create an uniquely-invasive cell. Thus, we speculate that loss of both LKB1 kinase activity and the CTD, which predictably occurs in lung adenocarcinoma patients with LKB1 truncating mutations, results in an aberrantly polarized and adhesive cell population that is superior at navigating the microenvironment during invasion.

Supplemental Information

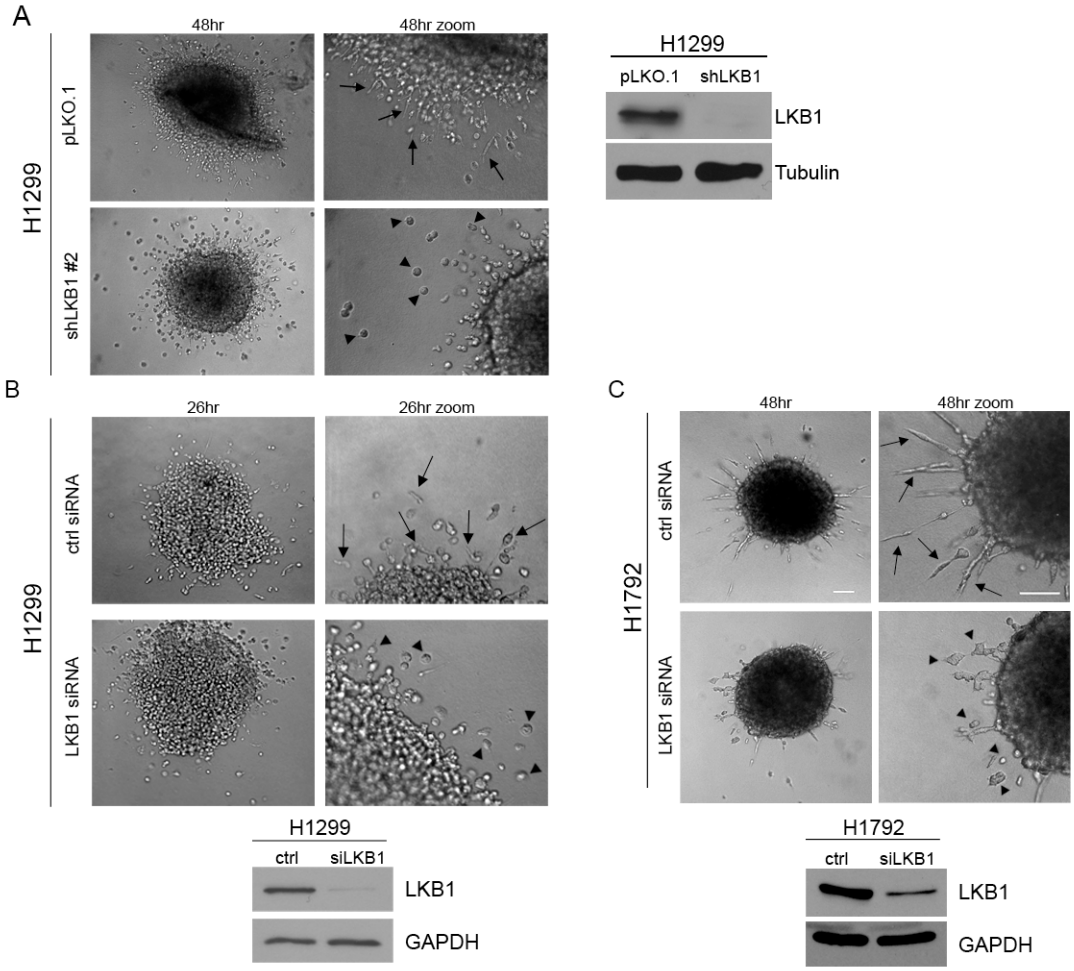

Figure S2.1. Knockdown of LKB1 causes a mesenchymal to amoeboid transition in invasive morphology in H1299 and H1792 cells. (A) LKB1 was stably depleted in H1299 cells using a second targeted shRNA. A spheroid invasion assay was performed and images taken at 48 hours. Western blot confirms LKB1 knockdown. Arrows=mesenchymal cells, arrowheads=amoeboid cells. (B) LKB1 was knocked down via targeted siRNA in H1299 cells with scrambled siRNA used as a control, and a spheroid invasion assay was performed as previously described. Invading cells were imaged after 26 hours of invasion time. Zoomed images are shown to the right. Western blot shows efficiency of the knockdown, and GAPDH was used as a loading control. (C) Experiment from B was repeated in H1792 (LKB1 wildtype) cells. Images were taken at 48 hours post-embedding. Scale = 100 μm.

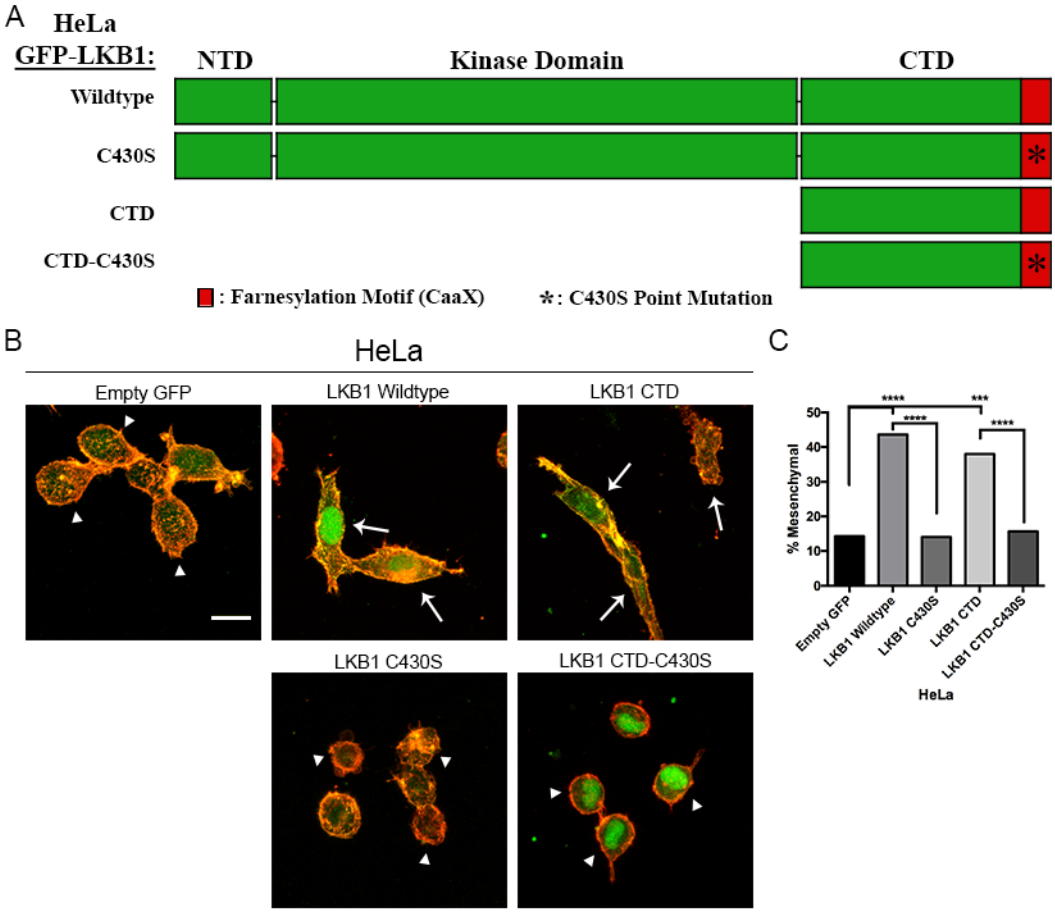

Figure S2.2. LKB1 regulates cellular polarization through its C-Terminal Domain in a farnesylation-dependent manner in HeLa cells. (A) Schematic showing HeLa (cervical carcinoma, LKB1-null) cells that were generated stably expressing GFP-tagged: wildtype LKB1, a C430S mutation to disrupt farnesylation, the C-terminal domain (CTD) alone, or the CTD alone with a C430S mutation. (B) Immunofluorescence of HeLa spheroids embedded in collagen and stained with phalloidin. Amoeboid and mesenchymal morphologies (described in Figure 1) were quantified as a percentage back to the total number of cells invaded in each spheroid. n=4 spheroids. Scale=20 μm. Arrows=mesenchymal cells, arrowheads=amoeboid cells. (C) The percentage of mesenchymal cells was quantified for each cell line at 24 hours post-embedding. ***=p≤0.001, ****=p≤0.0001.

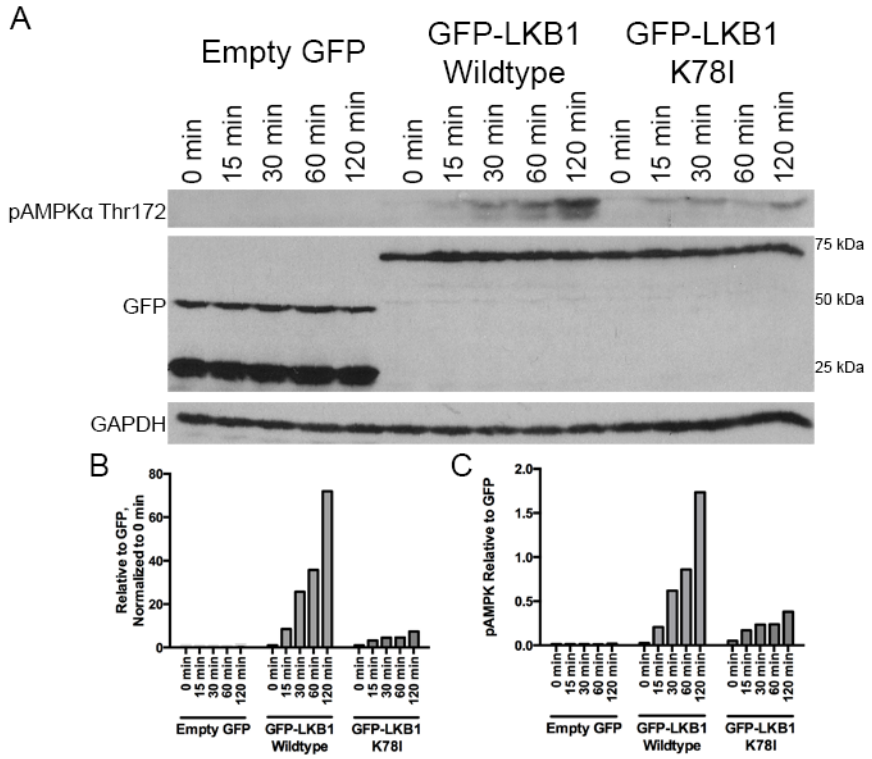

Figure S2.3. LKB1 K78I kinase dead exhibits reduced phosphorylation of AMPK. (A) Western blot show pAMPKα Thr172 expression at 0, 15, 30, 60, and 120 minutes of glucose deprivation in H157 cells re-expressing: Empty GFP, GFP-tagged LKB1 Wildtype, or GFP-tagged LKB1 K78I (kinase-dead mutant). (B) Densitometry of pAMPKα Thr172 normalized to GFP expression. (C) Densitometry of pAMPKα Thr172 normalized to GFP expression and 0 minute time point for each cell line.

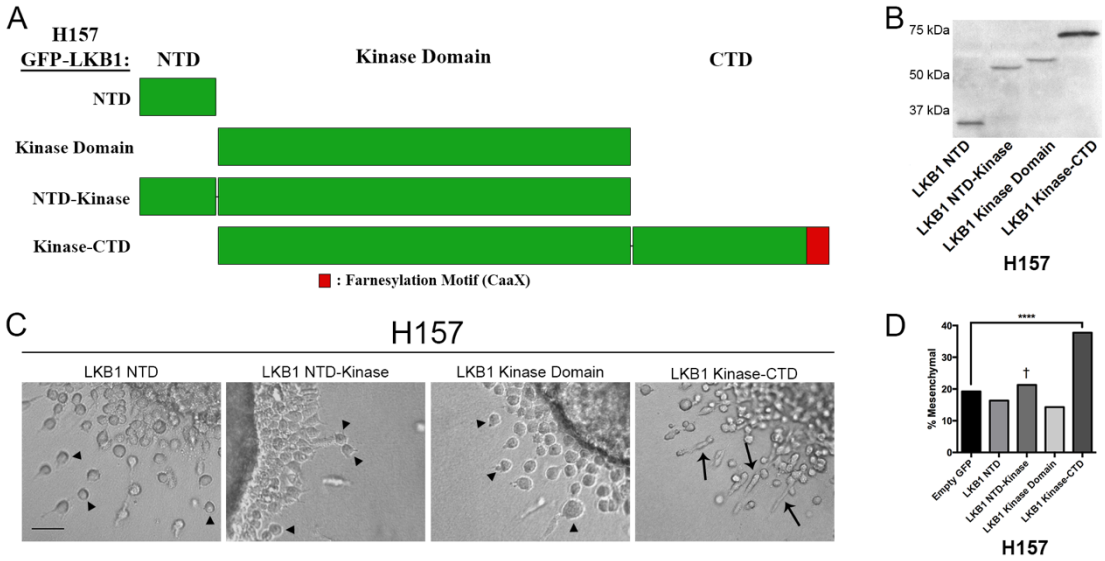

Figure S2.4. The LKB1 N-terminal and kinase domains do not restore mesenchymal polarization. (A) Schematic showing H157 (NSCLC, LKB1-null) cells that were generated stably expressing GFP-tagged: LKB1 N-terminal domain (NTD), N-terminal and kinase domains (NTD-Kinase), kinase domain, and kinase domain and C-terminal domain (Kinase-CTD). (B) Western blot probed with a GFP antibody verifying expression of the H157 stable cells. (C) Bright field imaging of H157 spheroids embedded in collagen. Amoeboid and mesenchymal morphologies were quantified as a percentage back to the total number of cells invaded in each spheroid. n=4 spheroids. Scale=50 μm. Arrows = mesenchymal cells, arrowheads = amoeboid cells. (D) The percentage of mesenchymal cells was quantified for each cell line at 24 hours post-embedding. †: H157 cells re-expressing LKB1 NTD-Kinase largely exhibited collective invasion. Quantifications of polarity are for the cells that invaded individually. ****=p≤0.0001.

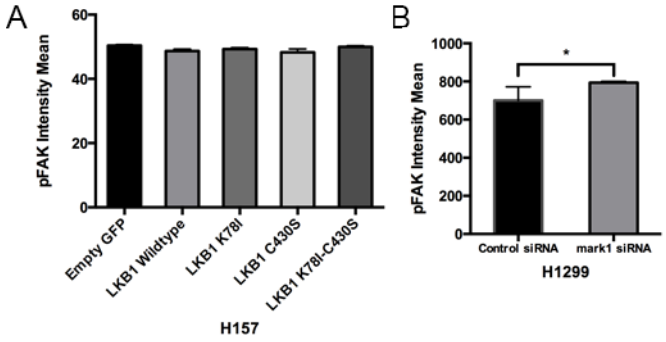

Figure S2.5. Individual pFAK site intensity mean exhibits little to no difference between LKB1 domains and a MARK1 knockdown phenocopying of these domains. (A) Individual pFAK site intensity mean in H157 cells re-expressed empty GFP control or GFP-tagged LKB1. (B) Individual pFAK site intensity mean in H1299 cells after MARK1 knockdown. *=p≤0.05.

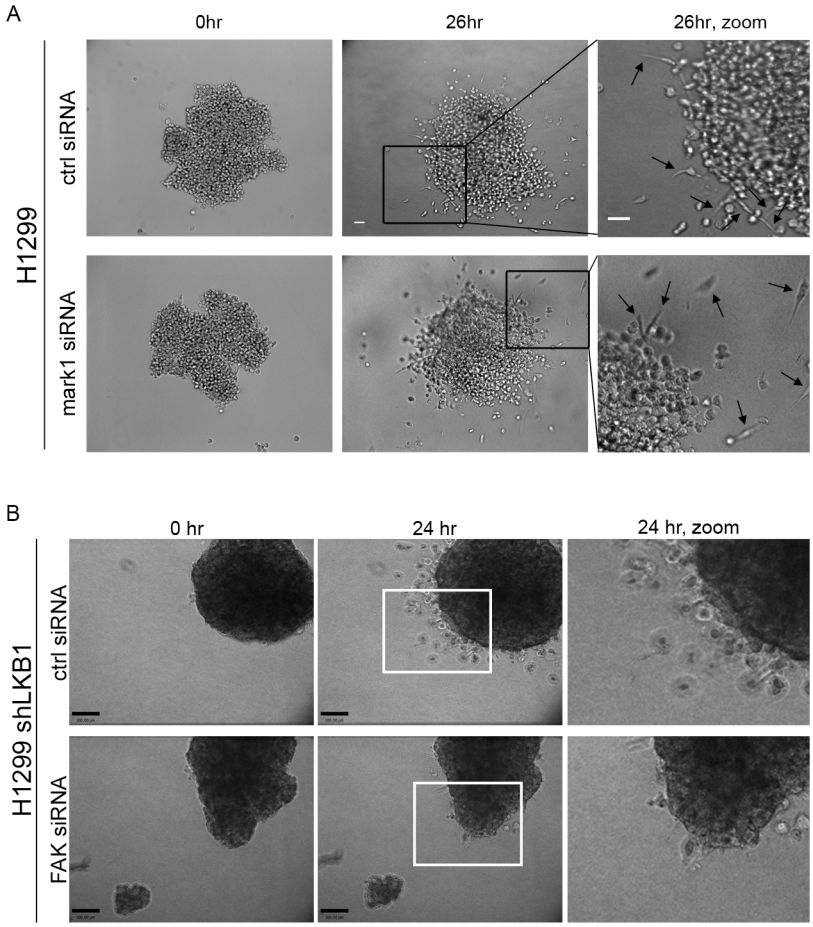

Figure S2.6. MARK1 knockdown does not impact invasive morphology, while FAK knockdown abolishes invasion in LKB1-depleted cells. (A) MARK1 was knocked down via targeted siRNA in H1299 cells (LKB1 wildtype). Scrambled siRNA was used as a negative control. Spheroids formed from control or MARK1 siRNA were embedded in a collagen matrix and imaged after 0 and 26 hours of invasion time. Zoomed images are shown on the right. Scale = 50μm. Arrows = mesenchymal cells. (B) H1299 shLKB1 cells were treated with either scrambled siRNA control or FAK targeted siRNA. Spheroids formed from these cells were then embedded in a collagen matrix and imaged at 0 and 24 hours post-embedding. Zoomed images are shown on the right. Scale = 100 μm.

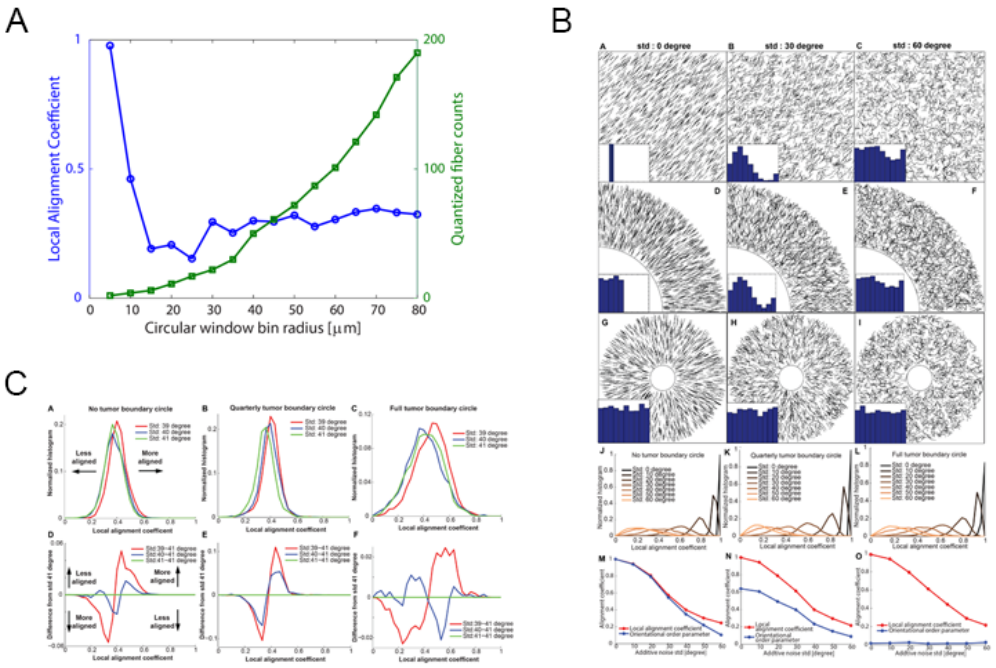

Figure S2.7. Optimization of the local alignment coefficient for quantifying heterogeneous collagen alignment. (A) Local alignment coefficient and quantized fiber counts for various circular window bin using figure 7Biii. (B) For each simulation, a line presents a fiber. Three different test cases were considered: collagen fibers without any tumor boundary (A-C), with a quarterly tumor boundary circle, located at the lower left corner of the box (D-F), with a tumor boundary circle, located at the center of the box (G-I). Inset figures show fiber angular histogram from 0-180^{o} (A-C, G-I) or 0-90^{o} (D-F). Aligned fibers are disturbed by adding an angle, sampled from normal distribution, where the mean is zero and standard deviation (std) is 30^{o} (B, E, H), and 60^{o} (C, F, I). Local alignment coefficient distribution for 7 std degree values for no tumor boundary circle (J), quarterly tumor boundary circle (K), and full tumor boundary circle (L). The average value of the local alignment coefficient and the orientational order parameter value for 7 std degree values for no tumor boundary circle (M), quarterly tumor boundary circle (N), and full tumor boundary circle (O), as compared to the orientational order parameter. (C) The local alignment coefficient distribution for 3 std degree values (39^{o}, 40^{o}, 41^{o}) for no tumor boundary circle (A), quarterly tumor boundary circle (B), and full tumor boundary circle (C). The difference curve of local alignment coefficient distribution from the baseline distribution curve (std 41^{o}) for three different std degree values for no tumor boundary circle (D), quarterly tumor boundary circle (E), and full tumor boundary circle (F).

