

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:

Elizabeth L. Zoeller

Date

Lysine demethylase 5B as a mediator of collective invasion

By

Elizabeth L. Zoeller
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Cancer Biology

Paula Vertino, Ph.D.
Advisor

Xiaodong Cheng, Ph.D.
Committee Member

Jin-Tang Dong, Ph.D.
Committee Member

Sumin Kang, Ph.D.
Committee Member

Adam Marcus, Ph.D.
Committee Member

Accepted:

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

Date

Lysine demethylase 5B as a mediator of collective invasion

By

Elizabeth L. Zoeller
B.S., University of North Carolina at Chapel Hill, 2008

Advisor: Paula Vertino, 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
2019

Abstract

Lysine demethylase 5B as a mediator of collective invasion

By Elizabeth L. Zoeller

Metastatic disease is the primary cause of death in cancer patients. Therapeutic targeting of the metastatic process has not performed well because of the complexity of the multi-step process and difficulty identifying molecular targets. Recent discoveries point to collective cell invasion, in which a group of cells invade in a coordinated, cooperative manner and can contain distinct functional populations such as cells that lead and cells that follow, as key to successful metastases formation both during the initial invasive steps leaving the primary site and the later establishment of tumors at a distant site. Given the many adaptations to changes in environment required of metastatic cells, survival of these cells depends on plasticity in gene expression, making epigenetic regulators promising anti-metastatic targets. In fact, many inhibitors of epigenetic enzymes are currently in development as anti-cancer treatments. Among these, lysine demethylase 5B (KDM5B) and fellow KDM5 family members are being rigorously investigated in efforts to develop effective and selective inhibitors. KDM5 family members are overexpressed across many cancer types. Moreover, KDM5B has been positively associated with invasive cell behavior in multiple cancers. To clarify the role of KDM5B in collective invasion, we utilized a non-small cell lung cancer model of leader and follower populations isolated from cells invading as a collective chain. Followers expressed less KDM5B overall and a mutant version of KDM5B that was absent from leaders, leading us to hypothesize that intact KDM5B is a requirement of leader cell behavior and decreased KDM5B activity defines follower cells. By measuring invasive behaviors after overexpressing wild-type or mutant KDM5B across leader and follower cells, we found that chain formation was enhanced by the presence of both KDM5B active and KDM5B deficient populations. Next, to determine the feasibility of preventing collective invasion through chemical inhibition of KDM5 enzymes, we tested several compounds in our invasive model. Although we did not successfully identify a treatment protocol against invasive behaviors, KDM5B and family members remain promising anti-metastatic targets. Here we uncover a novel role for KDM5B in collective invasion and take a step towards understanding and preventing the occurrence of metastasis.

Lysine demethylase 5B as a mediator of collective invasion

By

Elizabeth L. Zoeller
B.S., University of North Carolina Chapel Hill, 2008

Advisor: Paula Vertino, 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
2019

Table of contents

Chapter 1: Introduction	1
1.1 Metastasis and mortality.....	2
1.2 Heterogeneity in cancer.....	2
1.3 EMT and the metastatic process.....	3
1.4 Collective invasion and the origins of metastatic lesions.....	5
1.5 Leader and follower cells.....	7
1.6 Epigenetics in cancer.....	10
1.7 Functions of the lysine demethylase 5 family.....	12
1.8 Roles of lysine demethylase 5B in cancer.....	14
1.9 Dissertation goals.....	16
Chapter 2: Genetic heterogeneity within collective invasion packs drives leader and follower cell phenotypes	21
2.1 Abstract.....	22
2.2 Introduction.....	22
2.3 Results.....	24
2.4 Discussion.....	36
Chapter 3: Investigating KDM5B as an anti-metastatic therapeutic target	63
3.1 Introduction.....	64
3.2 Results.....	65
3.3 Discussion.....	70
Chapter 4: Discussion	86
4.1 KDM5B as a mediator of collective invasion.....	87
4.1.1 Summary of findings.....	87
4.1.2 Models of heterogeneity and cooperativity in invasion.....	93
4.1.3 Future considerations.....	97

4.2 KDM5B as a therapeutic target.....	101
4.2.1 Summary of findings.....	101
4.2.2 Future considerations.....	103
Chapter 5: Materials and methods.....	108
References.....	118

List of figures

Figure 1.1 Source and sink model of chemotaxis in cell migration and in melanoma.....	17
Figure 1.2 The process of contact inhibition of locomotion.....	18
Figure 1.3 Force generation, mechanosensation, and distribution during collective invasion..	19
Figure 1.4 Schematic of protein domains of KDM5B.....	20
Figure 2.1. ARP3 K240R and KDM5B L685W are validated mutations in H1299 leader and follower cells.....	42
Table 2.1 RNA-seq reveals leader- and follower-specific gene mutations.....	44
Figure 2.2. PTM hotspot analysis of ARP3 K240 suggests functional impact of the K240R mutation.....	45
Figure 2.3. ARP3 K240R confers leader-like properties when expressed in follower cells.....	47
Figure 2.4. Characterization of the follower-enriched KDM5B L685W mutation.....	49
Figure 2.5. Wild-type KDM5B suppress and KDM5B L685W drives chain-like invasion.....	50
Figure 2.6. Leader and follower cells are derived from two separate populations defined by mutational profile.....	52
Figure S2.1. Confirmation of leader- and follower-enriched mutations.....	53
Figure S2.2. ARP3 knockdown inhibits 3-D invasion.....	54
Figure S2.3. EV-rescued followers do not lead invasive chains.....	56
Figure S2.4. KDM5A but not KDM5C is downregulated in follower cells.....	57
Figure S2.5 Overexpression of wild-type KDM5B or KDM5B L685W do not affect growth rate.....	58
Figure S2.6. Wild-type KDM5B suppress and KDM5B L685W drives chain-like invasion in parental spheroid.....	59
Figure S2.7. Leader cells overexpressing wild-type KDM5B but not KDM5B L685W exhibit diminished chain formation but retain leader activity in leader-only spheroids.....	60
Figure S2.8. The relative proportions of the wild-type and variant mRNA expression reflects of the proportion at the genomic DNA level across all three populations.....	62
Figure 3.1 GSK-J4 and JIB-04 inhibit of breast cancer cell.....	75

Figure 3.2 KDM5B is overexpressed in H1299 leader cells.....	76
Figure 3.3 Knockdown of KDM5B produces varying results between shRNAs of different targeting sequences.....	77
Figure 3.4 Knockdown of KDM5B does not alter growth rate.....	79
Figure 3.5 Knockdown of KDM5B does not change the percentage of chains led by introduced leader cells.....	80
Figure 3.6 Treatment with KDM5 inhibitor KDM5-C70 fails to change invasive behavior	81
Figure 3.7 Recently developed KDM5 inhibitors suggest invasion can be inhibited chemically.....	83
Figure 3.8 CPI-48 does not significantly alter invasion.....	85
Figure 4.1 Model of invasive cooperativity within a heterogenous population.....	106
Figure 4.2. KDM5B may regulate leader behavior through HOXA5 and VEGF.....	107

List of abbreviations in alphabetical order:

ARP3 Actin relate protein 3

CIL contact inhibition of locomotion

CpG cytosine residue followed by a guanine residue from 5' to 3'

CSC Cancer stem cell

DMSO Dimethyl sulfoxide

DNMT DNA methyltransferase

ECM Extracellular matrix

EMT Epithelial-to-mesenchymal transition

FACS Fluorescence-activated cell sorting

FPS Function Probability Score

GI₅₀ half-maximal growth-inhibitory concentration

H2A Histone 2A

H2B Histone 2B

H3 Histone 3

H4 Histone 4

H3K4me3 Histone 3 lysine 4 trimethylation

H3K4me2 Histone 3 lysine 4 dimethylation

H3K4me1 Histone 3 lysine 4 monomethylation

HDAC Histone deacetylase

HAT Histone acetyltransferase

ITH Intratumor heterogeneity

JmjC Jumonji C domain

K240R Lysine to arginine substitution at residue 240

KDM Histone lysine demethylase

KDM5 Lysine demethylase 5 family

KDM5A Lysine demethylase 5A

KDM5B Lysine demethylase 5B

KDM5C Lysine demethylase 5C

KDM5D Lysine demethylase 5D

KMT Histone methyltransferase

L685W Leucine to tryptophan substitution at residue 685

LPA Lysophosphatidic acid

MBD Methyl-CpG binding domain

MET Mesenchymal-to-epithelia transition

NE Neuroendocrine cell type

Non-NE Non-neuroendocrine cell type

NSCLC Non-small cell lung cancer

OD Optical density

PTM Post-translational modification

SaGA Spatiotemporal Genomic and Cellular Analysis

SASA solvent accessible surface area

SCLC Small cell lung cancer

SRB Sulforhodamine B assay

TET Ten-eleven translocation enzyme

VAF Variant allele frequencies

VEGF vascular endothelial growth factors

Chapter 1

Introduction

1.1 Metastasis and mortality

Cancer is the second highest cause of mortality in the United States with more than 1.7 million new cases and over 600,000 deaths projected to occur in 2019 [1]. Of these cases, the vast majority of deaths are attributable to metastatic spread from the primary tumor site to distant organs in the body [2]. Metastasis occurs as cancer cells leave the tumor and organ in which they developed and move to another organ where they proliferate and establish new tumors at this secondary site. Metastasis becomes fatal when the growing metastatic tumors overwhelm the organs that house them and shut down vital functions for the human body.