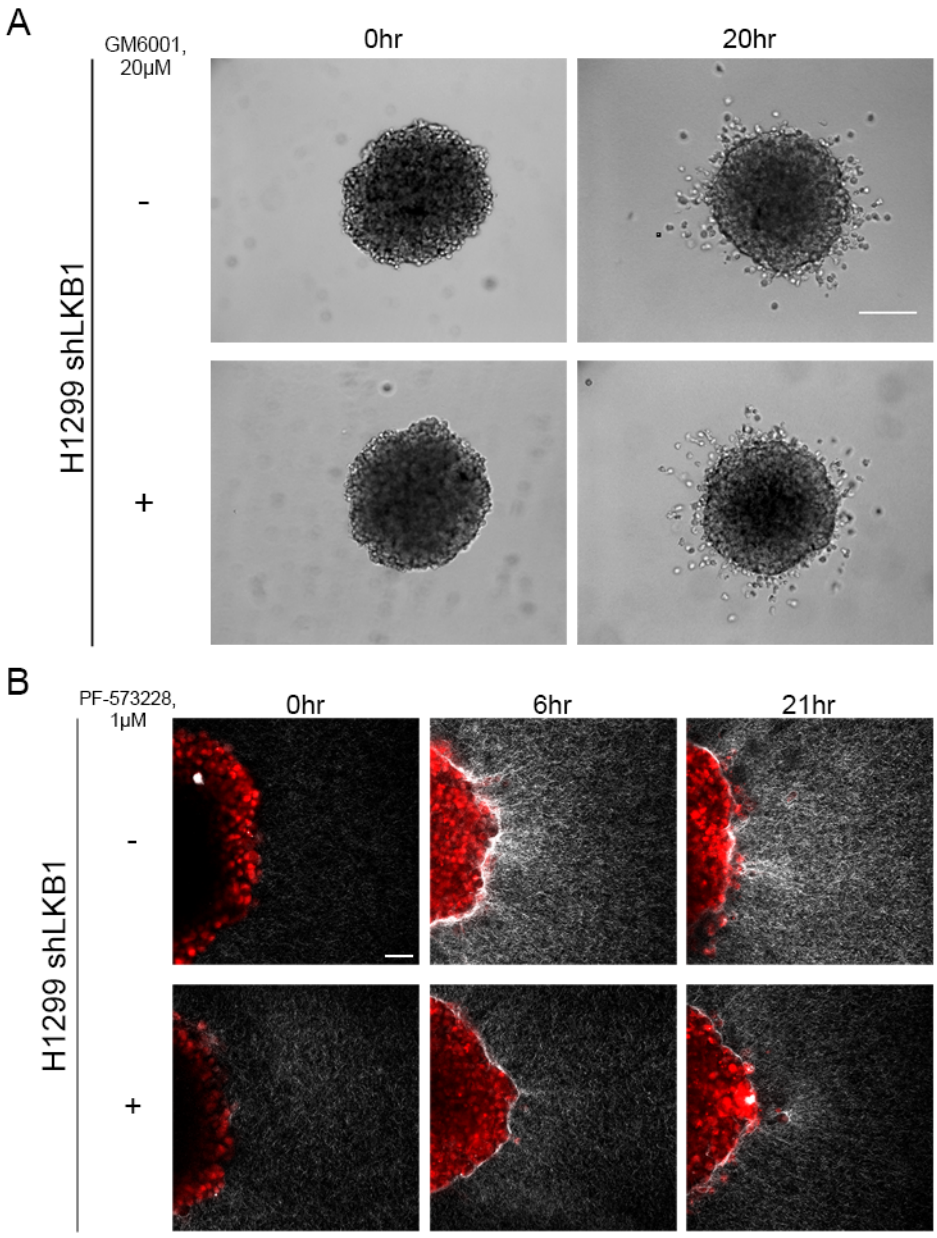

Figure S2.8. Pharmacological inhibition of MMPs does not impact collagen invasion, whereas FAK inhibition decreases collagen remodeling in LKB1-depleted cells. (A) H1299 shLKB1 spheroids were embedded in collagen in the presence of either DMSO control or 20μM of the pan-MMP inhibitor GM6001. Images were taken at 0 and 20 hours post-embedding. Scale = 100μm. (B) H1299 shLKB1 spheroids were dyed with CellTracker Red to visualize invading cells and embedded in a collagen matrix. The spheroids were imaged using SHG imaging in the presence of either DMSO vehicle control or 1μM FAK inhibitor PF-573,228. Images were taken at 0, 6, and 21 hours. Scale=50 μm.

Chapter 3: Developing an image-guided genomics technique to probe the mechanisms regulating collective invasion in lung cancer cells.

Adapted from "Image-guided genomics of phenotypically heterogeneous populations reveals a vascular signaling pathway during symbiotic collective cancer invasion." In Revision, *Nature Communications*.

J. Konen[1], E. Summerbell[1], B. Dwivedi[2], K. Galior[3], Y. Hou[4], L. Rusnak[1], A. Chen[1], J. Saltz[5], W. Zhou[2,6], L.H. Boise[2,6], P. Vertino[2,7], L. Cooper[4], K. Salaita[3], J. Kowalski[2,8], A.I. Marcus[2,6,*]

[1]Graduate Program in Cancer Biology, Emory University, Atlanta, GA
[2]Winship Cancer Institute, Emory University, Atlanta, GA
[3]Department of Chemistry, Emory University, Atlanta, GA
[4]Department of Biomedical Informatics, Emory University, Atlanta, GA
[5]Department of Biomedical Informatics, Stony Brook University, Stony Brook, NY
[6]Department of Hematology and Medical Oncology, Emory University, Atlanta, GA
[7]Department of Radiation Oncology, Emory University, Atlanta, GA
[8]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA
*Corresponding author

Abstract

To probe the phenotypic heterogeneity found in cell populations, we developed an image-guided genomics technique termed spatiotemporal genomic and cellular analysis (SaGA) that allows for precise selection and amplification of living and rare cells. SaGA was used on collectively invading 3-D cancer cell packs to create purified leader and follower cell lines. The leader cell cultures are phenotypically stable and highly invasive in contrast to follower cultures, which show phenotypic plasticity over time and minimally invade in a sheet-like pattern. Genomic and molecular interrogation reveals an atypical VEGF-based vascular signaling machinery that facilitates recruitment of follower cells but not for leader cell motility itself, which instead utilizes focal adhesion kinase-fibronectin signaling. While leader cells provide an escape mechanism for followers, follower cells in turn provide leaders with increased growth and survival. These data support a symbiotic model of collective invasion where different phenotypic cell types cooperate to promote their escape.

3.1 Introduction

A single tumor can harbor distinct genetic and epigenetic cellular sub-populations that drive tumor initiation and progression. This intratumor heterogeneity is proposed to be one of the major confounding factors of treatment causing relapse and poor clinical outcome(117). Genomic instability and epigenetic modifications generate intratumor heterogeneity(118-123) creating distinct genetic and epigenetic sub-populations or clones(121, 124-127). A branched tumor evolutionary architecture can emerge(128), (129) containing the plasticity to progress under harsh environmental conditions and thwart therapeutic attempts to eradicate the tumor (118, 124). It can be argued that until we discover how intratumor heterogeneity can be circumvented, precision oncology initiatives may fall short of expectations(118, 130-132).

Single cell sequencing methodologies(133-135) have improved the genomic, transcriptomic, and epigenomic resolution of clonal tumor populations; however, the phenotypic implications of these alterations remain unclear. This is partly due to experimental challenges and is compounded by phenotypic plasticity that allows cancer cells to adapt to local changes in the microenvironment, without changes to the genome itself (e.g., epithelial to mesenchymal transition(136)). Despite repeated observations that a small number of rare cancer cells or clones, hidden within a larger tumor population can drive tumor growth and spread(127, 137-142), studies linking single cell or clonal phenotypes with genomic data have been limited.

To probe the biology of rare and phenotypically heterogeneous cell populations, single cells or subclones need to be isolated based upon user-defined criteria, instead of a random isolation approach; therefore, we developed a technique to image live cells within a biologically relevant

3-D environment, select a cell or cellular group based upon user-defined criteria, extract the cell(s), and subject the cell(s) to genomic and molecular analyses. In this way, we can purify, amplify, and systematically dissect the biologies of rare cells. This new technique, termed spatiotemporal genomic and cellular analysis (SaGA), was used to dissect the phenotypic heterogeneity of collective cancer cell invasion in a 3-D lung cancer model. These data incorporate the first SaGA-derived leader and follower cell lines to reveal that leader cells utilize an atypical vascular signaling machinery to attract follower cells in invasive cell chains by secreting VEGF. In contrast, follower cells support leader cell growth by increasing their mitotic efficiency. This relationship argues for a cellular symbiosis within the collective invasion pack. Furthermore, these data provide proof of concept that SaGA is a powerful technology for dissecting phenotypic heterogeneity within cancer cell populations.

3.2 Methods

Cell lines and transfections

H1299 and H1792 human NSCLC cells (ATCC, Manassas, VA) were cultured in Roswell Park Memorial Institute (RPMI-1640) media supplemented with 10% fetal bovine serum and 100 units/mL of penicillin/streptomycin, and maintained at 37°C and 5% CO_2.

The gd2PAL-Dendra2 plasmid was obtained from the Gary Bassell lab (Emory University) and was stably transfected into H1299 cells using LT-1 transfection reagent (Mirus), and into H1792 cells using Lipofectamine 2000 (Invitrogen). Geneticin was used to select for Dendra2-expressing cells at 300 μg/ml concentration (H1299 cells) or 400 μg/ml (H1792 cells).

To create RFP-leader cells, RFP was subcloned from Lifeact and inserted into the pBabe-puro vector using BamHI and EcoRI. The mCherry-C1 vector was obtained from the Alexa Mattheyses lab (Emory University). mCherry was subcloned into the pBabe-puro vector using Afe1 and EcoR1-HF enzymes. Phoenix-ampho cells were infected as previously described(143). Puromycin (2μg/ml; EMD Millipore) was used to select cells expressing the mCherry or RFP plasmid, and expression was verified using immunofluorescence. mCherry-leader cells were made to increase brightness of red signal in mixing experiments.

Oligofectamine (Invitrogen) was used to introduce either two different FN1 siRNAs (Thermo Fisher Scientific) or FAK siRNA into leader cells, or KDR siRNA (Sigma) into H1299 or H1792 parental cells. Cells were treated with siRNA for 48 hours, and spheroids were formed after the second day of siRNA treatment and embedded 24-48 hours later.

Reagents and antibodies

Recombinant human VEGFA (R&D, Cat. no. 293-VE) was used for western blotting and live cell imaging experiments to stimulate VEGF signaling. Fibronectin antibody was used for western blotting and immunofluorescence imaging (Abcam, Cat. no. ab2413). Notch1 antibody and VEGF-neutralizing antibody were obtained from R&D Systems (Cat. no. AF3647 and MAB293). Primary antibodies against Dll4 and VEGFA were obtained from Santa Cruz (Cat. no. sc-365429 and sc-152). VE-cadherin, GAPDH, and phospho-FAK^{Y397} primary antibodies were obtained from Cell Signaling (Cat. no. 2158, 2118, and 8556). Vimentin antibody was from Sigma (V6630). E-cadherin and N-cadherin antibodies were obtained from BD Biosciences (Cat. no. 610181 and 610920). Alexa Fluor® 555 and 647 secondary antibodies (Invitrogen) and DAPI (4',6-diamidino-2-phenylindole, Invitrogen) were used for 3-D immunofluorescence. Horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson ImmunoResearch) were used for western blotting.

Western blot

Cellular protein expression was analyzed via western blotting as previously described(143). To analyze media for secreted proteins, cells were plated in serum-free media for 24 hours. Media samples were collected and cell debris eliminated via centrifugation. Proteins were precipitated using acetone, and samples were diluted and boiled in Laemmli sample buffer.

Spheroid formation and invasion assays

Spheroids were generated as previously described (143). Compacted spheroids were collected and resuspended in 2.0mg/ml Matrigel (BD Biosciences). Spheroids were plated in a 35mm glass bottom dish (In Vitro Scientific) and incubated at 37^{o}C overnight. To ensure invasion

occurred in 3-D and not along the glass bottom, the distance between the spheroid and the dish surface was measured and found to be an average of 76μm. Images were taken at 0 and 20-24 hours post-embedding using an Olympus IX51 microscope 4x (0.13 NA air), 10x (0.30 NA air) and 20x (0.45 NA air) with an Infinity2 CCD camera.

For drug treatments, spheroids were formed as described above. FAK inhibitor PF-562271 at 2μM, γ-secretase inhibitor RO4929097 at 10μM, or VEGR2 kinase inhibitor ZM323881 at 10μM, were added directly to the Matrigel during the embedding process, as well as to the growth media added on top of the matrix.

Spheroid microscopy

Fixed cell confocal

H1299 spheroids were fixed and stained for immunofluorescence as previously described(143). Spheroids were imaged with the Leica TCS SP8 inverted confocal microscope (40x oil HC PL APO, 1.30 NA) using 1.0μm z-stack intervals and sequential scanning (405nm DMOD Flexible, 488nm argon, 514nm argon).

Live cell confocal

H1299 spheroids were embedded in Matrigel and imaged using a Perkin Elmer spinning disk confocal microscope at 10x (Plan-Neofluar 0.30 NA) mounted onto a Zeiss Axiovert encased at 37°C with 5% CO_2. Transmitted light images were acquired every 10 minutes for 20 hours using a Hamamatsu Orca ER CCD camera with 2X2 binning. Quantification of chain dynamics was done using Volocity imaging software. H1299 parental, leader, and follower spheroids were imaged using a Leica SP8 inverted confocal microscope with live cell chamber

at 10x (HC Plan Fluotar 0.3 NA). Images were collected every 10 minutes using a 488nm argon laser, beginning ~6 hours post-embedding for 12 hours.

Spheroid image analysis

4D (x,y,z,t) spheroid dynamic images were first projected into 3D (x,y,t) to enhance contrast of dim branches(144). For each timepoint (t) and each position in (x,y) plane, the standard deviation of intensity in all z direction was calculated(144). The projected 3D (x,y,t) image sequences were filtered to remove background noise using Matlab function imgaussfilt3. Filtered images were segmented using 3D graph cuts method(145). The segmented images were polished using Matlab functions imclose and imfill to close gaps and fill holes. The basic morphology features of each spheroid were extracted using Matlab function regionprops. Branch number was quantified using Matlab function bwmorph to generate the skeleton of the spheroid and count the number of skeleton endpoints. The invasive radius was defined as the distance of the furthest point on the spheroid boundary to the centroid.

Invasive area was quantified by measuring both the total spheroid area around the outer perimeter and the inner spheroid core in ImageJ and taking the difference between the two measures. Spheroid circularity was utilized as an indirect measure of sheet-like invasion, and was quantified in ImageJ by measuring the spheroid outer invasive perimeter.

SaGA technique

H1299-Dendra2 or H1792-Dendra2 cells were plated for spheroids, embedded in Matrigel, and incubated overnight as described above. After ~16 hours of invasion, spheroid plates were imaged using the Nikon A1R live cell laser scanning confocal. Spheroids were imaged using

the 10x objective (0.3 NA DIC) and photoconversion was performed at a 3X zoom using the A1R software. The 405nm laser was used to photoconvert cells of interest at laser power 10-30%, which was found to not induce DNA damage in the cells as measured by yH2AX staining in cell nuclei (Fig S2.2B). To extract photoconverted cells, the protease dispase was faster and gentler on cells when compared to trypsin using FACS analysis (Fig S2.2C). Therefore, the Matrigel matrix was degraded and single cell suspension was achieved using dispase I at 1μg/ml with intermittent manual disruption via pipetting. The protease activity was inactivated using media and samples were centrifuged. Upon resuspension, the samples were analyzed via FACS for TexasRed and FITC expression (Fig S2.2D). To improve signal:noise for FACS cell isolation, photoconversion was optimized by varying the 405nm light excitation scan iterations, while considering cell viability post-photoconversion. Low photoconversion efficiency was defined as a red fluorescence signal <300 a.f.u. and high efficiency as >300 a.f.u. (Fig S2.2E, F). Both photoconverted (red) and non-photoconverted (green) populations were isolated from the cell sorter with two levels of gate stringency and imaged with fluorescence microscopy post-sorting to assess purity of red positive cells (P2 population; Fig S2.2G). In the low efficiency condition, low and high gate stringency resulted in a contaminated (non-photoconverted cells present) P2 population (Fig S2.2G, left). Using high efficiency photoconversion and a high gate stringency gave nearly a 100% pure P2 population (Fig S2.2G, right) and this approach was continued throughout. For purified leader or follower cell collections, 30-50 cells were sorted per well and expanded. For microarray analysis, 10 cells per well in triplicate were collected (see below).

Microarray transcriptome studies

Cells were processed using the Ovation® One-Direct System and Encore® Biotin Module

(NuGEN Technologies, Inc., San Carlos, CA). Biotin labeled cDNA was hybridized to the Affymetrix Human Gene ST 2.0 gene expression microarray and further processed on the GeneChip Instrument System for Array Cartridges (Affymetrix, Santa Clara, CA). All steps were carried out according to the manufacturer's protocol. Briefly, RNA from 1-20 cells was reverse transcribed using a proprietary RNA/DNA duplexed primer. ss-cDNA was converted to ds-cDNA and linearly amplified in a single primer isothermal amplification reaction (SPIA). Amplified cDNA was then fragmented and labeled. Biotin labeled cDNA was hybridized to the Human Gene ST 2.0 GeneChip at 45 degrees C for 40 hours. Hybridized microarrays were washed and stained on the Affymetrix GeneChip 450 fluidics station using the appropriate chip dependent fluidics script. Arrays were scanned and intensity data extracted using the Affymetrix 7G scanner and the Command Console software suite.

Microarray Analysis

The raw CEL files for all 10 samples were GC content adjusted, RMA background corrected, log2 transformed, quantile normalized, and mean probset summarized using Partek Genomics Suite v6.6 (PGS; Partek Inc., St. Louis, MO)(146). The probesets were annotated with the Affymetrix Human Gene 2.0 ST annotation file and expression values represented at the gene level.

The impact of photoconversion on gene expression was controlled by comparing noninvasive cells that have been photoconverted (NR) to noninvasive cells that were not photoconverted (IG). Any gene that significantly changed between these two conditions were denoted as alterations due to the photoconversion process and were excluded from analysis. Relative to these controls, significant expression differences between leader and follower cells were tested

with the four following hypotheses: 1) mean gene expression differences between leader vs. follower cells is greater than mean (absolute) expression differences between control (NR and IG) cells; 2) mean expression differences between leader vs. follower cells is greater than zero; 3) mean gene expression differences between followers vs. leaders is greater than mean (absolute) expression differences between control (NR and IG) cells; and 4) mean expression differences between followers vs. leaders is greater than zero. Hypothesis testing was done based on 500 permutations. For each permutation of the data, a t-statistic was defined by taking all pairs of differences among samples. In specific, nine pairs of expression differences between leaders and followers were formed by taking one sample from each of leader and follower cells. Likewise, nine pairs of expression differences between the two defined controls were formed. Using these 9 difference pairs, both a two- and one-sample t-statistic were defined for each gene for testing respective hypotheses 1 and 2; the differences were reversed for testing hypotheses 3 and 4. By permuting the expression difference pairs between groups (leaders and followers differences vs. controls differences), a p-value was estimated for each gene based on comparing the number of times the permuted data exceeded both the one- and two-sample t-statistic formed based on the observed data. Genes with $p<0.05$ were selected as the differentially expressed genes in the leaders (N=788) and followers (N=634). The heatmap of differentially expressed genes was generated using heatmap.2 R function. The biological pathways enriched among these differentially upregulated genes were searched against several curated databases using the functional interaction networks in Reactome FI Cytoscape plugin (147, 148).

Cell line genotyping

H1299 samples were processed according to the ABI AmpFLSTR Identifiler PCR

Amplification Kit protocol and analyzed on the ABI 3130xl Genetic Analyzer according to the manufacturer's directions. Amplicons were electrophoresed with the appropriate allelic ladder on the 3130xl Genetic Analyzer. Identification analysis was performed using GeneMApper ID software version 3.2.1.

Proliferation assays and mitosis analysis

For the proliferation assays, H1299 leader and follower cells were plated in triplicate in a 24-well plate. At days 1-3, cells were counted using an automatic cell counter (BioRad). For mitotic event analyses, H1299 RFP-leader and follower cells were plated in an 8-well LabTek glass bottom slide either alone or in co-culture. After 6 hours, cells were imaged every 5 minutes for 21 hours on the Leica SP8 inverted confocal microscope at 10x using a 488nm argon laser. Mitotic events were analyzed from these images using Leica Application Suite X software. The length of time from prophase to anaphase and anaphase to cytokinesis was determined by morphological features. The beginning of prophase was defined as the first image where the cell became spherical and increased Dendra2 fluorescence, the beginning of anaphase was defined as the first image where the chromosomes were visibly separated and the cell has begun elongating, and cytokinesis was defined as the first image where the two daughter cells are separated by a plasma membrane. The presence of a variety of mitotic defects were defined by morphological abnormalities. Cytokinetic instability was defined as what appears to be initially >2 daughter cells with excessive membrane blebbing and cell shape deformation during cytokinesis, but over time is corrected to 2 daughter cells. Cell death events were counted based on morphological changes consistent with cell death phenotypes (loss of all cell motility and membrane dynamics, shrinkage of cell, nuclear fragmentation, formation of apoptotic bodies, phagocytosis by neighboring cells, etc.).

Cell cycle analysis

H1299 follower and leader cells were plated in 100mm tissue culture dishes. After 24 hours, cells were washed and fresh RMPI-1640 media supplemented with 0% or 10% FBS was added to the cells. After 20 hours, cells were harvested and fixed in 95% ethanol at -20°C. Cells were stored at 4°C for 24 hours before staining with DNA staining buffer (4μg/mL DAPI, 0.25% Triton-X 100 in 1X PBS). DAPI expression was analyzed by flow cytometry on a BD FACSCanto-II cytometer using FACSDiva software. FlowJo software was used to exclude doublets and determine the distribution of cells within G0/G1, S, and G2 peaks.

Colony formation assays

H1299 parental, follower, and leader cells, or H1792 follower and leader cells, were plated in 35mm tissue culture dishes at 500 cells per plate. Cells were grown for 2 weeks, and media (RPMI-1640, 24-hour follower conditioned media, or 24-hour leader conditioned media) was refreshed every 3 days. To create conditioned media, 10×10^4 leader cells or 7.2×10^4 cells were seeded in a 6-well plate so as to reach approximately 70% confluence. After 24 hours, cells were washed twice with 1X PBS and then 1.5mL of RPMI-1640 without FBS was added to each well. After another 24 hours, media was centrifuged to remove cells and debris, and 24-hour conditioned media was added to colony formation assays. After 2-3 weeks, colony formation assays were stained with crystal violet (6% glutaraldehyde, 0.5% crystal violet in 1X PBS) for 30 minutes before rinsing thoroughly with water. Colony surface area and the number of colonies with more than 50 cells were quantified using Fiji imaging software (ImageJ).

Force sensor experiments

Reagents

Potassium phosphate monobasic (≥99.0%) and (3-aminopropyl) trimethoxysilane (97%, APTMS) were purchased from Sigma-Aldrich (St. Louis, MO). Fluorescent dye Alexa647 DIBO alkyne was purchased from Life Technologies (Grand Island, NY). 4-Azido-L-phenylalanine was purchased from Chem-Impex International (Wood Dale, IL). Ni-NTA Agarose (#30210) was purchased from Qiagen (Valencia, CA) P2 gel size exclusion beads were purchased from BioRad (Hercules, CA). Succinimidyl PEG NHS 2000 Da and Lipoic acid PEG NHS 3400 Da were purchased from Nanocs (Boston, Ma). Gold nanoparticles were purchased from nanoComposix (San Diego, CA).

Protein Expression

I27 based construct was designed with N-terminal ligand TVYAVTGRGDSPASSAA and two C-terminal cysteines for immobilization onto AuNPs. The pET22b plasmid encoding an I27 based sensor with a TAG codon was co-transformed with pEVOL-pAzF plasmid into electrocompetent BL21(DE3) *E. coli* cells. Cells were grown at 37 °C in the presence of ampicillin, chloramphenicol, and 0.2% glucose to an optical density (OD) of 0.2, at which 1 mM of 4-azido-L-phenylalanine was added. At an OD of 0.4, L-arabinose was added to a final concentration of 0.02% (w/v) and at an OD of 0.8, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Cells were shaken for 16 h at 30 °C, purified by Ni^{2+} affinity chromatography and stored at -80 °C in 0.1 M potassium phosphate buffer (pH 7.4).

Dye labeling

I27 based protein sensor was incubated with DIBO-A647 for 1 h at 37 °C, followed by

incubation at RT for 24 hours. The sensor was next purified using P2 gel size exclusion beads and the labeling ratio was quantified by UV-Vis absorption (NanoDrop).

AuNP Surface Preparation

Glass coverslips were piranha etched for 30 min, functionalized with an APTMS solution in acetone for 1 h and thermally annealed at 80 °C for 20 min. Subsequently, the surfaces were passivated with 5% (w/v) mPEG-NHS and 0.5% (w/v) lipoic acid-PEG-NHS in 0.1 M fresh sodium bicarbonate overnight at 4 °C. After passivation, 12 nM of AuNPs (diameter = 9 nm) were incubated onto the surface for 20 min.

Statistical Analysis

A two-tailed unpaired Student's *t*-tests was used to analyze statistical significance between two conditions in an experiment. For experiments with three or more comparisons, an ordinary one-way ANOVA with a Tukey's multiple comparisons test was used. Significance was assigned to p values<0.05. Error bars represent the mean ± SEM.

3.3 Results

Leader cells are a unique and essential invasive subpopulation during 3-D invasion

H1299 non-small cell lung cancer (NSCLC) tumor spheroids were embedded in a 3-D matrix and imaged over time (Fig S3.1). Invading cells displayed phenotypically heterogeneous, collective chain invasion with leader cells defined as the first cell of a chain with trailing follower cells (Fig 3.1). Upon leader cell detachment, the chain did not progress further, and this lack of plasticity was observed in 70% of all observed cases of leader cell detachment (Fig S3.1). Quantitation of invasive chain dynamics pre- and post-leader cell loss show that the distance traveled, displacement, and chain velocity significantly decreased upon leader cell detachment (Fig 3.1). A single chain plot over time demonstrates that invasion distance plateaus after leader cell detachment, and once leader cells detach, they attempt to return to the chain (Fig 3.1). Similar findings were observed in a second NSCLC line, H1792; however, follower chain progression occurs post-leader cell detachment but in the direction of the leader cells, suggesting that the chain is attempting to re-attach (Fig S3.1). Taken together, these data suggest that the leader cell is a specialized and essential cell type existing within a phenotypically heterogeneous cancer cell population.

Development and optimization of spatiotemporal genomic and cellular analysis technology

To probe phenotypic heterogeneity, we developed a technique termed spatiotemporal genomic and cellular analysis (SaGA) that allows for precise selection of user-defined living cells within a dynamic environment (Fig 3.2). H1299 or H1792 lung cancer cells were stably

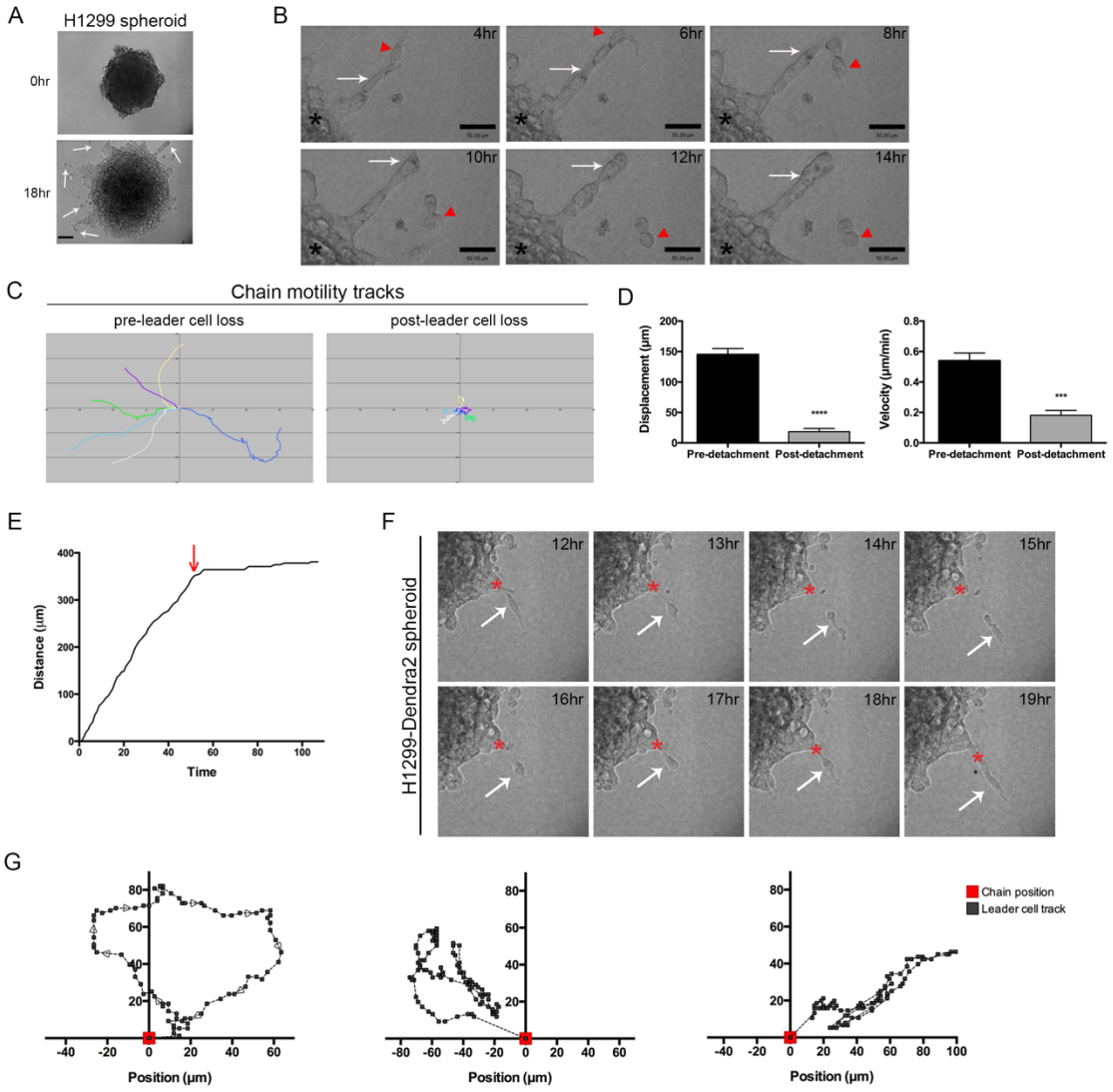

Figure 3.1. Leader cells represent a specialized subpopulation within the collective invasion pack. A. H1299 spheroids were imaged at 0 hours (top) and 18 hours (bottom) post-embedding. Arrows = leader-follower invasive chains. B-F) H1299 spheroids were imaged using live cell confocal imaging. B. Still images of live cell imaging of H1299 spheroids taken every 2 hours. Arrow = follower cells in invasive chain, arrowhead = leader cell. C. Invasive chains were tracked over time while intact (pre-leader cell loss) and after a leader cell becomes detached from the chain (post-leader cell loss). Cell track plots are shown. Colors represent a single chain. D. Quantification of track displacement and velocity of invasive chains from C. *$p<0.05$, ***$p<0.001$. E. Single invasive chain distance tracked over time. Arrow = point of leader cell detachment. F. Representative still images of live cell imaging of H1299 spheroids taken every hour. Arrow denotes the leader cell, asterisk denotes the follower cell chain. G. Three leader cells were tracked after the time of detachment from the rest of the chain. The position of each leader cell (gray track) is plotted relative to the position of the chain from which it detached (red point).

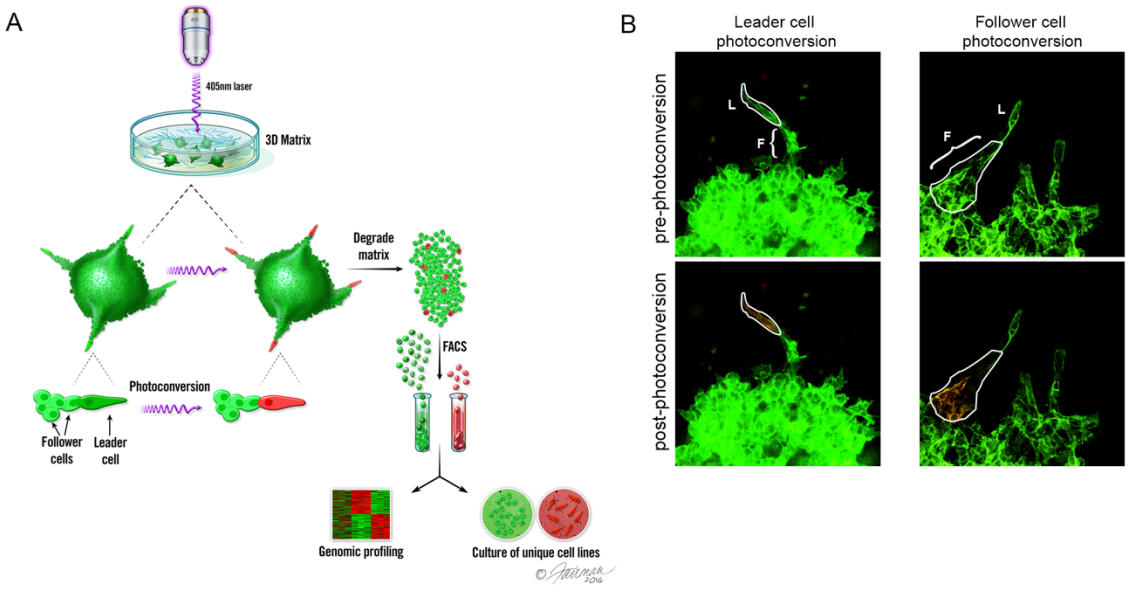

Figure 3.2. Overview of the spatiotemporal genomic and cellular analysis (SaGA) technique for probing heterogeneous cell phenotypes. A. Schematic showing the process of the SaGA technique (see methodology for details). B. Photoconversion examples using 3-D spheroids of H1299-Dendra2 cells. L = leader cell, F = follower cells.

transfected with Dendra2, a photoconvertible fluorophore with a plasma membrane-targeting palmitoylation tag (149), allowing us to define individual cells during imaging. Before photoconversion, all cells have green fluorescence (Fig 3.2, Fig S3.2, top) but upon user-defined exposure to 405nm laser, only the selected cell is photoconverted to emit red fluorescence (Fig 3.2, Fig S3.2, bottom). This process was tested in 3-D models, where a single leader cell or group of follower cells were photoconverted without any measurable photoconversion in neighboring cells (Fig 3.2) and without observable DNA damage (Fig S3.2). SaGA steps were then optimized as described in the Methods to specifically target, extract, and amplify purified leader and follower cells from a 3-D microenvironment.

SaGA-derived leader cells show a partial EMT and maintain their invasive potential during culture

SaGA was used to generate the first purified leader and follower cell lines from the parental line (H1299 or H1792 lung cancer cells) in 3-D collectively invading spheroids. Single nucleotide polymorphism (SNP) analysis verified that leader and follower lines created with SaGA originated from the H1299 parent line and not a contaminating cell type (Table S1). Follower cells had an epithelial-like morphology in 2-D culture, whereas leader cells were mesenchymal-like in shape (Fig 3.3, Fig S3.3). EMT marker assessment showed that leader cells have increased staining of the mesenchymal protein vimentin in 3-D spheroids compared to follower cells, which have little to no vimentin expression even in the few invasive cells (Fig 3.3). In contrast, the expression of the mesenchymal marker, N-cadherin, was decreased in leaders as compared to followers (Fig. 3.3), suggesting a partial leader-cell EMT. Both follower and leader cells are negative for the epithelial marker E-cadherin (Fig 3.3), consistent with H1299 cells. These data suggest that while leader cells have a more mesenchymal morphology

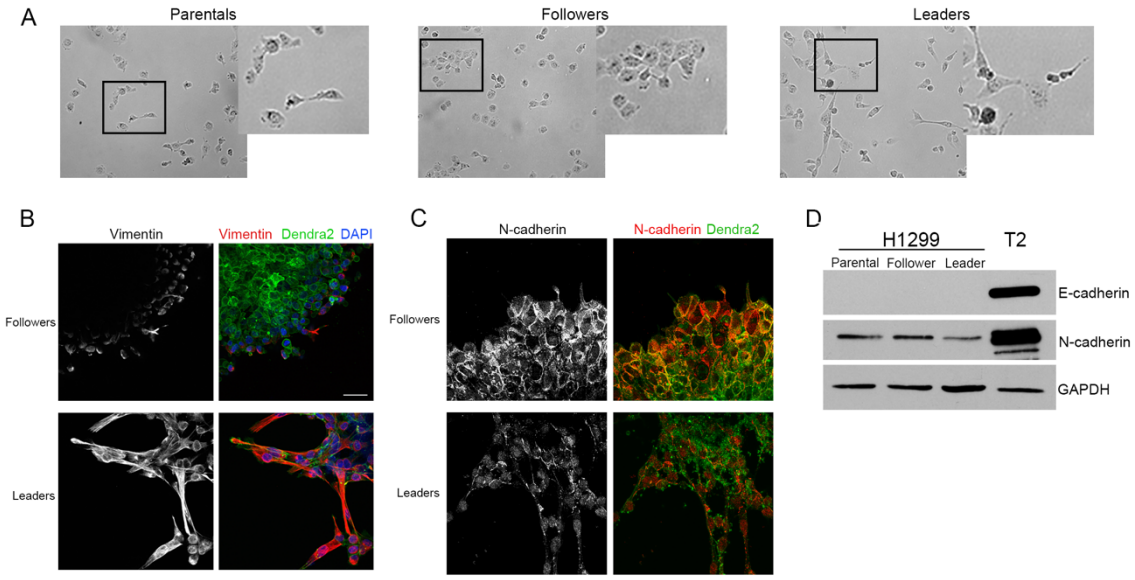

Figure 3.3. Purified leader cells show a partial EMT phenotype. A. Images of H1299 follower and leader purified cells as compared to the parental line. Zoomed images of each cell subtype are shown to the right. B. Immunofluorescence staining of vimentin (red) in follower and leader purified spheroids. All cells express Dendra2 (green) and DAPI (blue) was used to stain cell nuclei. Scale bar = 50μm. C. Immunofluorescence imaging of N-cadherin in purified leader and follower spheroids. Scale bar = 50μm. D. Western blot of H1299 parental, follower, and leader cells for E-cadherin and N-cadherin. The T2 mouse tumor derived cells were used a positive control for E-cadherin. GAPDH was a loading control.

and are vimentin positive, the traditional EMT signature alone cannot be utilized to identify leader cells.

The 3-D invasive properties of SaGA-derived leader and follower cells were assessed and leader cell spheroids show significantly more invasion over time than follower and parental spheroids (Fig. 3.4). Leader cell invasion resembles a network of interlinked cells when compared to sheet-like follower cell invasion (defined as invasion without the presence of leader-follower chains) (Fig 3.4), as well as parental cells, which are more chain-like. Leader and follower invasion showed similar patterns in H1792 purified spheroids (Fig S3.3). Quantitative analysis of spheroid invasive area and branch number over time show both are significantly increased in H1299 leader spheroids compared to follower spheroids (Fig 3.4). Leader cells maintain this invasive morphology and pattern in purified culture whereas follower cells revert back to parental-type invasion after 1-2 months in culture (not shown).

Leader cells promote leader-follower chain invasion

To characterize leader-follower chains, leader cells were added to follower cell spheroids at 1%, 10%, and 50% of total cell content. Follower cells show little to no invasion (Fig 3.5), and what little invasion does occur if a sheet-like pattern as previously seen (Fig 3.4); however, adding 1% leader cells to the follower cell spheroid restored leader-follower chains that are morphologically similar to the parental line, and this trend continued in a leader cell dose-dependent manner (Fig 3.5). The addition of 10% or 50% leader cells significantly increased the chain number and total invasive area (Fig 3.5). To test if leader cells retain leadership in these mixing experiments, RFP-expressing leader cells were made to track leader cells. RFP-leader cells are found at the leading tip of invasive chains 85% of the time (Fig 3.5). Leader

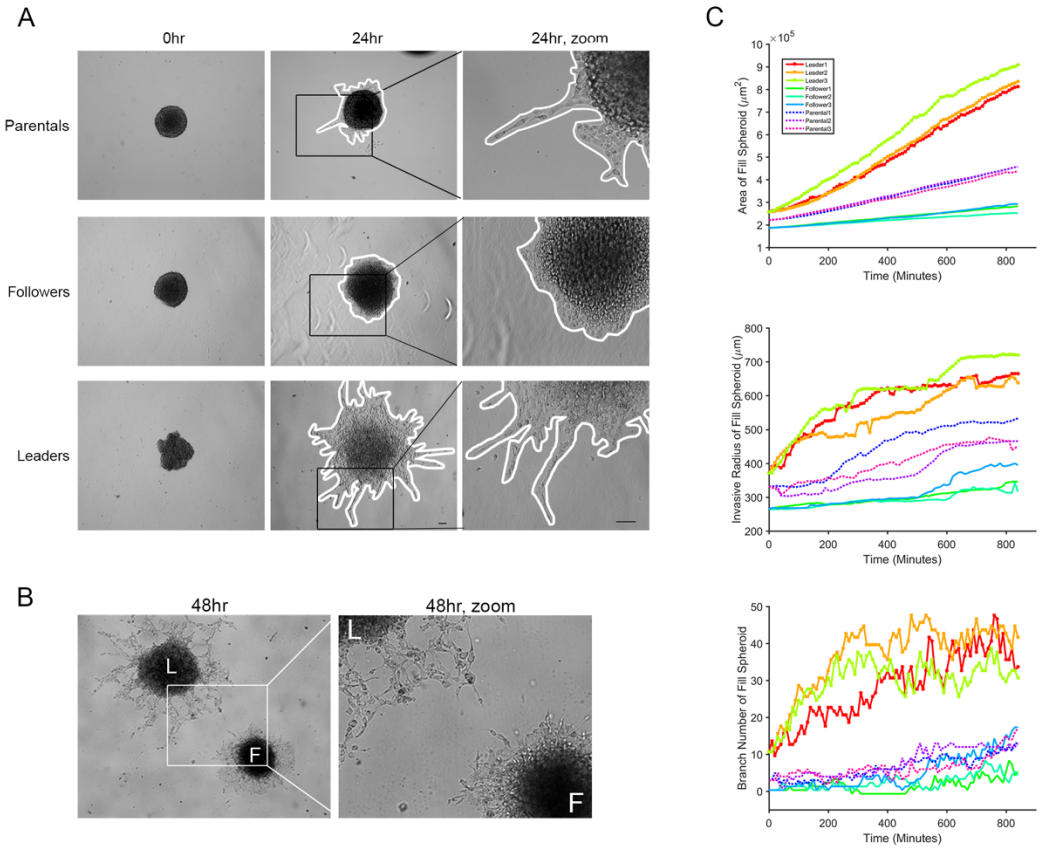

Figure 3.4. Leader cells maintain their invasive phenotype even when cultured as a purified population, suggesting they are a stable phenotype. A. Spheroid invasion assays comparing invasion at 24 hours in H1299 parental, follower, and leader spheroids. Zoomed images are shown to the right. Scale bar = 50μm. B. Leader (L) and follower (F) spheroids were embedded in proximity to each other in the same matrix. Images were taken at 48 hours post-embedding, with a zoomed image to the right. C. H1299-Dendra2 parental, follower, and leader spheroids were embedded and imaged every 10 min for 14h using live cell confocal microscopy. Fluorescence images in both time and z were extracted as individual tiffs and analyzed for area, invasive radius, and branch number. Three spheroids were analyzed per condition.

cells were rarely observed in a non-leading position within the invasive chain, even when leader cells comprise 50% of the entire population, suggesting either that follower cells are selected for or leader cells are excluded from non-leading positions.

We analyzed if leader cells could promote follower cell motility using 2-D motility assays, since single cell motility and cell-cell interactions could be more easily visualized. Co-culture experiments were performed by mixing follower cells with RFP-leader cells, and cell motility was quantitatively assessed. Follower cells plated alone show active lamellipodia dynamics but had limited net movement (Fig 3.5). Adding ~30% leader cells to the follower cell population significantly increased follower cell motility (Fig 3.5). A similar set of experiments were performed to determine if this increased follower cell motility is due to leader cells themselves or a factor secreted by leader cells. Conditioned media from leader cells was sufficient to promote the motility of follower cells (Fig 3.5). Together, these data indicate that leader cells can stimulate the motility of follower cells via a secreted factor.

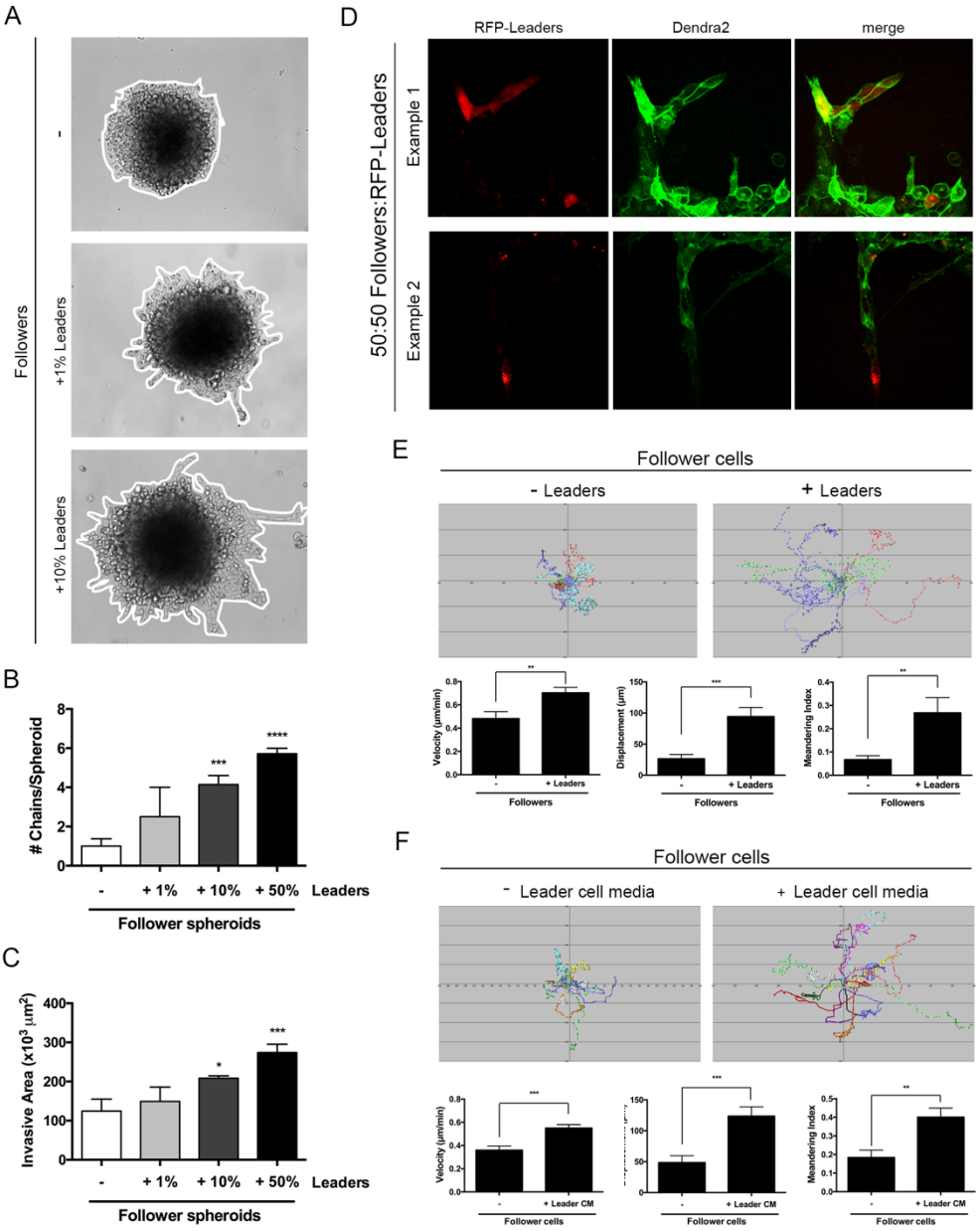

Figure 3.5. Leader cells promote the motility and invasion of follower cells through a secreted factor. A. Spheroids were formed from follower cells either alone (-) or in combination with increasing amounts of leader cells. Images were taken after 24 hours of invasion. B. Quantification of the number of chains per spheroid in the experiment from A. ***p<0.001, ****p<0.0001. C. Quantification of invasive area from the experiment in A. *p<0.05, ***p<0.001. D. RFP-expressing leader cells were mixed with follower cells at a 50:50 ratio. Spheroids from this mixture were embedded in Matrigel and imaged after 20 hours of invasion. E. Follower cells were plated in 2-D either alone or in co-culture with RFP-leader cells. Images were taken every 10 minutes using a spinning disc confocal. Single follower cells were tracked over time. Cell track plots as well as quantification of these tracks are shown. **p<0.01, ***p<0.001. F. Follower cells were plated in 2-D in either their own or in leader cell media that had been conditioned for 24 hours. Follower cell tracks are plotted and quantified below. **p<0.01, ***p<0.001.