The deadly effects of metastasis are well exemplified by the course of lung cancer. Within all cancer types, lung cancer accounts for one quarter of deaths caused by cancer and has less than a twenty percent five year survival rate [1]. If lung cancer is diagnosed when the disease remains localized to the primary site, the five-year survival rate is 54% [3]. However, as the disease progresses to a regional stage and then to stages with distant metastasis, this survival rate drops from 26% to 4% respectively [3]. Metastatic spread of tumor cells is clearly associated with decreased survival. Targeting metastasis is an opportunity to save lives.

1.2 Heterogeneity in cancer

Our comprehension of the biology of cancer is further complicated by intratumor heterogeneity (ITH). ITH arises as genetic changes occur in a cell that expands to form a subclonal population within a tumor. ITH itself varies in degree from cancer type to cancer type and from tumor to tumor [4]. One study of glioma found that from 0 to 8,000 mutations were heterogeneous either in the primary tumor or between the primary tumor and recurrent or metastatic sites [5]. Within non-small cell lung cancer (NSCLC), two recent studies that sequenced multiple regions of primary lung tumors found that around 24% to 30% of detected mutations were omitted from at least one of the

sampled regions [6, 7]. These results indicate that almost a third of mutations occur in spatially distinct subclonal populations and could be missed if only one region of tumor were sampled. Recent advances in technology have greatly increased our ability to detect ITH. Multiregional sequencing, genetic screening of single cells, or single-cell–derived clones, and single-cell sequencing are some of these advanced technical tools [8-17]. Bioinformatic tools to process these sequencing data have also allowed for great advances in understanding the ITH landscape across tumors [4, 18-23]. However, our knowledge of ITH will always be limited to the samples that are collected for sequencing—that is multiple regions of tumors must be collected to fully capture degree of heterogeneity [4].

An important aspect of ITH is that this heterogeneity offers diverse phenotypes on which evolutionary selection can act [24, 25]. After the initial development of cancer and acquisition of malignant traits, most models of ITH are consistent with neutral growth—i.e. diversity is maintained because little to no selective pressure causes a single subclone to outcompete its neighbors [8, 26-28]. However, some conditions can exert selective pressure on certain subclonal populations [29, 30]. Two of the greatest sources of selective pressure are clinical treatment of the tumor and invasion/migration of tumor cells into a new environment [31, 32]. While the dominant subclone or subclones may be treatment sensitive, a lesser subclone may have acquired mutations that enable its daughter cells to thrive and expand even in the context of the therapeutic treatment [33-37]. In the case of invasion and metastasis, cells are entering a new environment and must demonstrate traits unknown to the bulk of the primary tumor [38-41]. As diversity in genotype increases so does the probability that one of these subclones gains the ability to resist treatment or to survive in a novel environment.

1.3 EMT and the metastatic process

The multiple steps leading to successful metastasis are collectively termed the metastatic cascade [42-45]. To complete the metastatic cascade, cancer cells must invade the local tissues

surrounding the primary tumor site, intravasate into the circulatory system, survive transit while circulating, including a possible period of dormancy that could last up to several years, arrest their circulation within the bloodstream, extravasate from the vascular walls and into distant tissues, establish a microscopic colony, and finally, survive and proliferate to form a detectable metastatic tumor [45]. Given the complexity of this route, metastasis is unsurprisingly an inefficient process with some estimates that as few as 0.01% of cells that leave the tumor and enter the blood stream actually develop into a secondary tumor [46, 47]. And yet, metastasized tumors are still major killers. Further reducing the efficiency of this process by close examination of each step is necessary for successfully targeting metastasis.

Much focus of research has been placed on the first step of the metastatic cascade—where a cell changes from fixed to motile. This change marks a fundamental shift in cell identity from a cell with no potential to metastasize to one that could. As tumor cells rapidly proliferate, they outgrow their supply of oxygen and nutrients provided by their inceptive vasculature [43, 48, 49]. Cancer cells often resort to several common strategies to handle the hypoxic environment—grow a new blood supply (i.e. angiogenesis), alter their metabolism, or escape the primary tumor site [48, 49]. In the context of metastasis, the last mode of accommodation is most relevant. Central to this shift is the epithelial-to-mesenchymal transition (EMT). Indeed, protein markers of migration and EMT, such as Twist, Snail, and COOH-terminal binding proteins, are upregulated in human tumor samples in the context of hypoxia [50-54]. While EMT has an important role during normal development as in embryogenesis and during wound healing in the adult, this programmed process is exploited by cancer cells to shed epithelial traits and gain those of mesenchymal cells as they move out of their native environment [55].

EMT is a spectrum of acquired traits with a number of intervening metastable phenotypes [56, 57]. At the epithelial end, cells demonstrate apical-basal polarity, a high number of cell-cell junctions,

and a noticeable lack of motility. At the mesenchymal end of the spectrum, cells display front-back polarity, lack cell-cell junctions, are capable of degrading and passing through the extracellular matrix, and importantly, are highly motile [55, 58-60]. Molecularly, epithelial cells are often defined by the expression of E-cadherin, occludins, and cytokeratins, whereas N-cadherin and vimentin mark the mesenchymal state [61].

When cancer cells halt their movement to sprout distant metastases, these cells must reestablish their stationary, epithelial traits in process called the mesenchymal-to-epithelial transition (MET) [62-68]. Provided the circulating tumor cells survive the physical forces and potential attacks from the immune system that cells are subject to after entering the blood stream, cells must migrate through the endothelial wall to extravasate [45, 69, 70]. At this point, many cancers may fail to thrive or enter a dormant state [71]. Those cells that do develop into full-fledged metastatic cancer likely regain some of the epithelial traits that made the primary tumor so successful at proliferation [45, 62]. In point of fact, quite a number of studies have shown that a MET reprogramming is essential for establishing a metastatic colony [45, 63, 64, 67, 72].

1.4 Collective invasion and the origins of metastatic lesions

While invasion by single cancer cells has been well studied, collective invasion by groups of cancer cells has recently become recognized as an important mode of movement. During normal development and malignant progression, cohesive groups of cells move in a coordinated, cooperative manner [73]. This process is called “collective migration” in the context of normal development and other benign conditions, such as wound healing, and called “collective invasion” when the cells are cancerous and invading the surrounding stroma [74]. During embryogenesis, collective migration occurs in such processes as migration of neural crest cells and morphogenesis of ducts, glands, and vessels [75-80]. Collective invasion is characterized by the maintenance of cell-cell junctions

throughout the movement of a group of cells and by the supracellular organization mediated by the actin cytoskeleton that integrates cell forces along a front to rear polarity [75, 81]. Cancer cells co-opt the programming underlying these processes in normal development much like that which occurs during EMT as they invade collectively out of a tumor.

Despite having a higher velocity, single cell migration is less efficient than collective migration due to the quick movements of single cells along an inconstant course [73]. Collective migration is a more persistent, directional movement of cells [73]. Collective migration has even been proposed as the predominant form of invasion leading to established metastasis [82]. Early in modern studies of cancer, single cells and cluster of cancer cells were found in blood samples from cancer patients [44]. Early animal models demonstrated that these clusters could also survive passage through the lung and were more efficient at producing metastases than single cells following tail vein injection [83, 84]. More recent studies have focused on clonality within metastases. If just one single cell seeds a metastatic site, the resulting tumor would be monoclonal, but if seeded by a cluster of cells, the secondary tumor may be polyclonal [82]. In prostate cancer, frequent polyclonal seeding from the primary tumor has been found at secondary sites and from existing metastases to new sites in the body [85].

The possibility that a polyclonal metastatic site results from multiple re-seedings by single cells exists. Mouse models involving the use of primary tumors comprised of cells expressing fluorescent proteins of varying colors have helped untangle this conundrum [82, 86-88]. Three independent research groups found multicolor metastases when the primary tumor was multicolored, demonstrating that the metastases were indeed polyclonal as is consistent with seeding by a cluster of cells [86-88]. Moreover, one group working in an orthotopic breast cancer model established that when creating single color tumors (one color tumor per side, separate colors per tumor), mostly single-color metastases form, indicating a single tumor origin [86]. Similar results were obtained in an autochthonous pancreatic cancer model expressing a multi-colored lineage labeling system [87].

Strikingly, by quantifying the relationship of the extent of clonal mixing in the primary tumor with the frequency of detection of multicolored metastases in a spontaneous breast cancer model also utilizing a multi-colored lineage labeling system, a further study estimated that more than 97% of metastases were derived from tumor cell clusters [88]. Thus, collective groups of cells invading and migrating together through the body are significant contributors to the formation of metastases.

1.5 Leader and follower cells

Cells that are migrating (or invading) collectively can be divided into two functionally discrete groups—leaders at the front of the pack leading along the invasive/migrating edge and the followers trailing behind, attached by cell-cell junctions [89]. Leader cells are called such because they are located at the leading edge of invasion, sense migration cues in the microenvironment, and control where and how fast the group of cells move [73]. Each leader cell is polarized into two distinct regions: the leading free edge of the cell which is exposed to the surrounding environment, including a plethora of external signals such as chemoattractants, and the trailing portion of the cell in which cell-cell junctions form connections with the follower cells and delimit the range of motion of the leading cell [73, 74, 89, 90]. Follower cells, by default, are the cells that follow the leaders behind the invasive front but may have a more significant role than just that of passive passengers.

In fact, cells are selected as leaders by their responses to these external cues, which include the extracellular matrix (ECM), soluble factors, and even interactions with other cells [73]. Integrins act to mediate signaling between cells and the ECM and respond to the stiffness and composition of the ECM [91, 92]. During intradermal tumor xenografts in mice, the ECM provided directional cues to multicellular streams of cells [73, 93, 94]. Integrin-mediated signaling triggered by interactions with the ECM during *in vitro* wound healing assays resulted in rearrangements of the cytoskeleton, structural reorganization, and morphological polarization—canonical characteristics of leader cells [73, 90, 95].

In turn, leaders also modify and remodel the ECM as they travel through the ECM and enlarge the migration tract. Leaders alter the ECM through the focal-adhesion generated traction which reorients matrix fibers and thereby promotes directional guidance and cell streaming. Leaders further alter the ECM through the secretion of matrix metalloproteases, which directly remodel ECM fibers, and secretion of ECM components, which change the nature and structure on which the cells migrate and increases the polarity of the migrating cell pack [73, 93, 94, 96, 97].

Additionally, collective migration can be stimulated by soluble factors, such as nitrous oxide, vascular endothelial growth factors (VEGFs), fibroblast growth factor, epidermal growth factor, and several other cytokines [90, 98-109]. These soluble factors can promote cell polarization and/or protrusions, often through downstream signaling aimed at actin polymerization, and active intracellular signaling which controls gene expression patterns characterizing leader cell behavior [73, 110-113]. Movement along a chemical gradient is called chemotaxis and requires a source and a sink (Fig. 1.1A) [114]. Within a collectively migrating pack, the follower cells can aid in the creation of a chemical gradient by acting as a sink, by removing the chemical or signaling molecular from the environment, either by destruction or expressing substrate-binding receptors unique to the rear of a pack [73, 115, 116]. One example of such behavior occurs in melanoma cells for which lysophosphatidic acts strong attractant and which break down lysophosphatidic acid (LPA), resulting in a higher concentration of this attractant outside of the group of cells than within the group (Fig. 1.1B) [117].

Followers further polarize leader cells through the use of contact inhibition of locomotion (CIL). CIL occurs when a moving cell encounters another cell and changes its trajectory of motion away from the second cell (Fig. 1.2A). This process is an active one in which cell protrusions collapse at the side of collision and extend from the opposite end and is mediated by RHOA and RAC signaling [118]. At the point of contact, interactions of cell surface molecules, including cadherins [115, 119,

120], ephrin and EPH receptors [121-123], members of the planar cell polarity pathway [124-126], and syndecan 4 [127], ultimately lead to the activation of RHOA and inhibition of RAC around the contact (Fig. 1.2B) [73, 128-130]. This switch in the activity of RHOA and RAC creates major changes in the cytoskeleton through microtubule catastrophe and disassembly of focal adhesions, and actin contractility and collapse of protrusions at the contact site [128, 131-134]. Conversely, activation of RAC at the end opposite the contact results in microtubule and microfilament polymerization, stabilization of focal adhesions, and generation of new protrusions [129, 135-137]. The new protrusions complete the last step of CIL by permitting movement away from the site of contact.

Some studies of leader cells have demonstrated that they can generate enough force to drag a large number of followers behind them [138]. At their leading edge, leader cells generate traction forces through focal adhesions which associate with acto-myosin cables to promote contraction of the cell body and to transmit these forces across several rows of followers (Fig. 1.3A,B) [139, 140]. Followers are able to respond to mechanical stimuli through mechanotransducers, such as talin in focal adhesion and α -catenin in adherens junctions, which change conformation because of the stretching force applied to them [141, 142]. For instance, stretching α -catenin exposes a protein domain capable of binding vinculin, which permits more recruitment of actin and strengthens the adherens junction (Fig. 1.3C) [143-145]. Stress forces distribute in manner across several rows of cells in a monolayer that cannot be explained by a model in which only leader cells generate traction, thus implicating that follower cells themselves can generate some traction [146]. One concept that can explain how these follower cells may also create pulling forces in the same direction as leader cells is that of plithotaxis, in which cells move in the direction of highest principle stress as a means to minimize intercellular shear stress [81, 147-149]. All in all, there are many physical forces and molecules contributing to the physiology of collective invasion.

1.6 Epigenetics in cancer

Epigenetics refers to heritable changes in gene expression and chromatin organization not caused by alterations in the DNA sequence [150]. The epigenome maintains differentiated states in normal tissues but, when aberrantly regulated, contributes significantly to the development and progression of cancer [151]. According to recent next-generation sequencing studies, more than half of human cancers possess mutations in the enzymes that organize the chromatin [151-153]. Additionally, the epigenome of cancer cells is profoundly disrupted when compared to that of normal tissues [152, 154, 155]. A fundamental aspect of epigenetic profiles is their reversibility. As epigenetic proteins regulate gene transcription and cell fate, their activity can profoundly alter phenotype in a relatively short period of time [156]. Moreover, these changes in phenotype can be reversed. The term plasticity encompasses this ability of cancer cell cells to change phenotype [156]. This plasticity is essential as cells transition between epithelial and mesenchymal states or adapt and develop resistance to therapeutics [57, 157].

The basic unit of chromatin at which epigenetic enzymes act is the nucleosome, which consists of the 147 base pairs of DNA wrapped around this protein complex and the eight histone proteins (two each of H2A, H2B, H3, and H4) that compose the complex [158]. Epigenetic enzymes regulate the relative accessibility of genes to transcription machinery and are divided into three functional classes—writers, readers, and erasers [159]. Writers covalently modify the molecules of the nucleosome, readers recognize these modifications and recruit more proteins to further modify or remodel the surrounding genomic region, and finally, erasers remove the modifications, which allows for plasticity in gene expression [151, 159]. Each of these steps represent a potential therapeutic target—either in a single agent or combination therapy protocol.

For DNA, cytosine residues, when followed immediately by guanine residues (CpG site), may be modified by methylation through DNA methyl transferase enzymes (DNMTs) [160]. DNA

methylation is then read by proteins containing methyl-CpG binding domains (MBD) [161]. Ten-eleven translocation (TET) enzyme family members iteratively oxidize the methylated cytosines [150, 162-164]. Each of these oxidized modifications may have a role in regulating gene expression, but importantly, TET enzymes act as erasers of DNA methylation by creating a hydroxymethylated cytosine that can be targeted for removal by DNA repair machinery [165-169]. The interpretation of DNA methylation is context dependent. DNA methylation in promoters containing a large proportion of CpG sites (so called CpG islands) represses gene expression [170]. However, in actively transcribed gene bodies, CpG sites are often methylated [170]. DNMTs were among the first epigenetic targets for which cancer chemotherapeutics were developed. The DNMT inhibitors azacytidine and decitabine are cytidine analogs that trap DNMTs onto the DNA in which these analogs are incorporated and lead to the degradation of DNMTs [171]. These inhibitors are broad reprogrammers that cause large scale changes in gene expression and target all cells entering S-phase as opposed to cancer cells expressing particular gene profiles [151, 172]. While approved for use in myelodysplastic syndrome and acute myeloid leukemia, DNMT inhibitor use is limited in solid tumors due to their toxicity, short half-lives, and possibly lack of specificity [151, 173-179].