Leader-follower invasion represents a VEGF-dependent vascular mimicry

To probe the underlying biological mechanisms that drive leader and follower cell biology and communication, transcriptome profiling was performed. There were 788 candidate transcripts that were upregulated in leader cells compared to follower cells (Fig S3.4; $p<0.025$)) and 684 transcripts upregulated in follower cells compared to leader cells (Fig S3.4; $p<0.025$). These results incorporate controls for the SaGA photoconversion approach (see methodology) and therefore all significant transcripts in leaders vs. follower must not have significantly changed between control samples (Fig S3.4). Functional interaction networks revealed several significant networks related to VEGF, focal adhesion signaling, and RNA Pol II transcription (Fig 3.6, S3.4) that vary in leader cells compared to follower cells. Specifically, VEGF signaling transcripts were significantly increased in leader cells compared to follower cells, whereas focal adhesion signaling transcripts were more heterogeneous with some significantly increased or decreased in leader compared to follower cells.

VEGF signaling was particularly interesting since morphologic patterning of endothelial cells during vascular sprouting has leader-like tip cells with follower-like stalk cells, and Fig 3.4 supports the concept of a secreted factor stimulating follower cell movement. VEGF secretion was found to be upregulated in leader cells compared to parental and follower cells (Fig 3.6, S3.5) with the $VEGF_{165}$ isoform the most abundant (Fig 3.6). Next, we probed if blocking VEGF signaling could impact leader cell influence on follower motility and invasion. An inhibitory anti-VEGF antibody was added to leader cell conditioned media that bathed 2-D follower cell cultures, which was sufficient to inhibit leader cell stimulation of follower cell motility (Fig 3.6). To determine if this is observed during 3-D collective invasion, the anti-VEGF antibody was added to mixed spheroids with 90% followers and 10% mCherry-leaders.

In control treated cells, mCherry-leader cells were observed at the tip of 80% of invasive chains, whereas anti-VEGF treatment abolishes chain formation, and the percentage of chains positive for mCherry-leaders decreased to 20% (Fig 3.6). Chain invasion was significantly reduced in parental spheroids with anti-VEGF treatment, knockdown of VEGFR2 (*KDR)*, or treatment with the VEGFR2 kinase inhibitor, ZM323881 (Fig 3.6, S3.5). However, leader cell invasion itself was not dependent on VEGF signaling, as the total number of leader cells found in the entire invasive area (i.e. independent of location within the invasive area) was not significantly reduced in the mixed spheroids with VEGF neutralization (Fig 3.6). Additionally, purified leader spheroids remained highly invasive with anti-VEGF treatment (Fig 3.6). We hypothesized this may be due to expression of VEGFR1 decoy receptor, since leader cells had increased levels of VEGFR1 mRNA (*Flt1*; Fig. 3.6, S3.4). Leader cells have significantly more VEGFR1 expression when compared to follower cells (Fig 3.6, S3.5), supporting the concept that VEGFR1 decoy receptor could dampen VEGF autocrine signaling in leader cells.

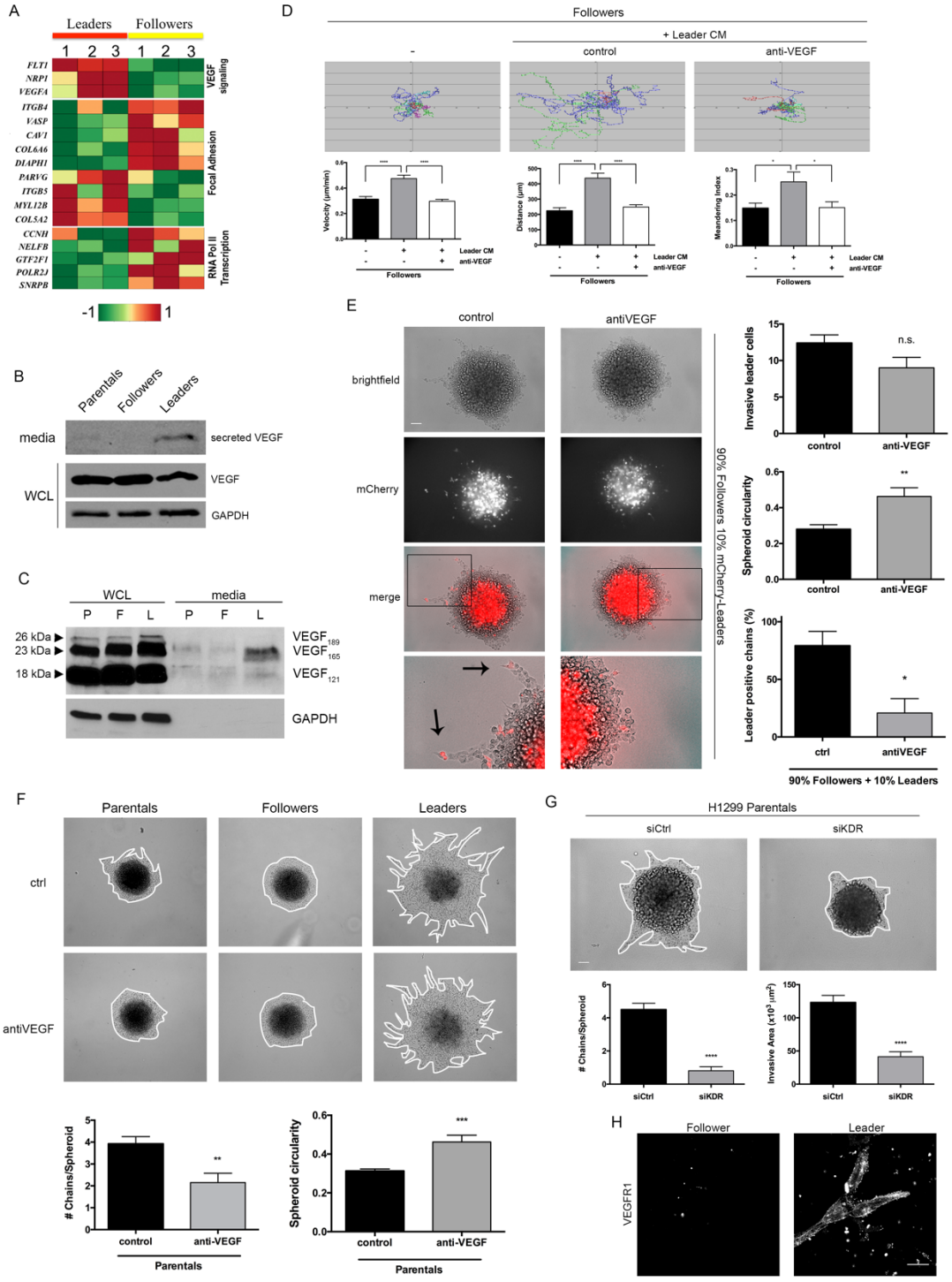

Figure 3.6. Gene and protein expression analyses reveal the VEGF pathway as significantly enriched in leader cells as compared to followers. A. Hierarchical clustering of differentially expressed genes involved in three major pathways: VEGF, focal adhesion, and RNA Polymerase II transcriptional regulation. The leader and follower experimental replicates are colored in red and yellow, respectively. The hierarchical clustering was performed using Pearson correlation distance measure and average linkage method. B. Western blot showing

VEGF expression in media and whole cell lysate (WCL) samples from H1299 parental, follower, and leader cells. GAPDH was used as a loading control. C. Western blot of various VEGF isoforms in H1299 parental (P), follower (F), and leader (L) whole cell lysates and media samples. GAPDH was used as a loading control. D. Live cell tracking analysis of follower cells plated in 2-D either alone (-), in leader conditioned media (CM), or leader CM pre-treated with a VEGF-neutralizing antibody (anti-VEGF). Cell track plots from each condition are shown, and bar graphs generated from these tracks are shown below. *$p<0.01$, ***$p<0.001$. E. The VEGF-neutralizing antibody in 3-D spheroid invasion assays. Mixed spheroids were generated using 90% follower cells, 10% RFP-positive leader cells. Spheroids were embedded in Matrigel with either the anti-VEGF antibody or vehicle control. Zoomed images are shown below. Arrows denote RFP-leader positive invasive chains. The total number of invasive leader cells, the spheroid circularity as an indirect measure of sheet-like invasion, and the percentage of leader-positive chains in each condition were quantified and graphed to the right. Scale bar = 100μm. *$p<0.05$, **$p<0.01$. F. Spheroid invasion assay of H1299 parental, follower, and leader spheroids in the presence of the anti-VEGF antibody as compared to vehicle control. The number of chains per spheroid and spheroid circularity for the parental spheroids were quantified and shown graphed below. n = 12-13 spheroids. **$p<0.01$. G. Spheroid invasion assay on H1299 parentals treated with either VEGFR2 (*KDR*) siRNA or scrambled siRNA control. The number of chains/spheroid and invasive area are graphed below. n = 10 spheroids. ****$p<0.0001$. H. Immunofluorescence imaging of VEGFR1 in follower and leader spheroids. Scale bar = 20μm.

We probed other mechanisms controlling tip-stalk cell maintenance during vasculogenesis and found that VE-cadherin, the major cadherin that links endothelial cells during vasculogenesis, is highly expressed in leader cells but not in follower cells (Fig S3.6), suggesting that leader-follower cell-cell contacts are maintained by VE-cadherin but follower-follower cell-cell contacts are not. We also focused on canonical Notch-Dll4 signaling (150) during collective invasion. Notch protein is highly expressed in follower cells but not in leader cells, while Dll4 is expressed in leader but not follower cells (Fig S3.6). This mimics tip-stalk cell expression patterning during vasculogenesis, further supporting the concept of a vascular mimicry during collective cell invasion. To probe if Notch1 acts as a leader cell suppressing signal, similar to its role in regulating tip cell formation in vasculogenesis, we utilized a γ-secretase inhibitor RO4929097, which inhibits the ability of γ-secretase to cleave of Notch1 and release the intracellular domain important for its transcriptional activity (Fig S3.6). Surprisingly, inhibition of Notch1 signaling via RO4929097 treatment did not promote the leader cell phenotype as seen previously in other models (151), but instead, blocked chain invasion in H1299 parental spheroids (Fig S3.6). We hypothesized that the role of Notch1 in follower cells in our system must differ from its canonical role in tip cell formation and we therefore tested whether follower cells may utilize Notch signaling to promote proliferation. This result showed that treatment with RO4929097 significantly decreased follower cell growth (Fig S3.6). These data taken together suggest that leader-follower cells utilize a non-canonical vascular mimicry program to promote formation of the invasion chain.

Fibronectin-FAK signaling drives leader cell invasion

Transcriptome data showed significant changes in cell adhesion pathways between leader and follower cells; therefore, activated focal adhesion kinase (FAK) was assessed in leader cells

compared to parental and follower cells. Leader cells show increased pFAK^{Y397} at the leading edge compared to followers, with larger and more elongated adhesion sites (Fig 3.7). We reasoned that enlarged pFAK-positive adhesion sites may help leader cells generate traction force during migration. To test this, leaders were plated onto a glass slide decorated with molecular tension probes that quantitatively map integrin forces using fluorescence microscopy (152) (Figure S3.7). Two hours after plating on the sensor, leader cells showed a positive tension signal at the sites associated with focal adhesions, thus indicating integrin-ligand forces that exceed 36 pN shortly after adhering (153) (Fig 3.7). In contrast, follower cells do not generate sufficient tension to unfold the probe within this time frame. To determine if this enhanced integrin-force is mediated through FAK and whether it is important for 3-D invasion, leader cell spheroids were treated with the FAK inhibitor, PF-562271. As compared to DMSO control, leader cells treated with FAK inhibitor had significantly reduced integrin force as well as reduced spheroid invasion (Fig 3.7). Interestingly, FAK inhibition in the follower cells promoted sheet-like invasion (Fig 3.7), suggesting differing functions of FAK in leader and follower populations.

To probe FAK-based signaling in leader cells, fibronectin was assessed since it is a major ligand for the integrin-FAK pathway. Leader cells had significantly more cellular and secreted fibronectin than parental and follower cells (Fig 3.7, S3.5). Leader spheroids had a vast 3-D fibronectin network but nearly no fibronectin was observed in follower spheroids (Fig. 3.7, Fig S3.5). Fibronectin siRNA depletion abolished leader cell invasion (Fig 3.7, S3.7), similar to pharmacological FAK inhibition, showing that this pathway is necessary for leader cell movement. Additionally, depleting FAK in leader cells prior to mixing in spheroids with follower cells reduces the number of leader-positive chains per spheroid (Fig S3.7), supporting

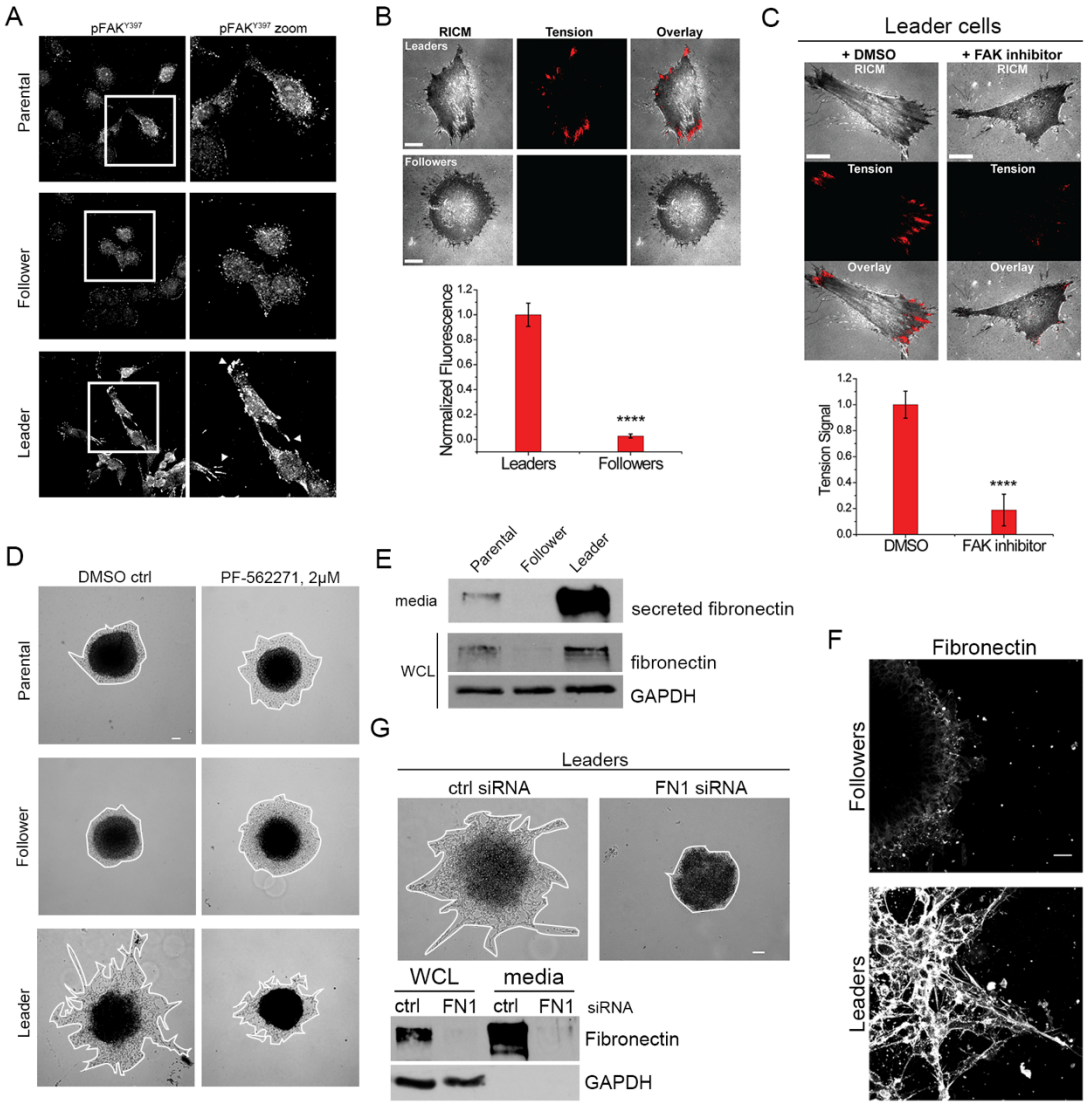

Figure 3.7. The fibronectin/FAK pathway drives leader cell invasion. A. Immunofluorescence imaging of pFAK^{Y397} in H1299 parental, follower, and leader cells plated in 2-D. Arrowheads denote large adhesion sites in leader cells. B. Representative reflection interference contrast microscopy (RICM) and integrin tension images of leader and follower cells as described in Supplementary information (S5)(153). An overlay of the RICM and integrin tension channels demonstrates that regions at the cell edge associated with FAs generate sufficient tension to unfold the titin-based probe for the leader cells. Plot below shows the average integrin tension signal for leaders and followers. Error bar represents the standard error of the mean (SEM) from 10 cells for each group collected from three chambers, ****$p<0.0001$. Scale bar = 10μm. C. Representative RICM and integrin tension images for leader cells incubated in the presence of either 0.1% DMSO control and PF-562271 FAK inhibitor. Error bar represents the standard error of the mean (SEM) from 10 cells for each group collected from three chambers, ****$p<0.0001$. Scale bar = 10μm. D. Spheroid invasion

assay with H1299 parental, follower, or leader cells in the presence of 2μM PF-562271 FAK inhibitor or DMSO control. Scale bar = 100μm. E. Western blot of fibronectin protein in the media or whole cell lysates (WCL) of H1299 parental, follower, and leader cells. GAPDH was a loading control. F. Immunofluorescence of extracellular fibronectin in follower and leader spheroids. Scale bar = 50μm. G. Fibronectin (*FN1*) was knocked down in leader cells using targeted siRNA. Scrambled siRNA was used as a negative control. Spheroids from *FN1* siRNA treated leader cells were assayed for invasion compared to control. Western blot analysis below confirms knockdown efficiency. Scale bar = 100μm.

the concept that FAK is required for leader cells to generate invasive chain motility. Taken together, these data show that leader cells depend upon fibronectin-FAK adhesion signaling to create force during invasion into the microenvironment.

Follower cells are highly proliferative and rescue leader cell growth and mitotic defects

The question remained as to why leader cells invade with follower cells, since purified leader cells are fully competent to invade alone. We hypothesized that follower cells may provide a benefit to invading leader cells. During the initial collection and expansion of leader and follower clones, we observed that leader cells grew at slower rates compared to follower cells (not shown); therefore, to test if follower cells are more proliferative, a basic proliferation assay was performed, showing that follower cells had increased cell counts after 3 days compared to leader cells (Fig 3.8, Fig S3.8). Similarly, colony formation assays showed that leader cells have little colony growth over time, whereas followers have a greater number of large colonies (Fig 3.8, Fig S3.8). Cell cycle analysis 20 hours post-serum starvation showed a large G1 population in leader cells compared to follower cells (Fig. 3.8, S3.8); however, without serum starvation there are no differences in the cell cycle between these two populations (Fig S3.8).

To test the hypothesis that followers provide a growth or survival advantage to leader cells, leader cells were subjected to a colony formation assay in the presence of leader cell conditioned media (LCM) or follower cell conditioned media (FCM). Leader cells in LCM had low colony formation as measured by colony number and area (Fig 3.8); in contrast, adding FCM to leaders resulted in a significantly increased colony number and area. Strikingly, the addition of LCM to follower cells significantly inhibited colony growth as compared to followers grown in their own conditioned media (Fig 3.8). Taken together, these data show

that FCM can significantly increase colony growth of leader cells, whereas LCM inhibits colony growth of follower cells.

To probe why leader cells have growth defects, live cell imaging was used to assess mitosis in purified populations. Leader cells had a variety of mitotic defects compared to follower cells (Fig 3.9), with the most prominent being cytokinetic instability (defined as initially having >2 daughter cells with excessive membrane blebbing and cell shape deformation during cytokinesis but over time corrected to 2 daughter cells; Fig 3.9). Other defects include increased time from prophase through anaphase and eventual fusion of daughter cells. Overall, ~70% of all leader cells have mitotic defects, as compared to only 6% in follower cells (Fig. 3.9).

To test if leader cell mitotic defects could be rescued by follower cells, follower cells were co-cultured with RFP-leader cells, and leader cell mitotic defects were nearly abolished (Fig 3.9). Co-culture with follower cells specifically rescued the prophase-to-anaphase mitotic delay observed in leader cells and cytokinetic instability (Fig 3.9). A similar effect on mitotic defects was observed using FCM on leader cell cultures where unsuccessful mitotic defects was significantly decreased as compared to leaders cultured in LCM (Fig 3.9). In addition to effects on mitotic efficiency, FCM also impacted the percentage of cells undergoing cell death. The addition of FCM to leader cells significantly reduced cell death as compared to leaders cultured in LCM (Fig 3.9). Conversely, LCM had the opposite impact on follower cells and increased cell death (Fig 3.9). Interestingly, follower cells when co-cultured with leader cells also have an increase in cytokinetic instability as well as a decrease in the overall mitotic fraction, suggesting that leader cells may hinder follower cell growth. Taken together, these data support

a model whereby follower cells decrease mitotic defects and leader cell death while supporting leader cell colony formation, whereas leader cells increase these defects in follower cells thereby hindering follower cell growth.