Histones are the other major components of the nucleosome. Eight histones make up a nucleosome with two each of four different proteins, histone 2A (H2A), histone 2B (H2B), histone 3 (H3), and histone 4 (H4) [150]. The N-terminal tail of histones are loosely structured and provide an easily accessible substrate for post-translational modifications and proteins (such as readers, erasers, and chromatin remodelers) to bind to these modifications [180]. Histones are subject to at least 16 classes of modification at over 60 documented residues [150, 180, 181]. As with DNA modifications, the consequences of histone modifications are dependent upon the position of the modified residue on a histone tail and the modifications to other surrounding amino acids [158, 180]. Lysine methylation and lysine acetylation are the primary histone modifications and currently the most targeted for

therapeutic intervention [150, 151, 159]. Histone lysine methyltransferases (KMTs) and histone lysine acetyltransferases (HATs) are writers for these modifications, histone lysine demethylases (KDMs) and histone deacetylases are their erasers, and a large number of protein domains are able to detect these marks [151].

Broad inhibitors of all HDACs have been approved for lymphoma and multiple myeloma for several years now [151, 182]. Additionally, small molecule inhibitors that target the bromodomains of all BET proteins, which read acetylated lysines, are also in clinical trials [151]. Furthermore, readers, writers, and erasers of histone modifications offer the opportunity to develop therapies specifically targeted to cancer cells, which may only overexpress a particular histone modifying protein. For example, several inhibitors of EZH2, an HMT that methylates H3K27, are in clinical trials as this enzyme is positively correlated with poor prognosis and tumor aggressiveness across several cancer types and is subject to activating mutations in multiple types of lymphoma [151, 159, 183-192]. Currently, inhibitors of DOT1L, an H3K27 methyltransferase, and KDM1A, an H3K4 and H3K9 demethylase (also called LSD1), are also being tested in clinical trials [151, 159, 193-201]. Many more inhibitors with many more histone-related targets are in pre-clinical research across the globe.

1.7 Functions of the lysine demethylase 5 family

The lysine demethylase 5 family (KDM5s) consists of four closely related proteins (KDM5A, KDM5B, KDM5C, and KDM5D) which remove mono-, di-, and tri-methylation of H3K4 [202-207]. While traditionally considered transcriptional repressors because of their tendency to demethylate activating marks H3K4me_{2/3} from promoters, KDM5 can be activating or repressing depending on the context of the target methylation site [208]. Members of this family are additionally characterized by a split catalytic Jumonji domain with a DNA-binding ARID, AT-rich interacting domain, and PHD finger (PHD1) intervening between the N-terminal (JmjN) and C-terminal (JmjC), as well as a C₅HC₂

zinc finger and an additional one (PHD2) or two PHD fingers (PHD3, only in KDM5A and KDM5B) (Fig. 1.4) [202, 204, 208-212]. Besides the two halves of the Jumonji domain, the C₅HC₂ zinc finger is necessary for catalytic activity of the protein [204, 213]. The PHD domains aid in substrate recognition. In KDM5B and to some extent in KDM5A, the PHD1 domain recognizes H3K4me₀, which may stabilize target gene repression by preventing re-methylation or promoting removal of adjacent H3K4me_{2/3}, but in KDM5C, the PHD1 domain binds H3K9me_{2/3} [206, 207, 212, 214, 215]. The PHD3 domains of KDM5A and KDM5B bind the substrate H3K4me₃ [214].

The KDM5s play significant roles in normal development. While KDM5A knockout mice appear to be without defect, KDM5A is ubiquitously expressed in adult tissues, binds Rb to promote differentiation, and along with KDM5B, helps activate the zygotic genome, with peak expression at the blastocyst stage [216-220]. In adult tissues, KDM5B is only expressed highly in the testis and pregnant mammary gland and moderately expressed in the eye, prostate, and ovary with expression varying temporally and spatially throughout the embryo in early development [221, 222]. Studies of stem cells have shown that loss of KDM5B extends self-renewal in embryonic stem cells and accelerates induced pluripotent stem cell reprogramming, indicating that KDM5B has a central role in regulating differentiation [223, 224]. Three reports of KDM5B knockout mice demonstrated great variance in phenotype. In the first, complete KDM5B knockout, generated by replacing exon 1 through flanking the loxP-neomycin-LoxP expression cassette in a C57bl/129/ola mouse background, was early embryonic lethal around day E4.5 to E7.5 [216]. The second KDM5B knockout study, which utilized a lacZ-Neo-reporter cassette flanked by FRT sites was inserted between exons 5 and 6 with exon 6 of KDM5B was flanked by loxP sites to create a frameshift mutation and subsequent termination mutation in a C57BL/6N; Rosa26::CreERT2 mouse background, found major neonatal lethality due to respiratory failure and neural and skeletal defects [225]. By generating KDM5B knockdown mice of C57BL/6 through a gene trap cassette to create a truncated protein, the third

study produced mostly viable KDM5B null mice with decreased body weight, reduced female fertility, diminished serum estrogen levels, and delayed mammary development with supporting evidence of defects in mammary morphogenesis and gene expression signifying that KDM5B positively regulates mammary development, results supported by a mouse model expressing ARID deleted KDM5B [216, 222]. The differences between these studies have been attributed to differences in genetic background of the mice and colony maintenance [222, 226]. Expressed throughout adult tissues and necessary for embryonic development in mice, KDM5C, which is located on the X chromosome and is one of the few genes to escape X inactivation, has multiple documented inherited mutations linked to intellectual disability in humans [227-231]. Moreover, KDM5A and KDM5B have too been associated with disorders of intellectual disability [232-234]. The least studied member of this family is KDM5D. Located on the Y chromosome, KDM5D is expressed in all male tissues and is important in spermatogenesis [235].

1.8 Roles of lysine demethylase 5B in cancer

All four KDM5 family members have established roles in carcinogenesis. KDM5A is overexpressed in glioblastoma and gastric and hepatic cancers, positively regulates the expression of VEGF and angiogenesis, and positively regulates EMT and metastasis [236-242]. KDM5C is either a tumor suppressor or oncogene depending on context. In clear cell renal carcinoma, loss of KDM5C enhances transcription of heterochromatic noncoding RNAs and consequently leads to genomic instability [243-247]. Yet, KDM5C expression is also associated with poor prognosis in breast and prostate cancer and drives invasion in hepatocellular carcinoma [248-250]. Loss of KDM5D promotes prostate cancer and clear cell renal carcinoma [250-253]. However, of all the KDM5 proteins, KDM5B has the most established role in carcinogenesis and cancer progression. Initially, KDM5B was identified as a gene overexpressed in breast cancer, years before the first instance of lysine

demethylation was detected [254]. In the ensuing years, KDM5B has also been found to be overexpressed in melanoma, glioma, liver cancer, lung cancer, bladder cancer, colorectal cancer, gastric cancer, prostate cancer, ovarian cancer, osteosarcoma, head and neck cancer [255-263]. KDM5B represses tumor suppressors BRCA1, CAV1, and HOXA5 by directly demethylating H3K4 at these genes [204, 264, 265]. KDM5B overexpression can also result in increases of downstream targets, such as cell cycle regulators E2F1 and E2F2 and the long non-coding RNA MALAT1, which promotes migration and invasion [266-268]. Moreover, KDM5B is positively associated with EMT, migration, and invasive behavior in lung cancer, gastric cancer, hepatocellular carcinoma, breast cancer [259, 268-273].

Fascinatingly, KDM5 members have repeatedly been implicated in the persistence of therapy resistant cancer cell populations. Recently, KDM5A was found to support a small subpopulation of reversibly drug-tolerant cells in cell lines across several cancer types (lung cancer, gastric cancer, melanoma, colorectal cancer, and breast cancer) and treated with several targeted and DNA damaging therapies [274]. KDM5A has also been linked to resistance to DNA damaging agent temozolomide in glioblastoma and to EGFR inhibitor erlotinib in breast cancer [236, 275]. Similarly, KDM5B marks a slow-cycling, multi-drug resistant subpopulation of melanoma cells and is associated with chemoresistance in ovarian cancer, neuroblastoma, and breast cancer [249, 276-279]. Moreover, KDM5B is connected to radioresistance in oral squamous cell carcinoma and non-small cell lung cancer through its involvement in recruitment of DNA repair machinery [280, 281]. To combat therapy resistance and to target other oncogenic effects of KDM5 family members, a number of inhibitors are currently in development to curtail the KDM5 demethylase activity [282-286]. However, specificity in targeting just the KDM5 family, let alone individual members, has been limiting in this endeavor [208, 287].

1.9 Dissertation goals

The goals of this dissertation are to establish how KDM5B defines leader and follower cancer cell subtypes and drives their unique invasive behaviors and to assess the feasibility of KDM5B as an anti-metastatic therapeutic target. Previous work in H1299 non-small cell lung cancer (NSCLC) cells enabled us to examine leader and follower cells separately, both to understand the biology that differentiates the subpopulations and to determine how they differ in their behavior when cultured separately [288]. Leader cells invade extensively in chains and divide slowly, often with mitotic defects. Follower cells invade far less and maintain a round spheroid but proliferate much more quickly. Moreover, the cells have several underlying genetic differences, including a mutation in KDM5B, as shown in Chapter 2. These phenotypes are sustained over several passages, and study of them offers the possibility of discovering novel mechanisms by which known proteins contribute to metastasis. Firstly, in Chapter 2, we confirm that previously identified mutations from early RNA-seq performed on leader and follower cells are exclusive to and delimit each subpopulation through extensive resequencing. Secondly, the work in Chapter 2 determines the extent to which these mutations drive the disparate invasive behaviors of leader and follower cells. We ask two important questions. Can we make follower cells behave like leader cells, building and leading invasive chains, if we express leader specific mutations in the follower cells? Can we make leaders stay close to home, tight to the spheroid, and give up chain building like follower cells by forcing these leaders to express a follower specific mutation? Lastly, in Chapter 3, we ask if KDM5B is a good target for chemical inhibition of invasion and metastasis. While a large number of compounds are in the earliest stages of preclinical development, most focus on an endpoint of cell death or growth inhibition. We question instead if some of these compounds can alter or even prevent invasion of cancer cells. The ultimate goal of this dissertation is to contribute a body of work to the compendium of studies that may one day stop death from metastatic cancer.

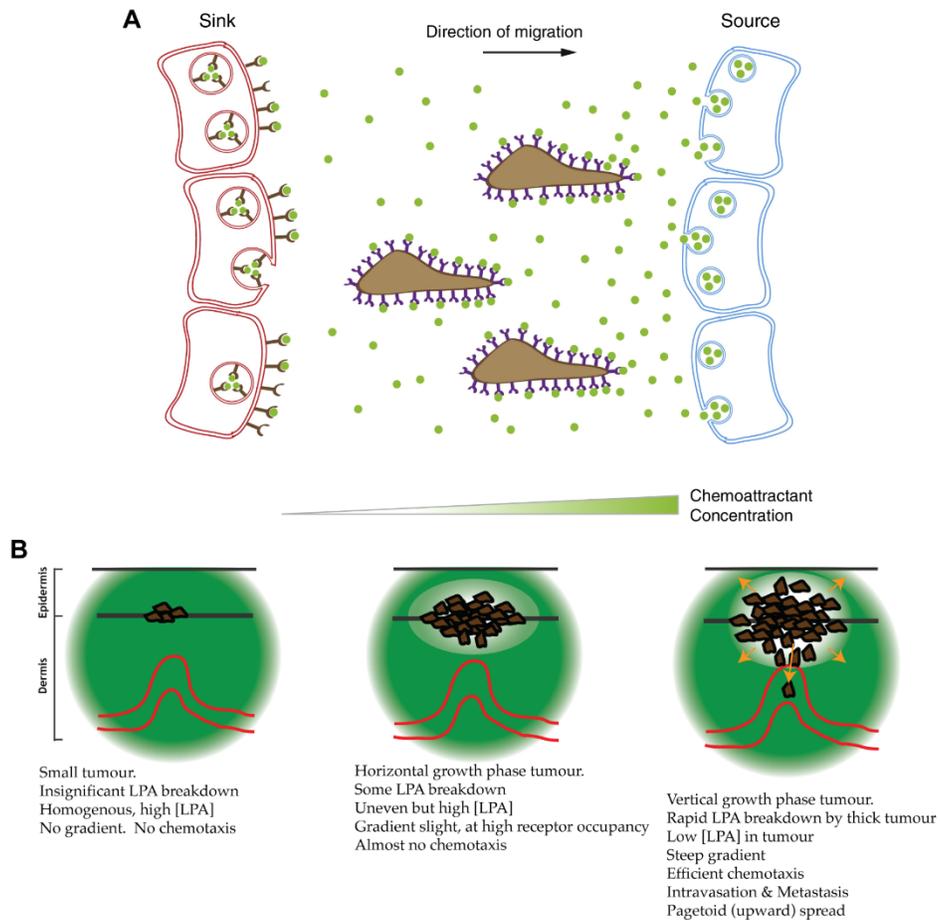


Figure 1.1 Source and sink model of chemotaxis in cell migration and in melanoma. Images adapted from Cai et al., 2014 [114] and Muinonen-Martin et al., 2014 [117]. (A) A cell migrates along a gradient of the green chemoattractant created by secretion from the source tissue and depletion within an opposite sink tissue, thus forming a gradient. (B) Depiction of growth of a melanoma tumor and subsequent degradation of LPA (green). As the tumor grows, LPA becomes depleted in areas of high cell density, generating a gradient of LPA concentration along which cells may migrate.

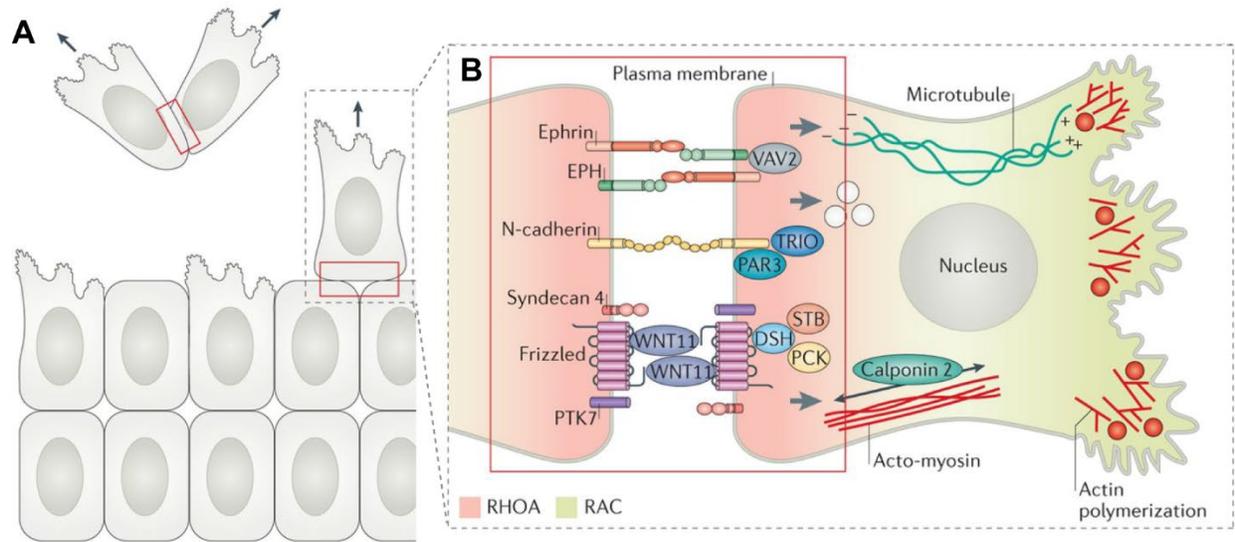


Figure 1.2 The process of contact inhibition of locomotion. Image adapted from Mayor and Etienne-Manneville 2016 [73]. (A) Cells changing direction (arrows) as a response to cell-cell contact (red boxes) (B) Interaction of molecules at site of cell-cell contact (red box) leading to RHOA activation (light red) and RAC inhibition (light green). At one the site of cell-cell contact, microtubules (teal threads) and focal adhesions disassemble (white circles), and protrusions collapse as actomyosin (red threads) contracts. Opposite to the site of cell-cell contact, protrusions form as microtubules polymerize and focal adhesions stabilize (red circles).

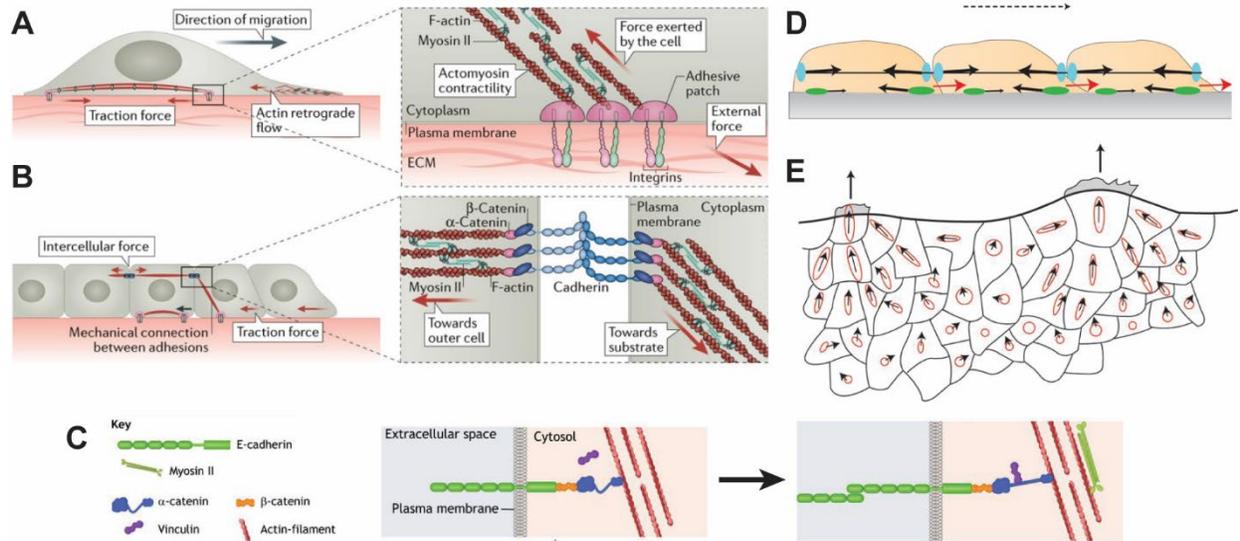


Figure 1.3 Force generation, mechanosensation, and distribution during collective invasion.