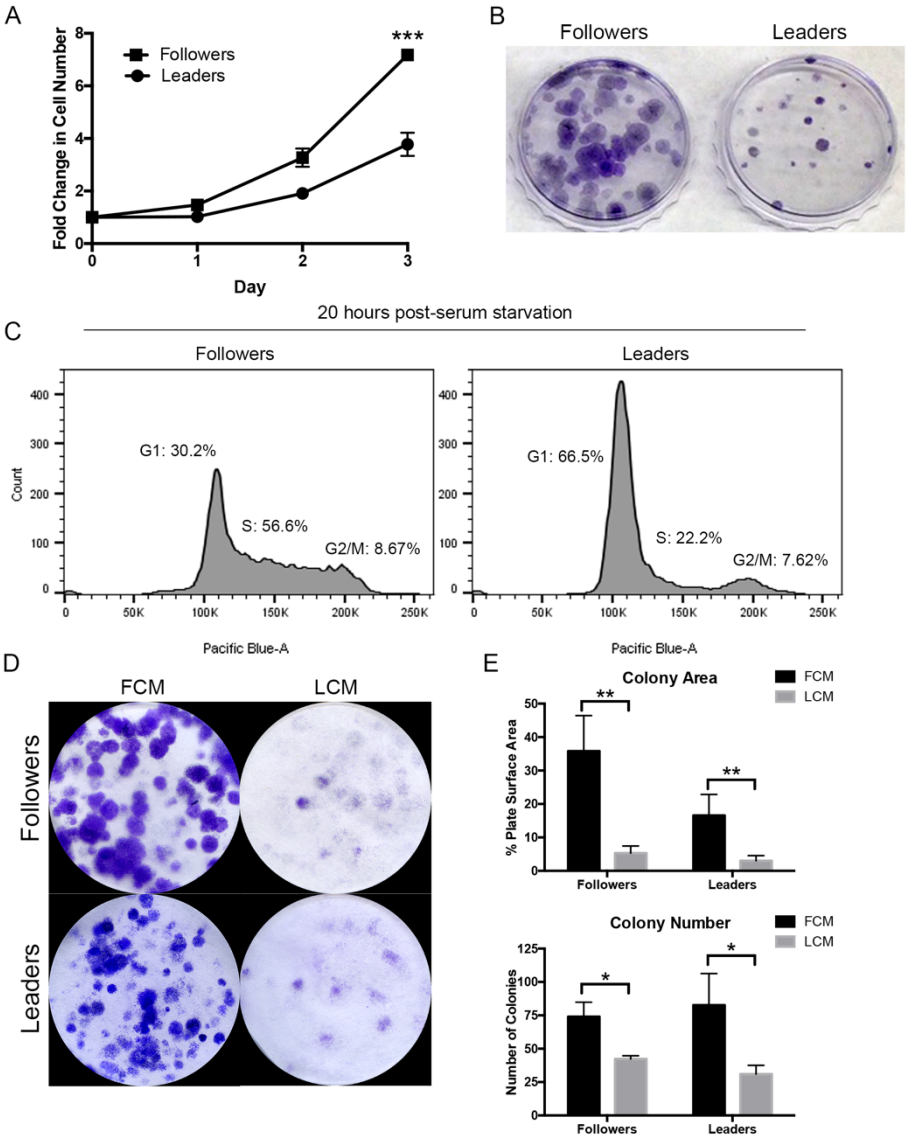

Figure 3.8. Follower cells represent a highly proliferative subpopulation that promotes the growth of leader cells via a secreted factor. A. Graph showing H1299 follower and leader cell growth over 3 days. ***p<0.001. B. Colony formation assay of H1299 parental, follower, and leader cells. C. Cell cycle analysis of H1299 follower and leader cells that were serum starved, and then released for 20 hours with normal growth media. D. Colony formation assay of H1299 follower and leader cells. Cells were plated in conditioned media from follower cells (FCM) or conditioned media from leader cells (LCM). Images taken 2 weeks after plating. E. The colony size and number of colonies from the experiment in D were measured using ImageJ software. *p<0.05, **p<0.01.

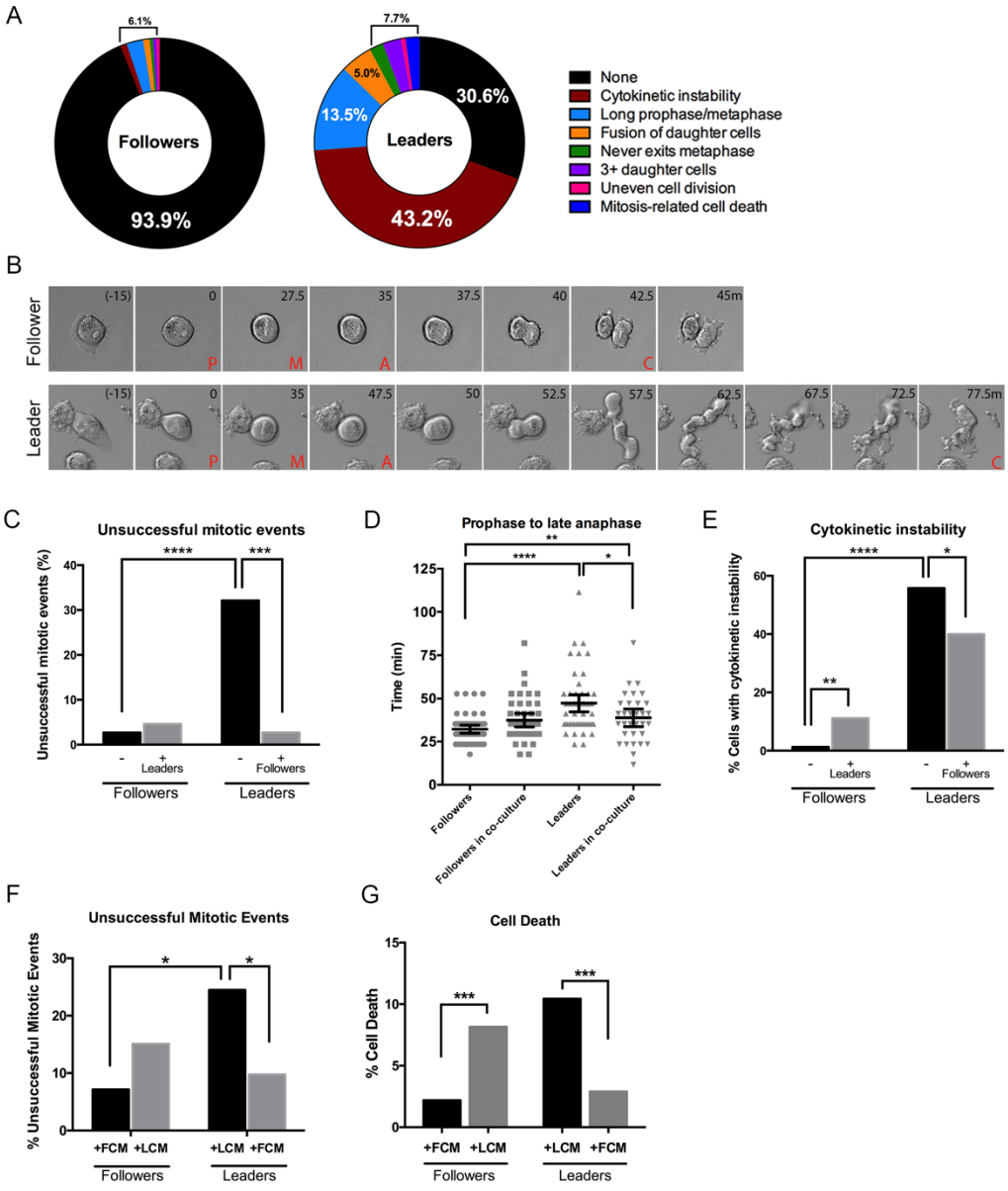

Figure 3.9. Mitotic defects observed in leader cells are rescued by the addition of follower cells. A-E. H1299 follower and RFP-leader cells were plated in 2-D alone or in a 50:50 mixed co-culture then imaged using live cell confocal microscopy. Mitotic events were analyzed in each condition. A. Graphs of each type of mitotic error noted during the live cell imaging as a percentage of all mitotic events. B. Still images of a mitotic event in a follower cell and a leader cell. Time in minutes. P = prophase, M = metaphase, A = anaphase, C = cytokinesis. C. Unsuccessful mitotic events were classified and graphed as a percentage of total mitotic events seen in followers, leaders, or each cell type in the co-culture condition. D. Dot plot of the amount of time each cell spent from prophase to the beginning of anaphase. n = 45 cells. Error bars represent the median and 95% confidence intervals. E. Cytokinetic instability events were graphed as a percentage of all dividing cells. F. Unsuccessful mitotic

events were counted and graphed from conditioned media from followers (FCM) or conditioned media from leaders (LCM) G. Leader and follower cells were cultured in LCM or FCM and cell death events were graphed as a percentage of total cells in the field of view. *p<0.05, **p<0.01 ***p<0.001, ****p<0.0001.

3.4 Discussion

SaGA combines microscopy, cell biology, and genomics to isolate and amplify phenotypically distinct subpopulations within a heterogeneous cell line. This technique lies at an emerging intersection of cell biology and genomics by combining single cell analysis with unbiased genomic datasets. We propose that SaGA can be used to isolate selected cells from phenotypically heterogeneous populations, including highly proliferative cells, drug-resistant cells, or other microscopy-amenable phenotypes. Furthermore, since a protein can be Dendra2-tagged (as opposed to palmitoylation-tagged Dendra2 used here), subcellular localization could be a stratifying phenotype to select cells with specific protein localizations.

We used SaGA here to probe the biological mechanisms that guide the phenotypic heterogeneity found in collectively cancer invasion (reviewed in (68, 154)). The leader cell phenotype is stable over many generations of culture and maintains an invasive and networked morphology when compared to follower or parental cells (Fig 3.4); therefore, leader cells are a stable phenotype and lack the phenotypic plasticity to revert back to a non-invasive phenotype. We do not observe follower cells taking on a leader cell position during collective invasion, which is in contrast to 2-D wounding studies where leader cells are replaced (155). In our studies, follower cells revert back to the parental phenotype (*i.e.,* gain collective invasion chains) suggesting that follower cells have greater phenotypic plasticity. We propose that leader cells are a specialized invasive cellular population, where the phenotype is hardwired into their genome or epigenome, and thus is less plastic. This is supported by previous studies that have found a specialized keratin-14 positive subpopulation capable of pioneering collective invasion in breast cancer (80). Additionally, recent work shows that a distinct epigenetic state of a breast

tumor cellular subpopulation promotes a transition to a more invasive cooperative cell invasion phenotype where canonical mesenchymal markers were insufficient to mark leader cells (81). Similarly, leader cells here also lack a complete EMT signature (Fig. 3.3), supporting the concept that invasive cancer cells can have partial EMT phenotypes that generate phenotypic plasticity (156).

Our transcriptomic data of SaGA-derived cell populations show several significant pathways enriched in leader or follower cells, including VEGF and adhesion signaling (Fig 3.6). VEGF signaling was an attractive candidate since leader-follower collective invasion morphologically resembles VEGF-driven vascular sprouting (157), which utilizes a leader-like tip cell and follower stalk cells. Our data support a model that resembles, but is not identical to, vascular sprouting signaling. Leader cells secrete VEGF, which is necessary for collective invasion pack formation, but not leader cell motility itself (Fig 3.6). Importantly, this differs from the endothelial cell vascular sprouting where a tip cell chemotaxes along a VEGF gradient through a hypoxic microenvironment. The leader-follower chain mimics the expression of canonical Notch-Dll4 endothelial cell expression patterns (150), further supporting the concept of a vascular signaling mimicry; however, Notch1 expression in follower cells does not appear to repress the leader phenotype, as is observed in canonical stalk cell maintenance (158, 159). Instead, in our case, inhibiting Notch1 reduces collective invasion (Fig S3.6), suggesting an atypical vascular signaling pathway. Other reports describe a tumor cell vascular mimicry where cancer cells upregulate endothelial-like gene expression programs to form vessel-like structures that act as a functioning blood supply (160). By imaging, vessel-like structures were not observed here but this possibility cannot be ruled out *in vivo*. Interestingly, while VEGF is required for pack formation, the actual motility of leader cells is instead dependent upon the

traction force created by fibronectin-FAK signaling (Fig 3.7). Fibronectin itself has been classically linked to cell invasion (161, 162), and leader cells secrete fibronectin at significantly higher levels than followers, and fibronectin is required for leader cell motility (Fig 3.7, S3.5). Taken together, this supports a model where formation of the collective invasion pack is VEGF-dependent and is a form of vascular mimicry, but motility itself utilizes fibronectin-FAK signaling.

One important question is why do cancer cells invade as a collective pack as opposed to single cells? One potential answer is that the multi-cellular pack provides a survival or invasive advantage to escaping cells. In circulating tumor cells, cells that invaded as groups had greater success and worse clinical outcomes (163, 164). Furthermore, studies show that tumor cell clusters rather than single cells seed polyclonal metastases in mouse models (163, 165-167), supporting the concept of collective invasion and/or metastasis. We observe that leader cells are competent to invade when follower cells are absent; however, when co-cultured with followers, single leader cell migration rarely occurs, suggesting leaders prefer migrating in packs. We show that followers provide a growth advantage to leader cells by increasing leader cell colony formation and correcting leader cell mitotic defects (Fig 3.8, 3.9). These data argue for a symbiotic relationship between leader and follower cells where follower cells increase leader cell mitotic success and leader cells provide followers with an escape mechanism (Fig 3.10). Interestingly, leader cell conditioned media caused follower cell death, increased their mitotic defects while decreasing their mitotic rate, and inhibited their colony formation (Fig 3.8, 3.9), suggesting leader cells can impact follower cell growth dynamics perhaps to maintain leader cell lineage within the cellular population.

Symbiosis usually involves a mutually beneficial relationship between different organisms (168); in this case however, the benefit occurs between two phenotypically distinct cellular populations within the collective invasion unit. Symbiosis in cancer has been proposed where cells cooperate to promote survival (169), especially in the context of heterogeneous sub-clonal populations (reviewed in (118)). Studies in a zebrafish melanoma model show that heterogeneous tumor cell populations cooperate to drive melanoma progression (170). Similarly, in a mouse xenograft model, inter-clonal cooperation occurs where tumor growth is driven by a minor cell subpopulation (171), and in breast cancer mouse models inter-clonal cooperation can be essential for Wnt-driven tumors (172). Taken together, these data suggest that therapeutic approaches aiming to disrupt the symbiotic ecosystem within the tumor cell community could potentially combat the dynamic evolutionary architecture of cancer.

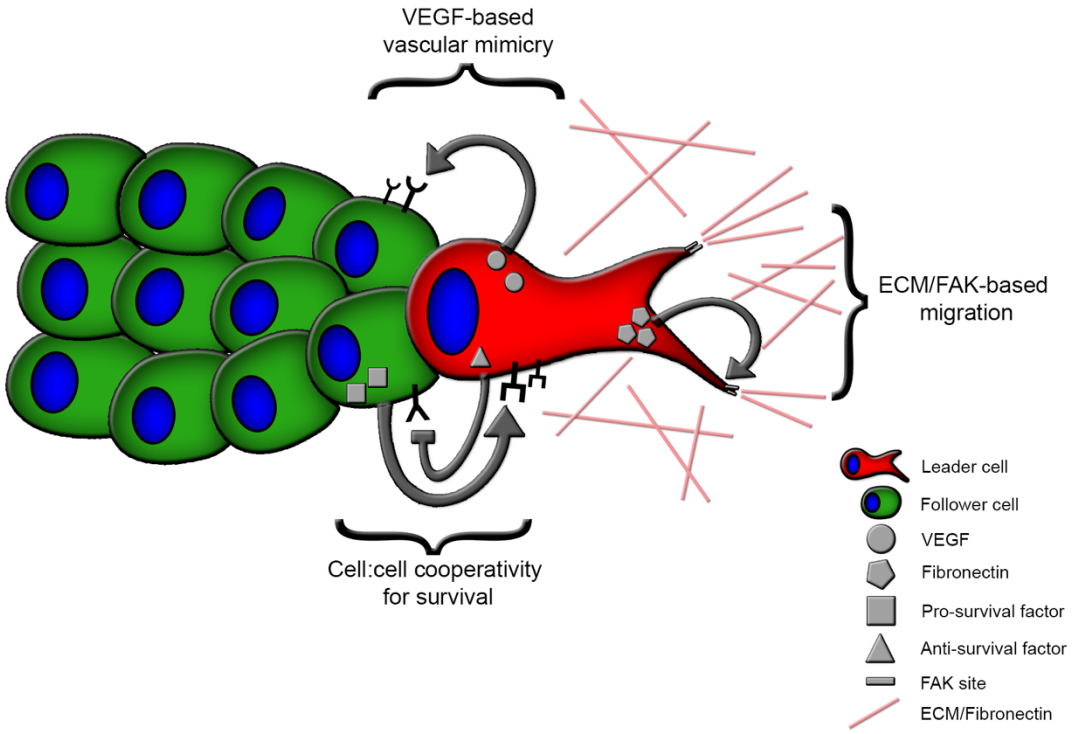

Figure 3.10. Model of cell symbiosis during collective cancer cell invasion. A collective invasion pack shows phenotypic heterogeneity consisting of at least two subpopulations, a highly invasive leader cell (red) at the front of the pack and follower cells (green) immediately attached to and following the leader cell. The leader cell has increased VEGF secretion, which promotes the motility and invasion of the follower cells. Other vasculogenesis signaling molecules are also utilized in the cell:cell communication between leader and follower cells, including VEGFR1, VEGFR2, VE-cadherin and Notch1; however, the activity of these proteins varies from the canonical functions. Thus, this atypical VEGF-based vascular signaling allows for successful chain formation during invasion. Additionally, leader cells secrete excess fibronectin which activates the canonical integrin/FAK pathway. This pathway allows for leader cells to create force to move the invasive pack forward into the microenvironment. Conversely, follower cells are a highly proliferative population that promote the growth of leader cells via a secreted factor, whereas leader cells secrete a factor that hinders follower cell growth. These data support a symbiotic relationship in the collective invasion pack in which the follower cells promote leader cell survival and leader cells promote follower cell escape.

Supplemental Information

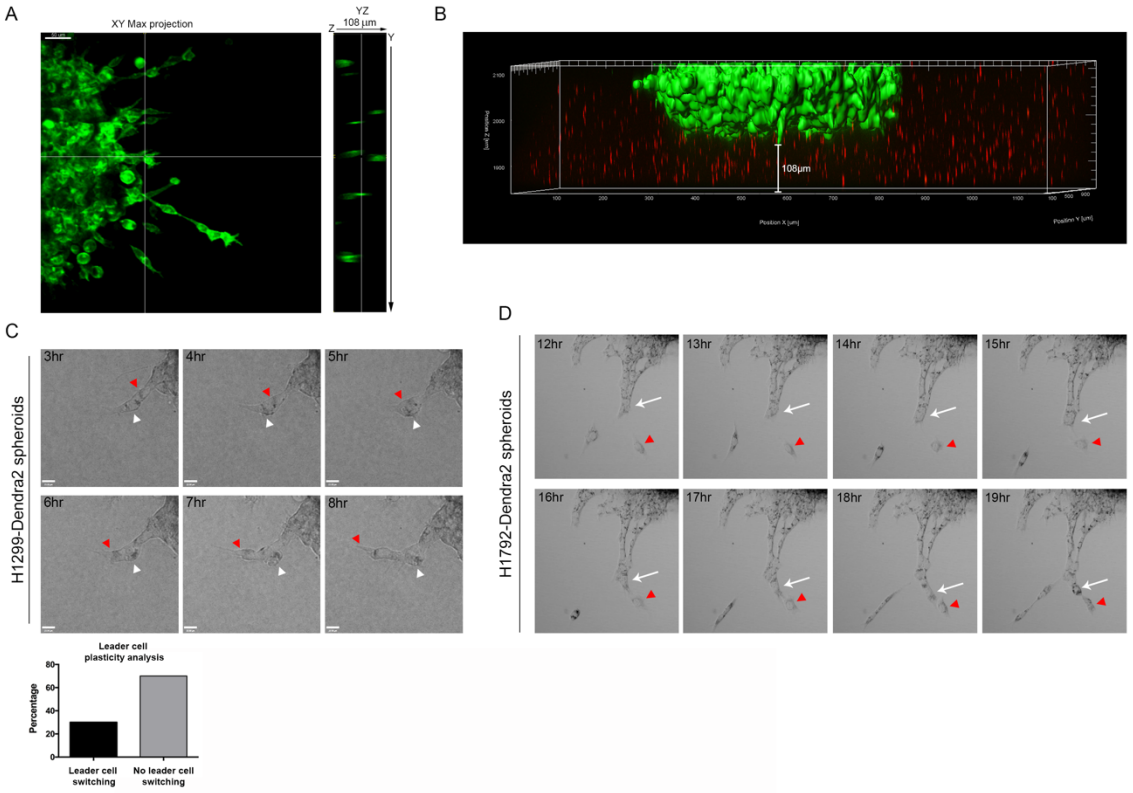

Figure S3.1. H1299 and H1792 spheroids show collective invasion in a 3-D Matrigel matrix. A. H1299-Dendra spheroid with multiple collective invasion chains. YZ projection on right shows that invasion occurs in multiple Z-planes. B. 3-D image of a H1299-Dendra2 spheroid embedded within Matrigel mixed with red fluorescent beads to show invasion occurs in a 3-D context. The bottom half of the spheroid and matrix are shown, and the distance between the spheroid and dish surface is marked. C. H1299-Dendra2 spheroid time lapse showing a single invasive chain with leader cell switching. White arrowhead = original leader cell. Red arrowhead = new leader cell. D. H1792-Dendra spheroids were embedded in Matrigel and imaged using live cell confocal microscopy. Images were acquired every ten minutes. Representative still images are shown, taken every 1 hour. Red arrowhead = leader cell. White arrow = follower cell within the invasive chain.