Image and text adapted from Ladoux and Mege 2017 [289], Chen et al. 2018 [290], and Gov 2011 [148]. (A) Model of a cell exerting traction forces through actomyosin contractility on the ECM at the integrin-based adhesions (inset) both at the front and at the back. (B) Molecular mechano-coupling of collectively invading cells and intercellular force transmission through cell-cell junctions to actomyosin, mediated by mechanosensitive cadherin complexes (inset). (C) Binding of vinculin (purple) to binding site of unfolded α -catenin (blue), exposed by force-dependent unfolding. (D) Schematic of the forces (arrows) produced in the two-dimensional cell layer during plithotaxis. Overall direction of forces and motion is depicted by the dashed black arrow. The full black arrows indicate the inwards pull of the actin-myosin skeleton at cell-cell adhesions (blue) and cell-substrate adhesions (green), which generate traction forces at the ECM. Additionally, each cell extends a small lamellipodia (cryptic lamellipodia) beneath a neighboring cell, exerting a pushing force (red arrows). (E) Top-view of the edge of the cell culture closing a wound. Where the red oval, marking local stress, is less circular, forces are more polarized. Cell motion (black arrows) is highly correlated with the direction and degree of maximal principle stress. The two cells at the edge are initiating pulling forces through develop large lamellipodia (gray).



Figure 1.4 Schematic of protein domains of KDM5B. Members of the KDM5 family are additionally characterized by a split catalytic Jumonji domain with a DNA-binding ARID, AT-rich interacting domain, and PHD finger (PHD1) intervening between the N-terminal (JmjN) and C-terminal (JmjC), as well as a C_5HC_2 zinc finger and two additional PHD fingers (PHD2 and PHD3). Besides the two halves of the Jumonji domain, the C_5HC_2 zinc finger is necessary for catalytic activity of the protein. The PHD domains aid in substrate recognition. In KDM5B and to some extent in KDM5A, the PHD1 domain recognizes H3K4me₀, which may stabilize target gene repression by preventing remethylation or promoting removal of adjacent H3K4me_{2/3}, but in KDM5C, the PHD1 domain binds H3K9me_{2/3}. The PHD3 domains of KDM5A and KDM5B bind the substrate H3K4me₃.

Chapter 2

Genetic heterogeneity within collective invasion packs drives leader and follower cell phenotypes

This chapter is adapted from a manuscript under consideration in the *Journal of Cell Science* and published by Elizabeth L. Zoeller*, Brian Pedro*, Jessica Konen, Bhakti Dwivedi, Manali Rupji, Niveda Sundararaman, Lei Wang, John R. Horton, Xiaodong Cheng, Chaojie Zhong, Benjamin G. Barwick, Elisabeth D. Martinez, Matthew P. Torres, Jeanne Kowalski, Adam I. Marcus, Paula M. Vertino

*Contributed equally to this work

This work is under revision at *Journal of Cell Science*.

Contributions: The work presented here is the result of a co-authorship with B.P. Each first author contributed equally to the work. E.L.Z. performed experiments in Figures 2.1.D,E regarding KDM5B, 2.4.C, 2.4.D protein expression blots, 2.5.A-E, S2.4, S2.5, S2.6, and S2.7 and graphed and diagramed Figures 2.4.A, 2.6 A-C, and S2.8.

2.1 Abstract

Collective invasion, the coordinated movement of cohesive packs of cells, has become recognized as a major mode of metastasis for solid tumors. These packs are phenotypically heterogeneous and include specialized cells that lead the invasive pack and others that follow behind. To better understand how these unique cell types cooperate to facilitate collective invasion, we analyzed transcriptomic sequence variation between leader and follower populations isolated from the H1299 non-small cell lung cancer cell line using an image-guided selection technique. We now identify fourteen expressed mutations that are selectively enriched in leader or follower cells, suggesting a novel link between genomic and phenotypic heterogeneity within a collectively invading tumor cell population. Functional characterization of two phenotype-specific candidate mutations shows that ARP3 enhances collective invasion by promoting the leader cell phenotype, and that relaxed KDM5B-enforced leader behavior may be necessary to support cooperative subpopulations in collective chains. These results demonstrate an important role for distinct genetic variants in establishing leader and follower phenotypes and highlight the necessity of maintaining a capacity for phenotypic plasticity during collective cancer invasion.

2.2 Introduction

Metastatic disease is the cause of 90% of deaths among cancer patients [2]. In non-small cell lung cancer (NSCLC), which comprises 80-85% of all lung cancer diagnoses, metastases are commonly observed in the bones, lungs, brain, liver, and adrenal glands. Patients presenting with metastatic disease have significantly worse prognoses than those with early-stage disease; for example, the 5-year survival rate for stage I NSCLC is 55%, while stage IV disease (in which distant metastases are present) carries a mere 4% five-year survival rate [291]. Successful colonization of distant sites requires that cells from the primary tumor gain the capacity to invade through the surrounding basement

membrane, travel through the bloodstream or lymphatic system, and ultimately expand to establish colonies at the metastatic site [42, 292].

Cells migrate through the microenvironment via multiple mechanisms. A classic example is the epithelial-to-mesenchymal transition, where cells lose expression of epithelial features such as E-cadherin and gain expression of mesenchymal proteins including vimentin and N-cadherin. This shift is thought to promote cellular detachment and enable cancer cells to undergo single-cell invasion. In contrast, collective cell migration refers to the coordinated movement of a group of cohesive cells [75]. This phenomenon is well-described in embryonic development and wound healing, and histological evidence from primary patient tumor samples [293-297], mouse models of metastasis [298, 299], and 3-D cultures [288, 294, 300-303] suggest that cells from solid tumors often migrate and invade in cohesive packs as well. These collective invasion packs and streams vary in width, shape, and cell number, as well as in the mechanisms guiding their movement [93, 304-309].

Understanding the mechanisms that underlie the outgrowth of metastatic clones is further complicated by the heterogeneous mix of cell populations within each tumor. This intratumor heterogeneity arises from cell-to-cell variation in the genetic background each expanding to create a unique subclonal population [4]. Superimposed upon this subclonal genomic heterogeneity is the potential for epigenetic heterogeneity reflected in variations in gene expression even among genetically identical cells. Recently, multiregional sequencing of primary lung tumors characterized the intratumor heterogeneity of NSCLC and showed that upwards of 24% to 30% of mutations went undetected in at least one sampling region, demonstrating that almost a third of mutations are occurring in spatially distinct subclonal populations and may have been missed in broad scale data based on single sampling, such as those analyzed as part of the TCGA project [6, 207]. These unique subpopulations may be endowed with properties that enhance attributes beneficial to tumor cells, such as resistance to drug therapy or the ability to invade and metastasize [40, 310-312]. For instance, the clonal profile of

metastatic disease often does not reflect the profile of the primary tumor, but instead includes one or just a few subpopulations from the primary site [88, 312-314].

One example of phenotypic heterogeneity associated with invasive behavior includes rare, specialized leader cells that lead collective invasive packs, and follower cells that adhere to and follow behind the leaders, both of which cooperate to achieve collective invasion [89, 288, 293, 294, 309]. We developed a novel platform (**S**patiotemporal **G**enomic and Cellular **A**nalysis, or **SaGA**) to isolate specific leader and follower cell populations from collectively invading NSCLC cells [288]. Characterization of these cell types revealed that isolated follower cells are highly proliferative but poorly invasive, while isolated leader cells are highly invasive, but poorly proliferative [288]. These cellular sub-types cooperate through an atypical angiogenic signaling pathway that is dependent upon VEGF. Previous data suggest a symbiotic relationship, in which both leader and follower cells are necessary for collective invasion to proceed successfully; however, key questions remain as to what drives the biology and emergence of leader and follower cells from a tumor cell population.

In this study, we aimed to elucidate the role of genetic heterogeneity on collective invasion. We analyzed invading leader and follower populations arising from a common H1299 parental NSCLC cell line grown as 3-D spheroids. Strikingly, this revealed mutational landscapes that differ significantly between leader and follower cells, including several expressed mutations that were found exclusively in one cell type or the other. To our knowledge this is the first identification of leader- and follower-specific gene mutations within the same collectively invading tumor cell population.

2.3 Results

Leader and follower cell populations contain distinct mutational profiles.

We previously developed the SaGA technique for isolation of leader and follower cells from collectively invading packs of human NSCLC cancer cells [288]. Briefly, cells expressing Dendra2

green-to-red photoconvertible fluorescent protein are formed into multicellular spheroids, embedded in Matrigel and allowed to invade for 24 hours. Leader or follower cells are then selected based upon physical positioning within invasive chains, optically highlighted by photoconversion using a 405nm laser, and separated by fluorescence-activated cell sorting (FACS) (Fig. 2.1A). Using this approach, we isolated three follower populations and two leader populations from H1299 parental cells. Following expansion of each population in 2-D culture, RNA-seq was performed in triplicate, using three separate passages of parental H1299 cells, the three separately isolated populations of follower cells, and the two separately isolated populations of leader cells (including two passages of one of the leader populations). Sequence variants were determined for each population (leader, follower, parental) independently by mapping the RNA-seq profiles to human reference genome Hg19 (GRCh37), resulting in a total of 6240 filtered variants combined in the three populations (see Methods for details). Notably, when comparing variant allele frequencies (VAF) via pairwise t-test analysis, a number of variants were disproportionately present in the leader versus follower populations. We therefore further filtered for those variants that exhibited >20% VAF in either leaders or the followers and <1% in the other (VAF student's t-test p-value <0.01 between leaders and followers). For the purposes of this study, we further excluded those located in 5' or 3' UTRs and known SNPs. Application of these criteria identified fourteen missense mutations – six leader-specific and eight follower-specific (Table 1). This represents the first identification of leader- and follower-specific mutations within the collective invasion pack.

Given the cell type specificity, we hypothesized that these mutations could be key contributors to the emergence of leader vs. follower phenotypes from the parental population. To test this, we chose two candidate mutations for further study – one leader-enriched, *ACTR3* chr2:114699797 A to G, which results in a mutation in *ARP3* at lysine 240 (*ARP3* K240R), and one follower-enriched, *KDM5B* chr1:202715414 A to C, resulting in a mutation at L685 (*KDM5B* L685W) (Fig. 2.1B). We

first confirmed these mutations by Sanger sequencing of genomic DNA and cDNA from the parental H1299 population as well as the isolated leader and follower populations (Fig. S2.1). Both variants were detectable at subclonal levels in genomic DNA, indicating that they were unlikely to have arisen *de novo* during transcription, but represented a subpopulation of genomic alleles in the parental population. Moreover, the selectivity for the leader or follower population observed at the RNA levels was preserved at the genomic DNA level (Fig. S2.1; Fig. 2.1C). Analysis of ARP3 and KDM5B expression in H1299 parental, leader and follower cells showed that ARP3 mRNA and protein levels were comparable between the populations (Fig. 2.1D,E). KDM5B mRNA and protein were significantly reduced in follower cells relative to leaders (Fig. 2.1D,E). Despite the reduced overall levels, the KDM5B mRNA in follower cells retained the ~2:1 ratio of wild-type vs. mutant observed at the genomic DNA level (Fig. 2.1C) suggesting that the follower-enriched KDM5B L685W mutation is expressed. Similarly, while there was some variation in the frequency of both mutations in the parental population between methods and DNA/RNA samples isolated at different times, there was little variation in the allelic balance in leader and follower populations, which maintained a consistent frequency of their respective mutations at both DNA and RNA level, suggesting that there is no allelic bias in the expression of the mutant version in either case. Thus, our selection of leader and follower cells based on phenotypic criteria also selected for subpopulations with distinct allelic balance of expressed mutations.

Predicted functional impact of the leader-enriched ARP3 K240R mutation.

We next sought to characterize the potential impact of the observed mutations. We started with the leader-enriched *ACTR3* mutation, which results in a K to R shift in ARP3 (K240R). ARP3 is a key component of the Arp2/3 complex that helps facilitate cellular migration by promoting lamellipodia protrusion [315]. Overexpression of ARP3 has been correlated with invasion, metastasis,

and poor survival in multiple cancer types, including gastric, colorectal, liver, and gallbladder [316-319]. Furthermore, multiple mass spectrometry studies indicate ARP3 K240 as a post-translational modification (PTM) site, with evidence of both ubiquitylation and acetylation [320-322] (Fig. 2.2A, inset). To evaluate the functional impact of the K240R mutation on the ARP3 protein, we used SAPH-ire (Structural Analysis of PTM Hotspots) [323], which predicts the functional potential of PTMs in protein families that have existing experimental and 3D structure data. SAPH-ire calculates a Function Probability Score (FPS) using a neural network model trained with an array of protein sequence and PTM-specific features extracted from PTMs with established functional impact. Consistent with these data, K240 had among the highest FPS values of all known modified residues within the ARP3 protein family and was among the top 90% of PTMs with well-established functional significance (i.e. 4 or more publications) across all protein families (Fig. 2.2A). SAPH-ire also revealed six experimental ubiquitylation sites in the ARP3 protein family between residues alignment positions 1298 – 1315, four of which correspond to ubiquitylation of ARP3 specifically (K240, K244, K251, and K254). Of these, K240 had the highest mean solvent accessible surface area (SASA) and was also proximal to a protein interaction interface (Fig. 2.2B). The high solvent accessibility, proximity to a protein-protein interface, and predicted functional impact of the ARP3 K240 site suggest that this mutation could indeed be altering the activity of ARP3 in leader cells.

ARP3 K240R promotes invasion and leader cell behavior.

We sought to determine the functional impact of the ARP3 K240R by introducing this mutation into follower cells and quantifying changes in 3-D invasion. Given the relatively low VAF (23.4%) of mutant *ACTR3* in leader cell DNA/RNA, we sought to replicate the leader cell *ACTR3* allelic balance by employing a rescue approach. ARP3 levels were first stably knocked down using two separate short hairpin RNAs (Fig. S2A,B), including one (shACTR3 #2) targeted to the 5' UTR of

endogenous *ACTR3*. Knockdown of ARP3 significantly reduced 3-D invasion in H1299 parental, leader and follower cells compared to pLKO.1 controls (Fig. S2.2C,D). Using a sulforhodamine B (SRB) assay to measure cell growth, we found little difference upon ARP3 knockdown until 96 hours in the leader and follower populations, and 120 hours in the parental population, when proliferation was decreased (Fig. S2.2E).

To test the functional consequences of ARP3 mutation, follower cells expressing shACTR3 #2 were then ‘rescued’ by stably expressing empty vector, mCherry-tagged wild-type ARP3, or mCherry-tagged ARP3 K240R, under the control of the moderate activity UBC promoter [324]. Higher expression was achieved for ARP3 K240R compared to wild-type ARP3 (Fig. 2.2C, upper bands), suggesting that ARP3 K240R may be more stable than the wild-type protein. When grown as a spheroid, embedded in a Matrigel matrix, and allowed to invade for 24 hours (Fig. 2.2D), ARP3 knockdown follower cells reconstituted with ARP3 K240R, and to a lesser extent those reconstituted with wild-type ARP3, exhibited significantly higher invasive area compared with unmodified followers or those reconstituted with empty vector (Fig. 2.2E). This indicates that ARP3 expression, and especially ARP3 K240R expression, can increase invasive capacity even in normally poorly-invasive follower cells.

Next, we sought to examine whether ARP3 K240R could specifically promote leader cell behavior (i.e. facilitate collective invasion and travel at the front of invasive chains) in a heterogeneous population. We created 3-D spheroids comprised of 90% unmodified H1299 followers and 10% ARP3 depleted followers rescued with either empty vector, wild-type ARP3, or ARP3 K240R (Fig. 2.3A). After 24 hours embedded in Matrigel, we observed a significant increase in invasive area and average number of chains per spheroid, and decreased circularity (indicating more chain-like and less sheet-like invasion) in the mixed spheroids containing 10% ARP3 K240R-rescued followers, as compared with the other three conditions (Fig. 2.3B,C). To determine whether the ARP3 K240R-

rescued followers were in fact leading these invasive chains, we used confocal fluorescence imaging to quantify the fraction of chains that exhibited mCherry-ARP3 K240R-rescued followers in the leader position. In spheroids containing 10% wild-type ARP3-rescued followers, we found rescued cells in the leader position in 53.8% of those chains (95% confidence interval 29.1% to 76.8%; Fig. 2.3D). By contrast, in spheroids containing 10% ARP3 K240R-rescued followers, rescue cells were found in the leader position in 87.2% of chains (95% confidence interval 78.0% to 92.9%; Fig. 2.3D). Thus, while both wild-type ARP3-rescued and ARP3 K240R-rescued cells led chains at a higher frequency than expected by random chance, the ARP3 K240R-rescued cells were more efficient in this regard. The specificity of the effect was further confirmed by mixing experiments with empty vector-rescued follower cells, which led only 17.7% of chains (95% confidence interval 6.19% to 41.0%; Fig. S2.3), suggesting that the enhanced leader ability of the ARP3 K240R cells was not simply due to the different cell types segregating within the spheroid. Together these data indicate that ARP3 K240R confers key leader-like behaviors onto follower cells, including increased invasive capacity, increased numbers of invasive chains, and a greater ability to lead those chains.