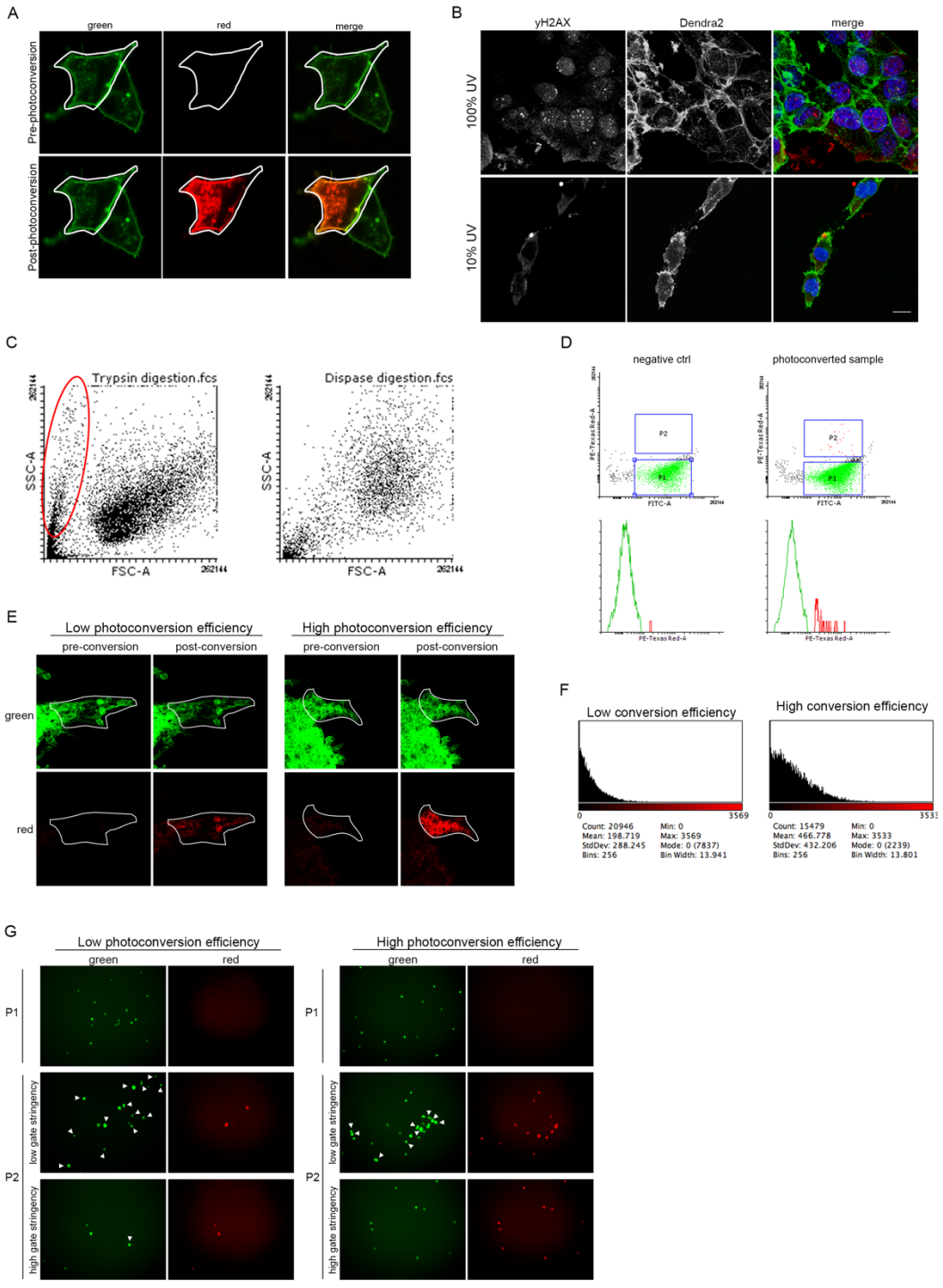

Figure S3.2. The development and optimization of spatiotemporal genomic and cellular analysis (SaGA) to isolate cells based upon invasive phenotype. A. H1299-Dendra2 cells in a 2-D monolayer were photoconverted using the 405nm laser. White outline around left cell shows the region of interest for photoconversion. B. Immunofluorescence imaging of yH2AX in H1299-Dendra2 spheroids. Invading cells were exposed to either 100%

or 10% 405nm laser power, which is comparable to the laser power utilized during photoconversion. C. H1299-Dendra2 spheroids embedded within a Matrigel matrix were degraded using either trypsin or dispase. Single cell suspensions were then analyzed via FACS and forward and side scatter plots are shown. Large cellular debris fraction is marked in the red circle. D. Photoconverted cells were analyzed via FACS using TexasRed to detect red positive cells. Gate was set using a negative control (H1299-Dendra2 cells that were not photoconverted). P1 = non-photoconverted cells, P2 = photoconverted cells. Histograms of Texas Red expression in each sample is shown below. E. Examples of low photoconversion efficiency (<300 a.f.u.) and high photoconversion efficiency (>300 a.f.u.) for optimization of sorting. F. Histogram of post-photoconversion images in E. Histograms represent only the red channel intensity. G. Using low (left) or high (right) photoconversion efficiency settings from E, photoconverted cells (P2) were collected on two different gate stringencies on the FACS machine. P1 denotes non-photoconverted cells as a negative control. Images were taken immediately after sorting to identify the sorting purity. Arrowheads = contaminating non-photoconverted cells.

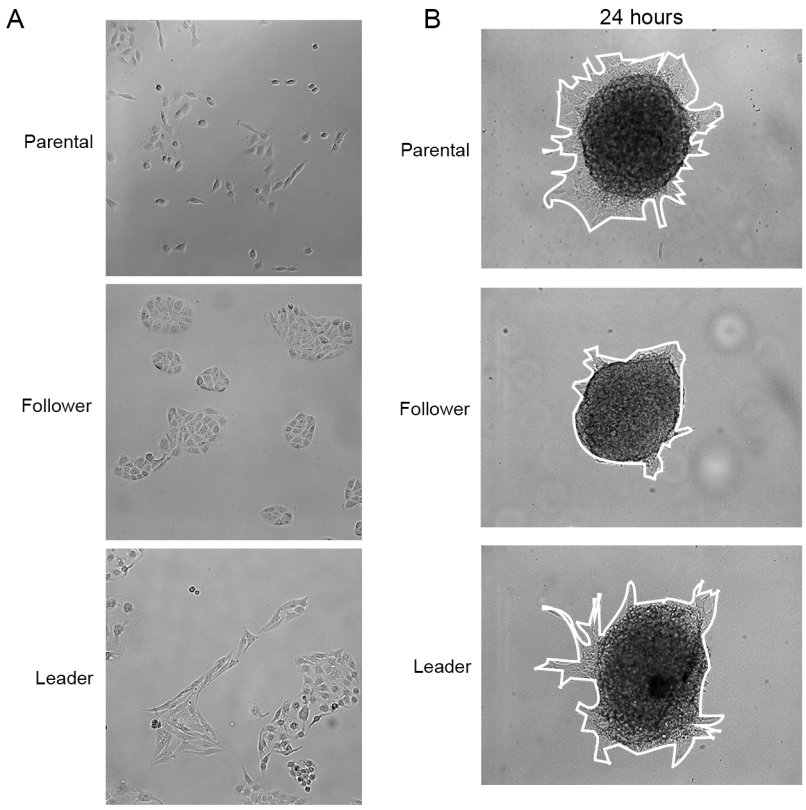

Figure S3.3. H1792-Dendra2 leader and follower cells show different morphology in 2-D culture and invasive patterning in 3-D spheroids. A. Brightfield images were taken of H1792-Dendra2 parental, follower, and leader cells in 2-D culture. B. H1792-Dendra2 parental, follower, and leader spheroids were embedded in Matrigel and imaged after 24 hours of invasion. White outline shows invasive area.

External ID	TH01		D21S11		D5S818		D13S317		D7S820		D16S539		CSF1PO		AMEL		vWA		TPOX	
2800M Control DNA	6	9.3	29	31.2	12	12	9	11	8	11	9	13	12	12	X	Y	16	19	11	11
H1299 Follower#1	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8
H1299 Follower#2	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8
H1299 Follower#3	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8
H1299 Leader#1	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8
H1299 Leader#2	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8
H1299-Dendra2 parent	6	9.3	32.2	32.2	11	11	12	12	10	10	12	13	12	12	X	X	16	18	8	8

Table S3.1. SNP analysis of H1299 leader and follower purified clones. H1299-Dendra2 parental cells collected and submitted for SNP analysis, and compared to the two leader and three follower purified clones. Identification analysis was performed using GeneMApper ID software.

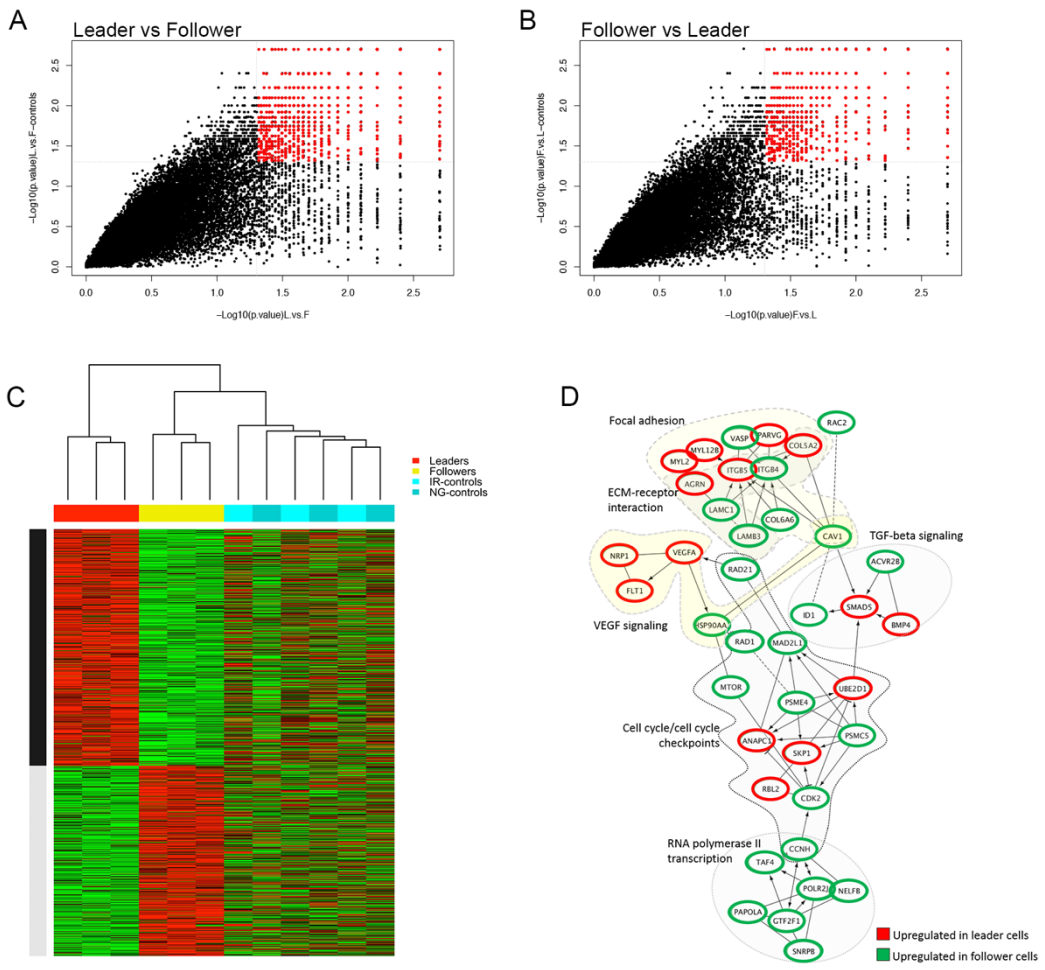

Figure S3.4. Gene expression analysis of purified leader and follower cells. A. Scatter plot showing permutation results of (-log10) p-values based on testing mean (pairwise among samples) expression differences greater in leaders (L) versus followers (F) (x-axis), and testing mean (pairwise among samples) expression differences between leaders and follows greater than differences in controls (y-axis). Each dot represents a gene. Genes whose mean expression was significantly (p.value $\leq$ 0.05) greater based on testing both hypotheses are highlighted in red. B. Analysis from A was repeated for expression differences greater in followers (F) versus leaders (L). C. Hierarchical clustering of 1422 significantly differentially expressed genes in leader (red; N=788) and follower (yellow; N=634) cells compared to controls (IR and NG; cyan shades). The hierarchical clustering was performed using Pearson correlation distance measure and average linkage method. The results represented are gene level expression log2 transformed after quantile normalization. D. Reactome FI network of significantly over expressed genes in the leader and follower cells relative to controls. Each node (or circle) represents a gene. Colors in the network indicate the over-expressed genes in the leader (red) and followers (green), respectively. The direct experimental and predicted interactions between the genes are shown by solid and dashed black lines, respectively. The known activating/catalyzing and inactivating (or inhibiting) interactions are shown by arrow-headed ("->") and bar-headed lines ("-|"), respectively. Only the GO processes with more

than three genes are shown. The network is created using Reactome FI network tool55 in Cytoscape54. Some of the pathways highlighted in the network are signaling by RNA polymerase II transcription, TGF-beta, focal adhesion, ECM-receptor interaction, cell cycle/cell cycle checkpoints, and VEGF signaling pathway.

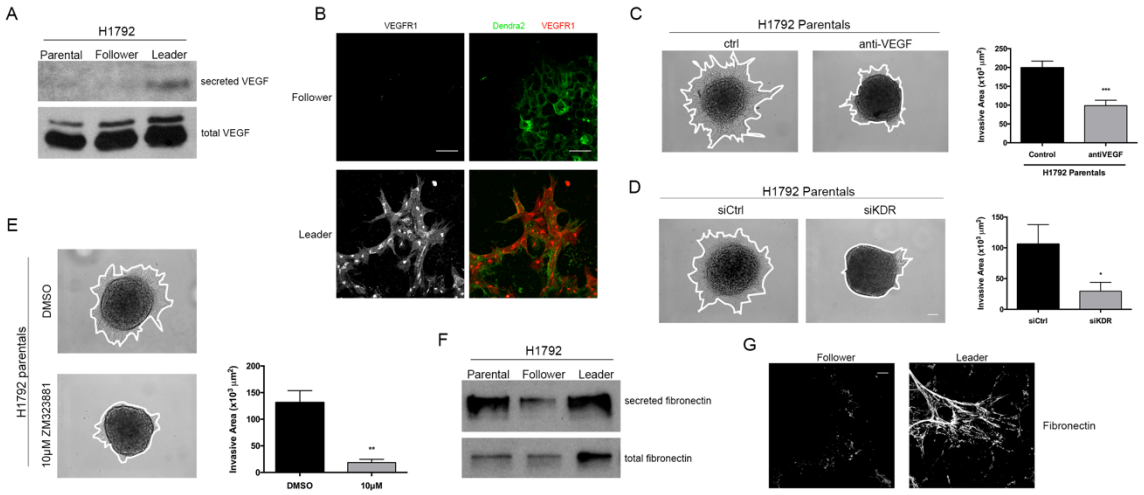

Figure S3.5. H1792 leader cells upregulate VEGF and fibronectin secretion as compared to follower cells. A. Western blot showing expression of VEGF in media samples (secreted) and whole cell lysates from H1792 parental, follower, and leader cells. B. Immunofluorescence imaging of VEGFR1 in H1792 follower and leader spheroids. Scale bar = 50μm. C. H1792 parental cells were treated with VEGF-neutralizing antibody (anti-VEGF) or vehicle control. Invasive area was measured and graphed below. n = 6 spheroids. **p<0.01. D. H1792 parental spheroids were analyzed for invasion after knockdown of VEGFR2 (*KDR*) via targeted siRNA. Scrambled siRNA was used as a negative control. Quantification of invasive area is shown to the right. n = 4-5 spheroids. *p<0.05. E. H1792 parental spheroids were treated with VEGFR2 kinase inhibitor, ZM323881 at 10μM, and invasion was compared to vehicle control (DMSO). A graph of invasive area is shown to the right. n = 5 spheroids. **p<0.01. F. Expression of fibronectin in both cell lysates and media (secreted) samples collected from H1792 parental, follower, and leader cells. G. Immunofluorescence imaging of extracellular fibronectin in H1792 follower and leader spheroids. Scale bar = 20μm.

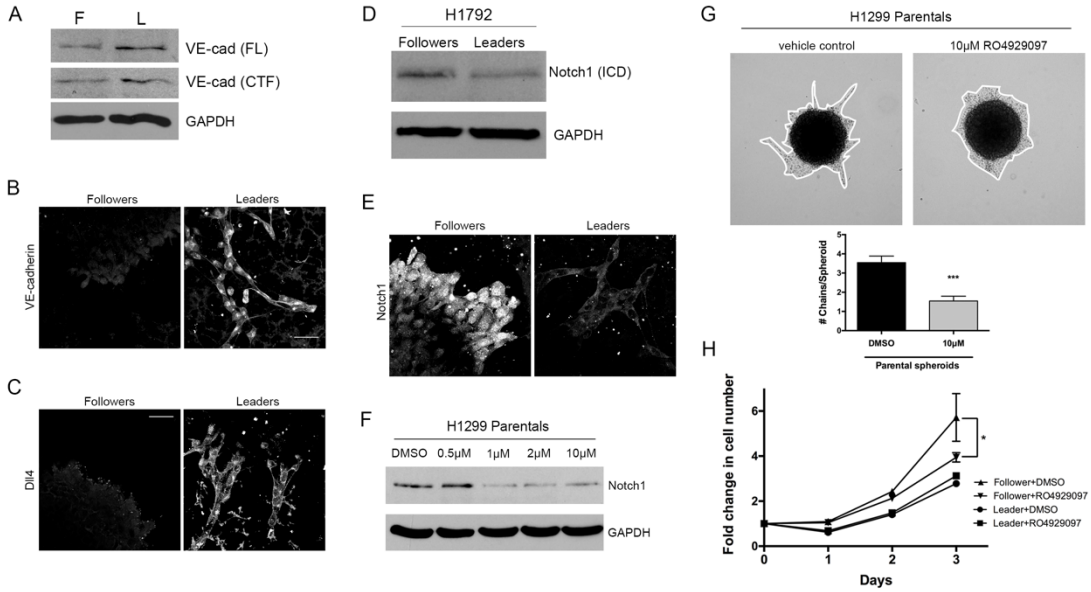

Figure S3.6. Leader cells mimic expression profiles seen during tip-stalk cell vasculogenesis. A. Western blot showing VE-cadherin expression, both full length and the C-terminal fragment, in follower and leader cell lysates. GAPDH was used as a loading control. B. Immunofluorescence imaging of VE-cadherin in follower and leader cell spheroids. Scale = 50μm. C. Immunofluorescence imaging of Dll4 in follower and leader spheroids. Scale = 50μm. D. Western blot showing Notch1 intracellular domain (ICD) expression in H1792 follower and leader cells. GAPDH was used as a loading control. E. Immunofluorescence imaging of Notch1 in follower and leader spheroids. Scale = 50μm. F. Western blot showing Notch1 intracellular domain expression in H1299 parental cells. Cells were treated with increasing doses of the γ-secretase inhibitor, RO4929097, or DMSO control. GAPDH was used as a loading control. G. H1299 parental spheroids were embedded in Matrigel in the presence of 10μM of RO4929097 or vehicle control. Images were taken 24 hours after embedding, and the number of invasive chains per spheroid was quantified. n = 11 spheroids. ***$p<0.001$. H. H1299 follower and leader cells were plated in 2-D with 10μM RO4929097 γ-secretase inhibitor and counted every day for 3 days. Graph shows fold change in total cell number over time. *$p<0.05$.

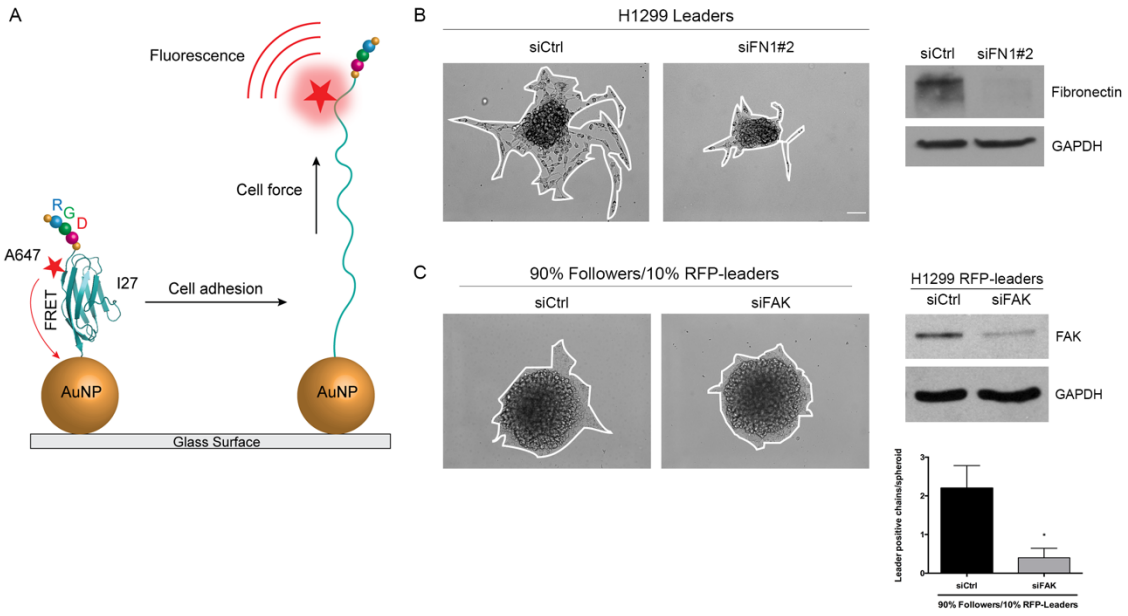

Figure S3.7. Adhesion studies in leader cells. A. Tension probe utilized to study adhesion forces (PDB ID: 1TIT). The experimental details used to prepare the probes and their response function is reported in (153). Briefly, the tension sensor is comprised of an engineered protein modified at the N-terminus with RGD integrin binding motif and with two cysteines at the C terminus for immobilization onto a 9 nm gold particle. Alexa647 organic dye was site-specifically conjugated to a non-canonical amino acid (p-azidophenylalanine) that was engineered near the N-terminus and allowed for copper free click chemistry. In the resting state, the titin-based probe is folded and the dye is quenched by the gold nanoparticle. When integrin receptors recognize the RGD sequence and unfolds the protein, this leads to separation of the dye from the gold nanoparticle. This results in dequenching of the dye, and a significant enhancement of the fluorescence signal at the sites of the integrin adhesions. Importantly, the tension fluorescence indicates a force magnitude that exceeds 36 pN.B. H1299 leader spheroids were formed from cells depleted of fibronectin (*FN1*) via a second targeted siRNA. Scrambled siRNA was used as a negative control. Western blot confirming knockdown is shown to the right. C. FAK siRNA or scrambled siRNA control was introduced into H1299 RFP-leader cells. Western confirming knockdown is shown to the right. Leader cells were plated as mixed spheroids at a 90:10 ratio with H1299 follower cells. RFP-leader positive chains were calculated and the graph is shown in the lower right. n = 5 spheroids. *$p<0.05$.

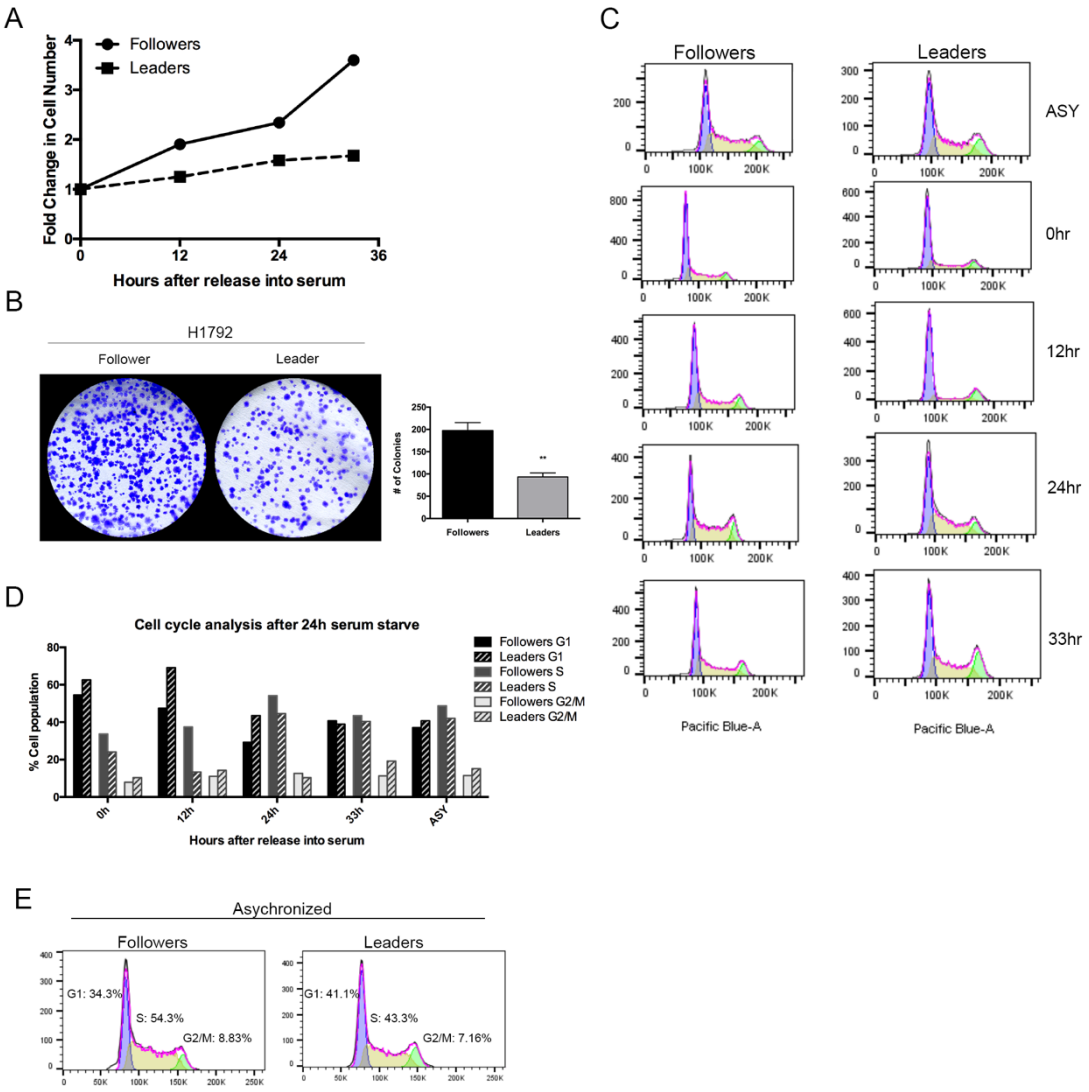

Figure S3.8. Cell cycle analysis of follower and leader cells. A. A sample of H1299 follower and leader cells were collected prior to fixation for the cell cycle analysis and counted using an automatic cell counting machine. Fold change in total cell number is shown over the time course. B. H1792 follower and leader cells were plated for a colony formation assay and stained with crystal violet. Quantification of the number of colonies is shown to the right. n = 3. **p<0.01. C-D) H1299 follower and leader cells were serum starved for 24 hours, then released by the addition of normal growth media for various lengths of time. Cells were collected for both cell growth (A) and cell cycle analysis. C. Cells were fixed at various times after release from serum starvation and stained with DAPI for cell cycle analysis. D. Quantification of each cell cycle phase from C in follower and leader cells. E. H1299 follower and leader cells were fixed and stained with DAPI for cell cycle analysis using flow cytometry. Cells were not first synchronized before fixation.

Chapter 4: Summary and Future Directions

The past several decades have shown tremendous progress in understanding the complex process of invasion and metastasis. As 3-D culturing, imaging, and genomics techniques have improved, more has been discovered about the cell-cell interactions and molecular mechanisms that drive cells to become invasive and ultimately metastatic. However, the work presented here has attempted to address some holes that exist in studying 3-D invasion in both a LKB1-dependent and –independent manner.

4.1 LKB1 regulation of 3-D lung cancer cell invasion

4.1.1 Summary of findings

The work on LKB1 has been extensive over the past several decades. After its identification as the major cause of the autosomal dominant disorder PJS, and then as a tumor suppressor frequently mutated and lost in sporadic lung adenocarcinomas, focus shifted to learning about this serine/threonine kinase and what role it plays in tumor progression. The seminal publication using a clinically relevant mouse model of lung cancer further propelled LKB1 into the limelight as a bona fide tumor suppressor, as its loss in combination with Kras activation significantly increased tumor progression and metastasis in comparison to inactivation of other tumor suppressors such as p53 (43). Now the question remained as to what molecular pathways are aberrantly active when LKB1 is lost. The kinase activity of LKB1 has been well studied and impacts a variety of downstream processes; thus, losing kinase activity would predictably have far-reaching effects. More focus has also been given to the C-terminal domain of LKB1 in recent years, as evidence suggests that the functional farnesylation

motif within this domain can impact LKB1 localization and signaling. How each of these domains independently impacts cell signaling and the downstream consequences remained unknown, especially within the physiologically relevant context of 3-D invasion. One recent publication focused on this question and found that LKB1 is vital for haptotaxis in a 3-D context (61); however, this study was completed in melanoma cells and they were not able to distinguish independent roles for kinase and farnesylation activity in this system. Thus, further studies on LKB1 function in a 3-D invasive context specifically in lung cancer were warranted, and were the focus of the work presented in Chapter 2.

Using 3-D multicellular tumor spheroids as our model system, we found that loss of LKB1 in lung cancer cells resulted in a mesenchymal to amoeboid transition (MAT). Additionally, LKB1-depleted amoeboid cells had significantly higher velocities during invasion than LKB1-competent amoeboid cells, suggesting that LKB1 loss provided an advantage to cells during invasion. Using various truncates and mutants of LKB1, we determined that kinase activity is not required for the regulation of this morphological transition. Instead, the C-terminal domain (CTD), specifically farnesylation of the CTD, is necessary to regulate 3-D invasive morphology as loss of CTD farnesylation resulted in MAT. Furthermore, CTD-dependent farnesylation controlled directionality of the cells during invasion. Taken together, our data suggests that LKB1 farnesylation within the CTD controls cellular polarity and directionality during 3-D migration; thus, loss of this functional motif results in a cell with aberrant polarity, a well established characteristic of invasive cancer cells.

The question remained as to what role the kinase domain plays during 3-D lung cancer invasion, so we sought to address this by analyzing the activity of FAK, as LKB1 has been

shown to play a role in regulating FAK activation. In our 3-D invasion model, we confirmed that LKB1 does repress the activity of FAK in a kinase-dependent manner. By blocking the activity of FAK via siRNA or pharmacological targeting, the velocity of LKB1 amoeboid cells was inhibited, suggesting that this hyperactive FAK resulting from loss of LKB1 kinase function promotes cell invasion via increased velocity. Additionally, our data supported that LKB1 kinase activity represses collagen remodeling, and loss of LKB1 kinase activity resulted in an increase in the matrix realignment and could ultimately be blocked by inhibiting FAK. These data point to the kinase activity of LKB1 to be essential for the repression of FAK and ultimately the repression of collagen remodeling during 3-D invasion.

Taken together, our findings support the concept that the functional domains of LKB1 have independent and varying roles in regulating 3-D lung cancer cell invasion, including the regulation of cell polarity, directionality, adhesion, and cell-matrix interactions. Patients harboring LKB1 truncating mutations, which is the majority of patients that present with LKB1 mutations in the clinic, would thus have disruptions in all of these pathways. Based on our findings, these patients would have tumors with hyper-invasive cell types that are particularly adept at navigating the tumor microenvironment, thus contributing to the tumor metastatic potential and patient survival.

4.1.2 Future directions

While the work outlined in this dissertation begins to address the *in vitro* functions of LKB1 in regulating 3-D invasion, numerous questions remain to be explored.

On the molecular level, there are several areas to add to our breadth of work on LKB1 functional domains and their roles in regulating invasion into a 3-D collagen gel. For example, one of the major known functions of LKB1 kinase activity is to phosphorylate and activate AMPK during times of energy stress. While it is well known that invasive cancer cells have differing metabolic requirements than noninvasive cancer cells (173-176), less is known about the metabolic requirements of specific types of invasive cells. It remains to be addressed if the significantly faster moving amoeboid cells are dependent on a particular metabolic pathway and whether this would be different than the requirements of a slower moving mesenchymal cell. The role that AMPK plays in regulating the balance between energy consumption and conservation could provide insight into the molecular mechanisms involved in creating these invasive cells. Additionally, AMPK is a known regulator of cell polarity, including functions independent of LKB1 (177). Thus, the role of AMPK in 3-D invasion with and without functioning LKB1 kinase activity should be explored. Other molecular questions also remain, such as which integrins are involved in promoting collagen remodeling during invasion and how LKB1 differentially regulates RhoA and cdc42 activity to control cellular polarization.

The molecular and cellular biology studies provided in Chapter 2 suggest that LKB1 loss impacts cell invasion *in vitro*, and previous works have shown that it is also vital in regulating metastasis *in vivo* (43). While these canonical studies thrust LKB1 into the limelight as a potential metastasis suppressor, little characterization was completed about the impact of LKB1 loss on the mode of invasion and metastasis, or what domains of LKB1 were critical. Using our panel of LKB1-depleted cells with rescue of either mutant or truncated forms, these questions could begin to be addressed.

As discussed in these works, lung adenocarcinoma patients that harbor LKB1 mutations show no clear mutational pattern; instead, mutations occur across the whole gene body and are mostly nonsense, and thus truncating, mutations. These patients would lose either just CTD function or CTD and kinase functions. However, patients also present with missense mutations that would predictably have differing repercussions on the cell biology of the tumor. Therefore, studying the impact of these clinically relevant mutations, and comparing nonsense to missense mutations, would be invaluable for understanding what downstream pathways are differentially active in each of these cases. Data generated from this type of analysis would potentially cluster LKB1 mutant patients into differing categories based on their transcriptome profiles, and ultimately, could impact the treatment strategies utilized for each subset of patients with different LKB1 mutations.

4.2 Cell symbiosis between invasive subpopulations in collective invasion

4.2.1 Summary of findings

Histological analyses of solid tumor samples collected from patients reveals that most solid tumor types invade into the microenvironment as collective units. Because it is clearly a clinically relevant characteristic of tumor biology, the field of collective invasion has become extensive, especially in terms of breast cancer invasion. One recent publication identified a basal epithelial program to be activated in the invasive leader cell population, with a focus on cytokeratin 14 expression (80). However, these leader cells were not identified as a distinct genetic subclone within the breast tumors, but instead, were a transient phenotype. Another study discovered a population of "trailblazer" cells that, similarly to leader cells, can pioneer invasion into a 3-D matrix (81). This study was also completed in breast tumors and again found them to be an epigenetically regulated population that occurs spontaneously. Additionally, the role of the following cells within the collective chain was not examined. Thus, we wanted to address several questions with our studies. Firstly, do highly invasive leading cells exist as a distinct, stable population and secondly, what cell-cell interactions are involved in building and maintaining the collective chain. Lastly, we wanted to address these questions in the context of lung cancer collective invasion as little work has begun to address this tumor type.

To model collective invasion *in vitro*, we again utilized our multicellular lung cancer spheroids embedded within 3-D Matrigel, and found that in this system, leading cells pioneer invasion into the microenvironment with following cells immediately behind. Quantitative analysis of live cell confocal videos revealed that these collective invasion chains move significantly faster,

farther, and with more directionality when a leader cell is attached to the chain. In contrast, when a leader cell becomes detached, the remaining cells in the collective chain are not able to promote invasion in 70% of the cases observed. Additionally, the leader cells that become detached attempt to reconnect with the follower cells, suggesting that these cells want to exist as a cohesive unit. Taken together, our preliminary data suggest that leader cells in our lung cancer 3-D invasion model are a specialized subpopulation that warrant further study to identify key pathways involved in their maintenance.

In order to specifically isolate, amplify, and analyze leader and follower invasive subpopulations, we developed and optimized the spatiotemporal genomic and cellular analysis (SaGA) image-guided genomics technique. Using this technique, we precisely selected and amplified purified leader and follower cells, creating the first ever cell lines with these invasive subpopulations. As our initial data suggested, we found that the leader cells are in fact a specialized subpopulation and even when cultured as a purified population are highly invasive when compared to followers, which show little to no invasion on their own. Leader cells are able to promote follower cell motility and invasion, thus rescuing chain invasion in followers. We performed a microarray on leaders and followers, and pathway analysis revealed that leaders upregulate several components with the VEGF pathway. We confirmed this via molecular studies and identified VEGF as a key factor in building the collective chain; leader cells secrete $VEGF_{165}$, to which follower cells respond. Blocking this pathway inhibits chain invasion both in the parental population and in spheroids containing a mixed population of leaders and followers; however, leader cells do not elicit an autocrine response to this VEGF gradient. Instead, their invasion is dependent upon the canonical cell adhesion pathway.

The question still remained as to why leader cells preferentially invade with follower cells when embedded as a mixed population. We observed that leader cells, when cultured as a purified population, often lagged in their growth. When performing growth assays, leader cells have markedly decreased growth when compared to followers, which are a highly proliferative population. Cell cycle analysis revealed leader cells to have a large G1 population, and live cell studies also identified that a large percentage of leader cells demonstrate various defects during mitosis. We hypothesized that followers may provide a growth or survival advantage to leaders. Indeed, when co-cultured with follower cells or exposed to follower cell conditioned media, leader cell mitotic defects and leader cell growth are significantly rescued. These data suggest that follower cells provide leader cells with a growth or survival advantage via a secreted factor.

Taken as a whole, our data begins to dissect leader and follower invasive subpopulations and the underlying mechanisms involved in the success of the collective unit. These populations exist in a symbiotic relationship, wherein leader cells provide follower cells with a way to escape the primary spheroid and follower cells in turn represent a proliferative population that provides leaders with a growth advantage.

4.2.2 Future directions

The development and implementation of SaGA provided a way for us to link a particular cellular phenotype to a genomic or transcriptomic profile. Our goal was to learn about the collective chain in lung cancer invasion; however, the potential impact of SaGA is far-reaching, as it can be utilized to study any observable cellular phenotype. When discussing cancer biology in particular, relatively rare cell populations are thought to be vital in driving tumor

progression, drug resistance, and metastasis. Traditional sequencing technologies fall short in either isolating these rare populations, or cannot link a phenotype or underlying cell biology to a particular genomic profile. SaGA can thus be utilized as a tool to probe tumor heterogeneity with the ability to isolate rare, yet dynamic cell types and assign a genomic profile to each variable phenotype observed within a heterogeneous tumor. While well outside the scope of the work presented here, the development of SaGA provides vast future directions in which to study tumor biology.

In terms of the work presented here, there are several questions that remain unanswered, as well as many new directions in which to take the research. For example, several key molecular studies should be addressed in order to complete the story presented in Chapter 3. It is still unknown how follower cells promote the growth and/or survival of leader cells, and this information is necessary to fully understand how these subpopulations co-depend on each other and may provide a novel therapeutic opportunity. The data presented suggests that follower cell co-culture or even just follower cell conditioned media is sufficient to correct mitotic defects observed in leader cells. Additionally, there is a significant reduction in the number of leader cells that undergo apoptosis. Thus, there is an increase in leader cell growth over time when cultured with follower conditioned media as shown by colony formation assay. However, the factor that followers are secreting to promote these effects is unknown, as well as how a secreted molecule could rectify mitotic defects. There is some precedence in the literature that secreted factors can impact mitosis or mitotic synchrony (178-180). Our microarray data may provide some insights, as there are several soluble factors that were upregulated by followers that could provide leader cells with a growth advantage. Additionally, the molecular mechanisms involved in the building of the invasive chain are known only

superficially. Our data points to the VEGF signaling pathway as being necessary for the formation of the invasive chain; however, the downstream players are still unknown. We have some data that suggests follower cells respond to this VEGF gradient via VEGFR2, as VEGFR2 knockdown and kinase inhibitors reduce collective invasion (Fig 3.6, S3.5). However, this has yet to be solidified, as well as what molecules are activated downstream of the receptor activation in order for follower cell motility and invasion to be promoted. There are many effectors of the VEGF signaling pathway, so this mechanism needs to be further elucidated.

As mentioned early, there are many new questions that can now be addressed with our leader and follower purified lines. One obvious direction is whether these purified subpopulations actually have differing metastatic potential. While the *in vitro* studies clearly support that leaders and followers have different invasive capabilities, *in vivo* studies to illustrate that these characteristics are also true in a mouse model would greatly strengthen the work presented here. Pilot studies have been performed using subcutaneous flank xenografts with the H1299 parental, follower, and leader cells. Six of 10 parental cell injections and 9 of 10 follower cell injections formed primary tumors (not shown). Interestingly, the leader cells formed no primary tumors, supporting the *in vitro* data that these cells do not grow well when cultured as a purified population. As our model predicts, both subpopulations may be necessary for successful metastasis to occur, so performing *in vivo* mixing experiments with limiting dilutions of RFP-leader cells would be important to address whether metastases from these animals have both subpopulations and if so, at what ratios. Taking this a step further and identifying a leader-specific biomarker would allow us to extend our studies into patients. We could use this biomarker to identify leader cells in histological samples and correlate the presence of leader

cells with patient survival. Another future direction is to perform a drug screen to identify differential drug targets for leaders and followers. Because leader cells have both mitotic defects as well as a large G1 population, this cell population may be particularly resistant to common chemotherapeutic agents. Therefore, finding a leader cell-specific drug may reveal a novel sensitivity in these cells that would make them more susceptible to treatment.

Lastly, our main goal with the development of SaGA was to assign genomic profiles to particular rare cell phenotypes. While we were able to perform a microarray and obtain information about gene expression differences, our bigger question is whether these two phenotypes arise from differing mutations; in other words, are leader and follower cells actually two independent subclones with identifying mutations? To address this, we performed RNA-sequencing on leader and follower cultures and compared back to the parental cell line. Using this approach, we identified numerous differentially expressed mutations in these subpopulations and have since validated several hits via traditional sequencing methods (not shown). Thus, future studies will focus on identifying which of these mutations may be biologically relevant and therefore be considered "drivers" of the leader cell phenotype. If one or more of these mutations are found to be imperative for the leader cell phenotype, a sequencing panel to perform deep sequencing on patient samples could be created to ultimately search for these mutations in patients and correlate with survival data. Because leader cells represent a rare population, these mutations would likely be missed with normal sequencing methods. By narrowing the coverage to only a few genes and performing deep sequencing, these rare mutations could then be exposed in patients. Overall, these data could help precision medicine efforts to learn more about the subpopulations that exist within a heterogeneous tumor.

References

1. Cancer Facts & Figures, 2015. American Cancer Society. http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2015/. Accessed on 09-06-2016.
2. SEER Stat Fact Sheets: Lung and Bronchus. National Cancer Institute. http://seer.cancer.gov/statfacts/html/lungb.html. Accessed on 09-06-2016.
3. Shim HS, L.d.H., Park EJ, Kim SH Histopathologic characterisitics of lung adenocarcinomas with epidermal growth factor receptor mutations in the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society lung adenocarcinoma classification. *Arch Pathol Lab Med* **135**, 1329-1334 (2011).
4. Pao, W. *et al.* Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. *PLoS Medicine* **2**, e73 (2005).
5. Stella, G.M. *et al.* EGFR and KRAS mutational profiling in fresh non-small cell lung cancer (NSCLC) cells. *Journal of Cancer Research and Clinical Oncology* **139**, 1327-1335 (2013).
6. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069-1075 (2008).
7. Sanchez-Cespedes, M. *et al.* Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res* **62**, 3659-3662 (2002).
8. de Bruin, E.C. *et al.* Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* **346**, 251-256 (2014).
9. de Bruin, E.C., McGranahan, N. & Swanton, C. Analysis of intratumor heterogeneity unravels lung cancer evolution. *Molecular & Cellular Oncology* **2**, e985549 (2015).
10. Zhang, J. *et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256-259 (2014).
11. Mehlen, P. & Puisieux, A. Metastasis: a question of life or death. *Nature reviews. Cancer* **6**, 449-458 (2006).

12. Spano, D., Heck, C., De Antonellis, P., Christofori, G. & Zollo, M. Molecular networks that regulate cancer metastasis. *Seminars in cancer biology* **22**, 234-249 (2012).
13. Chambers, A.F., Groom, A.C. & MacDonald, I.C. Metastasis: Dissemination and growth of cancer cells in metastatic sites. *Nature reviews. Cancer* **2**, 563-572 (2002).
14. Valastyan, S. & Weinberg, Robert A. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147**, 275-292.
15. Ma, L. & Weinberg, R.A. Micromanagers of malignancy: role of microRNAs in regulating metastasis. *Trends in Genetics* **24**, 448-456.
16. Mina, L.A. & Sledge, G.W. Rethinking the metastatic cascade as a therapeutic target. *Nat Rev Clin Oncol* **8**, 325-332 (2011).
17. Kalluri, R. & Weinberg, R.A. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation* **119**, 1420-1428 (2009).
18. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nature reviews. Molecular cell biology* **15**, 178-196 (2014).
19. Futterman, M.A., García, A.J. & Zamir, E.A. Evidence for Partial Epithelial-to-Mesenchymal Transition (pEMT) and Recruitment of Motile Blastoderm Edge Cells During Avian Epiboly. *Developmental dynamics : an official publication of the American Association of Anatomists* **240**, 1502-1511 (2011).
20. Leroy, P. & Mostov, K.E. Slug Is Required for Cell Survival during Partial Epithelial-Mesenchymal Transition of HGF-induced Tubulogenesis. *Molecular biology of the cell* **18**, 1943-1952 (2007).
21. Kessenbrock, K., Plaks, V. & Werb, Z. Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. *Cell* **141**, 52-67 (2010).
22. Nabeshima, K., Inoue, T., Shimao, Y. & Sameshima, T. Matrix metalloproteinases in tumor invasion: role for cell migration. *Pathol Int* **52**, 255-264 (2002).
23. Stamenkovic, I. Matrix metalloproteinases in tumor invasion and metastasis. *Seminars in cancer biology* **10**, 415-433 (2000).
24. Carmeliet, P. & Jain, R.K. Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. *Nature reviews. Drug discovery* **10**, 417-427 (2011).
25. Wyckoff, J.B. *et al.* Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res* **67**, 2649-2656 (2007).

26. Joyce, J.A. & Pollard, J.W. Microenvironmental regulation of metastasis. *Nature reviews. Cancer* **9**, 239-252 (2009).
27. Douma, S. *et al.* Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB. *Nature* **430**, 1034-1039 (2004).
28. Schafer, Z.T. *et al.* Antioxidant and oncogene rescue of metabolic defects caused by loss of matrix attachment. *Nature* **461**, 109-113 (2009).
29. Gupta, G.P. *et al.* Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* **446**, 765-770 (2007).
30. Padua, D. *et al.* TGFβ primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
31. Weis, S., Cui, J., Barnes, L. & Cheresh, D. Endothelial barrier disruption by VEGF-mediated Src activity potentiates tumor cell extravasation and metastasis. *The Journal of Cell Biology* **167**, 223-229 (2004).
32. Klein, C.A. Parallel progression of primary tumours and metastases. *Nature reviews. Cancer* **9**, 302-312 (2009).
33. Lu, P., Weaver, V.M. & Werb, Z. The extracellular matrix: A dynamic niche in cancer progression. *The Journal of Cell Biology* **196**, 395-406 (2012).
34. Barkan, D., Green, J.E. & Chambers, A.F. Extracellular Matrix: A Gatekeeper in the Transition from Dormancy to Metastatic Growth. *European journal of cancer (Oxford, England : 1990)* **46**, 1181-1188 (2010).
35. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
36. Kunz-Schughart, L.A. & Knuechel, R. Tumor-associated fibroblasts (part I): Active stromal participants in tumor development and progression? *Histology and histopathology* **17**, 599-621 (2002).
37. Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nature reviews. Molecular cell biology* **10**, 445-457 (2009).
38. Friedl, P. & Wolf, K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature reviews. Cancer* **3**, 362-374 (2003).
39. Stupack, D.G. *et al.* Potentiation of neuroblastoma metastasis by loss of caspase-8. *Nature* **439**, 95-99 (2006).

40. Kapitanović, S. *et al.* nm23-H1 expression and loss of heterozygosity in colon adenocarcinoma. *Journal of Clinical Pathology* **57**, 1312-1318 (2004).
41. Dong, J.-T. *et al.* <div xmlns="http://www.w3.org/1999/xhtml of the KAI1 Metastasis Suppressor Gene during the Progression of Human Prostatic Cancer Infrequently Involves Gene Mutation or Allelic Loss . *Cancer Research* **56**, 4387-4390 (1996).
42. Yang, X., Welch, D.R., Phillips, K.K., Weissman, B.E. & Wei, L.L. KAI1, a putative marker for metastatic potential in human breast cancer. *Cancer letters* **119**, 149-155.
43. Ji, H. *et al.* LKB1 modulates lung cancer differentiation and metastasis. *Nature* **448**, 807-810 (2007).
44. Yoon, K.A. *et al.* Germline mutations of the STK11 gene in Korean Peutz-Jeghers syndrome patients. *British journal of cancer* **82**, 1403-1406 (2000).
45. Olschwang, S. *et al.* Peutz-Jeghers disease: most, but not all, families are compatible with linkage to 19p13.3. *Journal of Medical Genetics* **35**, 42-44 (1998).
46. Schumacher, V. *et al.* STK11 genotyping and cancer risk in Peutz-Jeghers syndrome. *Journal of Medical Genetics* **42**, 428-435 (2005).
47. Guldberg, P. *et al.* Somatic mutation of the Peutz-Jeghers syndrome gene, LKB1/STK11, in malignant melanoma. *Oncogene* **18**, 1777-1780 (1999).
48. Matsumoto, S. *et al.* Prevalence and specificity of LKB1 genetic alterations in lung cancers. *Oncogene* **26**, 5911-5918 (2007).
49. Wingo, S.N. *et al.* Somatic LKB1 mutations promote cervical cancer progression. *PloS one* **4**, e5137 (2009).
50. The Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-550 (2014).
51. Hezel, A.F. & Bardeesy, N. LKB1; linking cell structure and tumor suppression. *Oncogene* **27**, 6908-6919 (2008).
52. Shackelford, D.B. & Shaw, R.J. The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nature reviews. Cancer* **9**, 563-575 (2009).
53. Marcus, A.I. & Zhou, W. LKB1 regulated pathways in lung cancer invasion and metastasis. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* **5**, 1883-1886 (2010).

54. Hawley, S.A. *et al.* Complexes between the LKB1 tumor suppressor, STRAD alpha/beta and MO25 alpha/beta are upstream kinases in the AMP-activated protein kinase cascade. *Journal of biology* **2**, 28 (2003).
55. Shaw, R.J. *et al.* The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc Natl Acad Sci U S A* **101**, 3329-3335 (2004).
56. Hardie, D.G. & Alessi, D.R. LKB1 and AMPK and the cancer-metabolism link - ten years after. *BMC biology* **11**, 36 (2013).
57. Jaleel, M. *et al.* Identification of the sucrose non-fermenting related kinase SNRK, as a novel LKB1 substrate. *FEBS letters* **579**, 1417-1423 (2005).
58. Lizcano, J.M. *et al.* LKB1 is a master kinase that activates 13 kinases of the AMPK subfamily, including MARK/PAR-1. *The EMBO journal* **23**, 833-843 (2004).
59. Baas, A.F. *et al.* Complete polarization of single intestinal epithelial cells upon activation of LKB1 by STRAD. *Cell* **116**, 457-466 (2004).
60. Zhang, S. *et al.* The tumor suppressor LKB1 regulates lung cancer cell polarity by mediating cdc42 recruitment and activity. *Cancer Res* **68**, 740-748 (2008).
61. Chan, K.T. *et al.* LKB1 loss in melanoma disrupts directional migration toward extracellular matrix cues. *The Journal of Cell Biology* **207**, 299-315 (2014).
62. Carretero, J. *et al.* Integrative Genomic and Proteomic Analyses Identify Targets for Lkb1-Deficient Metastatic Lung Tumors. *Cancer Cell* **17**, 547-559 (2010).
63. Goodwin, Jonathan M. *et al.* An AMPK-Independent Signaling Pathway Downstream of the LKB1 Tumor Suppressor Controls Snail1 and Metastatic Potential. *Molecular Cell* **55**, 436-450 (2014).
64. Kline, E.R., Shupe, J., Gilbert-Ross, M., Zhou, W. & Marcus, A.I. LKB1 Represses Focal Adhesion Kinase (FAK) Signaling via a FAK-LKB1 Complex to Regulate FAK Site Maturation and Directional Persistence. *The Journal of biological chemistry* **288**, 17663-17674 (2013).
65. Gao, Y. *et al.* LKB1 inhibits lung cancer progression through lysyl oxidase and extracellular matrix remodeling. *Proc Natl Acad Sci U S A* **107**, 18892-18897 (2010).
66. Roy, B.C. *et al.* Involvement of LKB1 in epithelial-mesenchymal transition (EMT) of human lung cancer cells. *Lung cancer (Amsterdam, Netherlands)* **70**, 136-145 (2010).

67. Yao, Y.H. *et al.* Attenuated LKB1-SIK1 signaling promotes epithelial-mesenchymal transition and radioresistance of non-small cell lung cancer cells. *Chinese journal of cancer* **35**, 50 (2016).
68. Friedl, P., Locker, J., Sahai, E. & Segall, J.E. Classifying collective cancer cell invasion. *Nat Cell Biol* **14**, 777-783 (2012).
69. Friedl, P. Prespecification and plasticity: shifting mechanisms of cell migration. *Current opinion in cell biology* **16**, 14-23 (2004).
70. Friedl, P. & Alexander, S. Cancer Invasion and the Microenvironment: Plasticity and Reciprocity. *Cell* **147**, 992-1009.
71. Friedl, P. & Wolf, K. Plasticity of cell migration: a multiscale tuning model. *The Journal of Cell Biology* **188**, 11-19 (2010).
72. Paňková, K., Rösel, D., Novotný, M. & Brábek, J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cellular and Molecular Life Sciences* **67**, 63-71 (2010).
73. Wolf, K. *et al.* Compensation mechanism in tumor cell migration: mesenchymal–amoeboid transition after blocking of pericellular proteolysis. *The Journal of Cell Biology* **160**, 267-277 (2003).
74. Wolf, K. *et al.* Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. *Nat Cell Biol* **9**, 893-904 (2007).
75. Sabeh, F., Shimizu-Hirota, R. & Weiss, S.J. Protease-dependent versus -independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. *J Cell Biol* **185**, 11-19 (2009).
76. Sahai, E. & Marshall, C.J. Differing modes of tumour cell invasion have distinct requirements for Rho/ROCK signalling and extracellular proteolysis. *Nat Cell Biol* **5**, 711-719 (2003).
77. Wolf, K., Müller, R., Borgmann, S., Bröcker, E.-B. & Friedl, P. *Amoeboid shape change and contact guidance: T-lymphocyte crawling through fibrillar collagen is independent of matrix remodeling by MMPs and other proteases*, Vol. 102. (2003).
78. Parri, M. & Chiarugi, P. Rac and Rho GTPases in cancer cell motility control. *Cell communication and signaling : CCS* **8**, 23 (2010).
79. Khalil, A.A. & Friedl, P. Determinants of leader cells in collective cell migration. *Integrative biology : quantitative biosciences from nano to macro* **2**, 568-574 (2010).

80. Cheung, Kevin J., Gabrielson, E., Werb, Z. & Ewald, Andrew J. Collective Invasion in Breast Cancer Requires a Conserved Basal Epithelial Program. *Cell* **155**, 1639-1651 (2014).
81. Westcott, J.M. *et al.* An epigenetically distinct breast cancer cell subpopulation promotes collective invasion. *J Clin Invest* **125**, 1927-1943 (2015).
82. Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T. & Drewes, G. Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. *The Journal of biological chemistry* **279**, 12804-12811 (2004).
83. Spicer, J. *et al.* Regulation of the Wnt signalling component PAR1A by the Peutz-Jeghers syndrome kinase LKB1. *Oncogene* **22**, 4752-4756 (2003).
84. Vallenius, T. *et al.* An association between NUAK2 and MRIP reveals a novel mechanism for regulation of actin stress fibers. *Journal of Cell Science* **124**, 384-393 (2011).
85. Zagórska, A. *et al. New Roles for the LKB1-NUAK Pathway in Controlling Myosin Phosphatase Complexes and Cell Adhesion*, Vol. 3. (2010).
86. Cheng, H. *et al.* SIK1 Couples LKB1 to p53-Dependent Anoikis and Suppresses Metastasis. *Science Signaling* **2**, ra35-ra35 (2009).
87. Barnes, A.P. *et al.* LKB1 and SAD Kinases Define a Pathway Required for the Polarization of Cortical Neurons. *Cell* **129**, 549-563 (2007).
88. Kishi, M., Pan, Y.A., Crump, J.G. & Sanes, J.R. Mammalian SAD Kinases Are Required for Neuronal Polarization. *Science* **307**, 929-932 (2005).
89. Nakano, A. & Takashima, S. LKB1 and AMP-activated protein kinase: regulators of cell polarity. *Genes to cells : devoted to molecular & cellular mechanisms* **17**, 737-747 (2012).
90. Watts, J.L., Morton, D.G., Bestman, J. & Kemphues, K.J. The C. elegans par-4 gene encodes a putative serine-threonine kinase required for establishing embryonic asymmetry. *Development (Cambridge, England)* **127**, 1467-1475 (2000).
91. Amin, N. *et al.* LKB1 regulates polarity remodeling and adherens junction formation in the Drosophila eye. *Proc Natl Acad Sci U S A* **106**, 8941-8946 (2009).
92. Martin, S.G. & St Johnston, D. A role for Drosophila LKB1 in anterior-posterior axis formation and epithelial polarity. *Nature* **421**, 379-384 (2003).

93. Partanen, J.I. *et al.* Tumor suppressor function of Liver kinase B1 (Lkb1) is linked to regulation of epithelial integrity. *Proc Natl Acad Sci U S A* **109**, E388-397 (2012).
94. Asada, N., Sanada, K. & Fukada, Y. LKB1 regulates neuronal migration and neuronal differentiation in the developing neocortex through centrosomal positioning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **27**, 11769-11775 (2007).
95. Etienne-Manneville, S. Cdc42--the centre of polarity. *J Cell Sci* **117**, 1291-1300 (2004).
96. Xu, X., Jin, D., Durgan, J. & Hall, A. LKB1 controls human bronchial epithelial morphogenesis through p114RhoGEF dependent RhoA activation. *Molecular and cellular biology* (2013).
97. Xu, X., Omelchenko, T. & Hall, A. LKB1 tumor suppressor protein regulates actin filament assembly through Rho and its exchange factor Dbl independently of kinase activity. *BMC cell biology* **11**, 77 (2010).
98. Havel, L.S., Kline, E.R., Salgueiro, A.M. & Marcus, A.I. Vimentin regulates lung cancer cell adhesion through a VAV2-Rac1 pathway to control focal adhesion kinase activity. *Oncogene* (2014).
99. Ladhani, O., Sanchez-Martinez, C., Orgaz, J.L., Jimenez, B. & Volpert, O.V. Pigment epithelium-derived factor blocks tumor extravasation by suppressing amoeboid morphology and mesenchymal proteolysis. *Neoplasia (New York, N.Y.)* **13**, 633-642 (2011).
100. Nogawa, H. & Mizuno, T. Mesenchymal control over elongating and branching morphogenesis in salivary gland development. *J Embryol Exp Morphol* **66**, 209-221 (1981).
101. Bredfeldt, J.S. *et al.* Automated quantification of aligned collagen for human breast carcinoma prognosis. *J Pathol Inform* **5**, 28 (2014).
102. Huang, Y.W., Baluna, R. & Vitetta, E.S. Adhesion molecules as targets for cancer therapy. *Histology and histopathology* **12**, 467-477 (1997).
103. Morley, S. *et al.* Trading in your spindles for blebs: the amoeboid tumor cell phenotype in prostate cancer. *Asian Journal of Andrology* **16**, 530-535 (2014).
104. Lo, B. *et al.* Lkb1 regulates organogenesis and early oncogenesis along AMPK-dependent and -independent pathways. *The Journal of Cell Biology* **199**, 1117-1130 (2012).

105. Friedl, P., Zanker, K.S. & Brocker, E.B. Cell migration strategies in 3-D extracellular matrix: differences in morphology, cell matrix interactions, and integrin function. *Microsc Res Tech* **43**, 369-378 (1998).

106. Mandeville, J.T., Lawson, M.A. & Maxfield, F.R. Dynamic imaging of neutrophil migration in three dimensions: mechanical interactions between cells and matrix. *J Leukoc Biol* **61**, 188-200 (1997).

107. Li, J. *et al.* Loss of LKB1 disrupts breast epithelial cell polarity and promotes breast cancer metastasis and invasion. *Journal of experimental & clinical cancer research : CR* **33**, 70 (2014).

108. Kakiuchi, M. *et al.* Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat Genet* **46**, 583-587 (2014).

109. Wang, K. *et al.* Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet* **46**, 573-582 (2014).

110. Han, X. *et al.* Transdifferentiation of lung adenocarcinoma in mice with Lkb1 deficiency to squamous cell carcinoma. *Nat Commun* **5**, 3261 (2014).

111. Erler, J.T. *et al.* Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* **440**, 1222-1226 (2006).

112. Haston, W.S., Shields, J.M. & Wilkinson, P.C. Lymphocyte locomotion and attachment on two-dimensional surfaces and in three-dimensional matrices. *J Cell Biol* **92**, 747-752 (1982).

113. Gadea, G., de Toledo, M., Anguille, C. & Roux, P. Loss of p53 promotes RhoA–ROCK-dependent cell migration and invasion in 3D matrices. *The Journal of Cell Biology* **178**, 23-30 (2007).

114. Kosla, J. *et al.* Metastasis of aggressive amoeboid sarcoma cells is dependent on Rho/ROCK/MLC signaling. *Cell communication and signaling : CCS* **11**, 51 (2013).

115. Rosel, D. *et al.* Up-regulation of Rho/ROCK signaling in sarcoma cells drives invasion and increased generation of protrusive forces. *Mol Cancer Res* **6**, 1410-1420 (2008).

116. Wyckoff, J.B., Pinner, S.E., Gschmeissner, S., Condeelis, J.S. & Sahai, E. ROCK- and myosin-dependent matrix deformation enables protease-independent tumor-cell invasion in vivo. *Curr Biol* **16**, 1515-1523 (2006).

117. Morris, L.G. *et al.* Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget* (2016).

118. McGranahan, N. & Swanton, C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* **27**, 15-26 (2015).
119. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A* **110**, 4009-4014 (2013).
120. Jacoby, M.A., Duncavage, E.J. & Walter, M.J. Implications of Tumor Clonal Heterogeneity in the Era of Next-Generation Sequencing. *Trends in Cancer* **1**, 231-241.
121. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* **366**, 883-892 (2012).
122. Brocks, D. *et al.* Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep* **8**, 798-806 (2014).
123. Easwaran, H., Tsai, H.C. & Baylin, S.B. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell* **54**, 716-727 (2014).
124. Yates, L.R. & Campbell, P.J. Evolution of the cancer genome. *Nat Rev Genet* **13**, 795-806 (2012).
125. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* **72**, 4875-4882 (2012).
126. Turtoi, A., Blomme, A. & Castronovo, V. Intratumoral heterogeneity and consequences for targeted therapies. *Bull Cancer* **102**, 17-23 (2015).
127. Greaves, M. & Maley, C.C. Clonal evolution in cancer. *Nature* **481**, 306-313 (2012).
128. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).
129. Yap, T.A., Gerlinger, M., Futreal, P.A., Pusztai, L. & Swanton, C. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med* **4**, 127ps110 (2012).
130. Burrell, R.A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338-345 (2013).
131. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: implications for targeted therapeutics. *British journal of cancer* **108**, 479-485 (2013).
132. Horswell, S., Matthews, N. & Swanton, C. Cancer heterogeneity and "the struggle for existence": diagnostic and analytical challenges. *Cancer letters* **340**, 220-226 (2013).
133. Wu, A.R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nature methods* (2013).
134. Kalisky, T., Blainey, P. & Quake, S.R. Genomic analysis at the single-cell level. *Annual review of genetics* **45**, 431-445 (2011).

135. Kalisky, T. & Quake, S.R. Single-cell genomics. *Nature methods* **8**, 311-314 (2011).
136. Scheel, C. & Weinberg, R.A. Phenotypic plasticity and epithelial-mesenchymal transitions in cancer and normal stem cells? *Int J Cancer* **129**, 2310-2314 (2011).
137. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nature reviews. Cancer* **12**, 323-334 (2012).
138. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-1558 (2013).
139. Torkamani, A. & Schork, N.J. Identification of rare cancer driver mutations by network reconstruction. *Genome research* **19**, 1570-1578 (2009).
140. Eifert, C. & Powers, R.S. From cancer genomes to oncogenic drivers, tumour dependencies and therapeutic targets. *Nature reviews. Cancer* **12**, 572-578 (2012).
141. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nature medicine* **17**, 313-319 (2011).
142. Martinez, P. *et al.* Parallel evolution of tumour subclones mimics diversity between tumours. *J Pathol* **230**, 356-364 (2013).
143. Konen, J. *et al.* LKB1 kinase-dependent and -independent defects disrupt polarity and adhesion signaling to drive collagen remodeling during invasion. *Molecular biology of the cell* **27**, 1069-1084 (2016).
144. Selinummi, J. *et al.* Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images. *PloS one* **4**, e7497 (2009).
145. Yuan, J., Bae, E. & Tai, X.C. in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on 2217-2224 (2010).
146. Dyson, H.J. & Wright, P.E. Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology* **6**, 197-208 (2005).
147. Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics (Oxford, England)* **27**, 431-432 (2011).
148. Wu, G., Feng, X. & Stein, L. A human functional protein interaction network and its application to cancer data analysis. *Genome biology* **11**, R53 (2010).
149. Gurskaya, N.G. *et al.* Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat Biotechnol* **24**, 461-465 (2006).
150. Hellstrom, M. *et al.* Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* **445**, 776-780 (2007).

151. Riahi, R. *et al.* Notch1–Dll4 signalling and mechanical force regulate leader cell formation during collective cell migration. *Nat Commun* **6** (2015).

152. Stabley, D.R., Jurchenko, C., Marshall, S.S. & Salaita, K.S. Visualizing mechanical tension across membrane receptors with a fluorescent sensor. *Nature methods* **9**, 64-67 (2012).

153. Galior, K., Liu, Y., Yehl, K., Vivek, S., Salaita, K. Titin-Based Nanoparticle Tension Sensors Map High-Magnitude Integrin Forces within Focal Adhesions. *Nano Letters* **16**, 341-348 (2016).

154. Wang, X., Enomoto, A., Asai, N., Kato, T. & Takahashi, M. Collective invasion of cancer: Perspectives from pathology and development. *Pathology International* **66**, 183-192 (2016).

155. Riahi, R. *et al.* Notch1-Dll4 signalling and mechanical force regulate leader cell formation during collective cell migration. *Nat Commun* **6**, 6556 (2015).

156. Tam, W.L. & Weinberg, R.A. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nature medicine* **19**, 1438-1449 (2013).

157. Benjamin, L.E., Golijanin, D., Itin, A., Pode, D. & Keshet, E. Selective ablation of immature blood vessels in established human tumors follows vascular endothelial growth factor withdrawal. *J Clin Invest* **103**, 159-165 (1999).

158. Blanco, R. & Gerhardt, H. VEGF and Notch in Tip and Stalk Cell Selection. *Cold Spring Harbor Perspectives in Medicine* **3**, a006569 (2013).

159. Hellstrom, M. *et al.* Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* **445**, 776-780 (2007).

160. Wagenblast, E. *et al.* A model of breast cancer heterogeneity reveals vascular mimicry as a driver of metastasis. *Nature* **520**, 358-362 (2015).

161. Damsky, C.H. & Werb, Z. Signal transduction by integrin receptors for extracellular matrix: cooperative processing of extracellular information. *Current opinion in cell biology* **4**, 772-781 (1992).

162. Werb, Z., Tremble, P.M., Behrendtsen, O., Crowley, E. & Damsky, C.H. Signal transduction through the fibronectin receptor induces collagenase and stromelysin gene expression. *The Journal of Cell Biology* **109**, 877-889 (1989).

163. Aceto, N. *et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110-1122 (2014).

164. Hou, J.M. *et al.* Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. *J Clin Oncol* **30**, 525-532 (2012).
165. Cheung, K.J. *et al.* Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad Sci U S A* **113**, E854-863 (2016).
166. Maddipati, R. & Stanger, B.Z. Pancreatic Cancer Metastases Harbor Evidence of Polyclonality. *Cancer Discov* **5**, 1086-1097 (2015).
167. Cheung, K.J. & Ewald, A.J. A collective route to metastasis: Seeding by tumor cell clusters. *Science* **352**, 167-169 (2016).
168. Margulis, L. *Symbiosis in cell evolution : life and its environment on the early Earth*. (W. H. Freeman, San Francisco; 1981).
169. Axelrod, R., Axelrod, D.E. & Pienta, K.J. Evolution of cooperation among tumor cells. *Proc Natl Acad Sci U S A* **103**, 13474-13479 (2006).
170. Chapman, A. *et al.* Heterogeneous tumor subpopulations cooperate to drive invasion. *Cell Rep* **8**, 688-695 (2014).
171. Marusyk, A. *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* **514**, 54-58 (2014).
172. Cleary, A.S., Leonard, T.L., Gestl, S.A. & Gunther, E.J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113-117 (2014).
173. Gatenby, R.A., Gawlinski, E.T., Gmitro, A.F., Kaylor, B. & Gillies, R.J. Acid-mediated tumor invasion: a multidisciplinary study. *Cancer Res* **66**, 5216-5223 (2006).
174. Han, T. *et al.* How does cancer cell metabolism affect tumor migration and invasion? *Cell Adhesion & Migration* **7**, 395-403 (2013).
175. Riganti, C., Gazzano, E., Polimeni, M., Aldieri, E. & Ghigo, D. The pentose phosphate pathway: an antioxidant defense and a crossroad in tumor cell fate. *Free radical biology & medicine* **53**, 421-436 (2012).
176. Wang, J.B. *et al.* Targeting mitochondrial glutaminase activity inhibits oncogenic transformation. *Cancer Cell* **18**, 207-219 (2010).
177. Lee, J.H. *et al.* Energy-dependent regulation of cell structure by AMP-activated protein kinase. *Nature* **447**, 1017-1020 (2007).

178. Alcala, S.E. *et al.* Mitotic asynchrony induces transforming growth factor-beta1 secretion from airway epithelium. *American journal of respiratory cell and molecular biology* **51**, 363-369 (2014).

179. Duckworth, C.A. *et al.* Progastrin-induced secretion of insulin-like growth factor 2 from colonic myofibroblasts stimulates colonic epithelial proliferation in mice. *Gastroenterology* **145**, 197-208, e193 (2013).

180. Gregori, J. *et al.* Enhancing the Biological Relevance of Secretome-Based Proteomics by Linking Tumor Cell Proliferation and Protein Secretion. *Journal of proteome research* (2014).