As noted above, the K240R-reconstituted ARP3 knockdown cells express higher levels of ARP3 than those reconstituted with the wildtype ARP3. To distinguish the impact of ARP3 dosage from that of ARP3 K240R mutation, we re-created the same rescue cell lines using a CMV promoter (Fig. 2.3E) and repeated the above experiments creating mixed spheroids with 10% rescued cells (Fig. 2.3F). At this higher protein expression level, there was no significant difference in invasive area between the groups (Fig. 2.3G). Additionally, while chain number increased and circularity decreased in spheroids containing either wild-type ARP3-rescued or ARP3 K240R-rescued followers (Fig. 2.3G), there was no significant difference between the two, suggesting that the functional difference between ARP3 K240R and wild-type ARP3 is mitigated at higher expression levels. Furthermore, both wild-type ARP3-rescued followers (81.5% of chains led; 95% confidence interval 63.3% to 91.8%) and

ARP3 K240R-rescued followers (84.2% of chains led; 95% confidence interval 62.4% to 94.5%) promoted leader cell behavior when mixed with 90% unmodified followers (Fig. 2.3H). Based upon these findings, it appears that increased dosage of ARP3 is sufficient to promote leader-like behavior, and that the selective ability of ARP3 K240R to lead invasive chains and drive collective invasion at lower expression levels may arise, in part, from increased effective dose of ARP3 protein.

Functional impact of the follower-enriched KDM5B L685W mutation.

Subsequently, we sought to assess the effects of a follower-specific mutation on collective invasion. As noted above, the KDM5B L685W mutation was selectively expressed in the parental and follower populations but was excluded from the leader population (Fig.1B). KDM5B is a lysine demethylase that catalyzes the removal of di- and trimethylation from histone H3 methylated at lysine 4 (H3K4me2, H3K4me3). As an epigenetic regulator, KDM5B is uniquely situated among the identified mutations to influence multiple pathways directly involved in guiding invasive behavior (Table 1) [57, 325, 326]. Indeed, KDM5B has been implicated in the regulation of invasive and migratory behavior across many cancer types, including hepatocellular, gastric, breast, lung, and prostate [261, 268, 269, 271, 327-329]. The point mutation at amino acid 685 observed in followers results in a leucine to tryptophan (L685W) substitution and lies in close proximity to the zinc finger domain, a region necessary for demethylase activity *in vitro* and in cells [204, 213] (Fig. 2.4A). Furthermore, structural modelling of the L685W substitution showed that residue 685 sits in a hydrophobic pocket within KDM5B, near the binding site of the N-terminal tail of histone H3 (Fig. 2.4B). As tryptophan has a much bulkier sidechain than leucine, the introduction of this large, aromatic sidechain has the potential to alter or disrupt the interaction of KDM5B with histone H3 (Fig. 2.4B).

Overall, the follower cells expressed less KDM5B protein than leader cells or the parental H1299 population (Figs. 2.1E, 2.4C), and, as noted above, at least a fraction of that (~30%) would be

expected to harbor the L685W mutation. KDM5A was also slightly decreased in follower cells, but expression of KDM5C was equal across populations (Fig. S2.4). Comparison of the global levels of the KDM5 substrates showed that while H3K4me2 and me3 were similar across populations, follower cells had reduced global levels of H3K4 monomethylation (H3K4me1), the end-product of KDM5B catalytic activity, suggesting reduced cellular H3K4 demethylase activity in follower cells (Fig. 2.4C). To test this directly, we generated parental H1299 cells stably overexpressing HA-tagged wild-type KDM5B or KDM5B L685W and performed *in vitro* demethylation assays on cell lysates (Fig. 2.4D). Whereas overexpression of wild-type KDM5B trended toward increased cellular H3K4 demethylation activity by approximately 2-fold, overexpression of KDM5B L685W had no impact on endogenous H3K4 demethylation activity despite similar levels of wild-type and mutant expression (Fig. 2.4D). Taken together these data indicate that the KDM5B L685W mutation seen in the followers is associated with reduced cellular KDM5 activity.

KDM5B L685W modulates collective invasion by tipping the balance toward follower cell behavior.

The above data suggests that KDM5B L685W may confer a partial loss of function/decreased activity on the follower cells, and further that this mutation is selected against in the leader population (addressed in more detail below). We therefore sought to determine how the expression of the L685W might behave in the context of an otherwise wild-type KDM5B background (to mimic the mixed genotype observed in followers) by expressing KDM5B L685W in leader cells and determining the impact on collective invasion in 3-D spheroid culture (Fig. 2.5A,B). Importantly, overexpression of wild-type KDM5B or KDM5B L685W did not affect cell proliferation leading us to conclude that any changes in spheroid phenotypes is attributable to alterations in invasive behavior, not expansion (Fig. S2.5). Interestingly, overexpression of wild-type KDM5B in leader cells suppressed their propensity

for chain-like invasion, as indicated by an increase in circularity and a decrease in the number of chains relative to cells expressing empty vector (Fig. 2.5B,C). However, invasive area was similar between leaders expressing empty-vector and wild-type KDM5B, suggesting that the wild-type KDM5B expressing leaders retain invasive capabilities but lose some of the collective (cooperative) behavior manifested as chain-like invasion. By contrast, overexpression of the KDM5B L685W mutant enhanced slightly the chain-like invasion, with spheroid behavior resembling the high chain number and low circularity observed in empty vector expressing leader cells (Fig. 2.5B,C). Parental-derived spheroids were similar in nature to leader spheroids in that spheroids expressing KDM5B L685W produced more chains than those expressing wild-type KDM5B (Fig. S2.6). However, empty vector expressing parental spheroids, unlike unmodified parental spheroids, produced negligible numbers of chains making comparisons to wild-type KDM5B impractical (Fig. S2.6).

To reconcile these findings, we considered the possibility that heterogeneity in KDM5B expression/activity may be required for cells in a population to take on the cooperative roles necessary to form an invasive chain. Even in spheroids derived from a purified leader population, such as those shown above, there must be at least some cells that take the trailing position, or evolve follower-like behavior, even in the absence of new genetic events. One interpretation of the above data is that overexpression of wild-type KDM5B may further enforce the leader cell phenotype and prevent the emergence of cells with follower behavior, thus suppressing chain-like co-operative behavior; by pushing the balance even further towards high KDM5B activity, the reduced population plasticity might decrease the probability that cells can take on the follower role. By contrast, expression of KDM5B L685W might allow for greater variation in KDM5B activity, promoting heterogeneity and the emergence of cells with follower behavior, further enhancing collective behavior. If this is the case, then one might predict that ectopic expression of wild-type KDM5B or KDM5B L685W would have relatively little impact on follower cells, which on their own do not exhibit significant cooperative

behavior. Indeed, when expressed in follower cells, the further expression of KDM5B or KDM5B L685W had a negligible impact on circularity or the chain number of follower-only spheroids (Fig. 2.5B,C). These data also indicate that increasing the concentration of wild-type KDM5B is in itself not sufficient to impose leader behavior on follower cells.

The benchmark test for leader cell activity is to determine how such cells behave when mixed with followers in 3-D invasion assays. As noted above for ARP3, leader cells can have a strong effect on chain formation, even when they make up a very small proportion of the total cells in a spheroid. We therefore tested directly how expression of KDM5B L685W affects the ability to lead in a heterogenous population by creating mixed spheroids comprised of predominantly (75%) unmodified follower cells into which leader cells overexpressing wild-type KDM5B or KDM5B L685W (25%) had been added and monitoring chain-like invasion (Fig. 2.5D,E). We found again that even in the context of unmodified followers, expression of wild-type KDM5B suppressed the chain-like behavior relative to control (vector only) leaders. In contrast, expression of KDM5B L685W not only retained but even somewhat enhanced the ability to lead invasive chains compared to those expressing either empty-vector or wild-type KDM5B (Fig. 2.5D,E). These data suggest that overexpression of KDM5B L685W mutant sustains chain-like cooperative behavior. Taken together, these data support the idea that wild-type KDM5B supports the leader phenotype in part by suppressing emergence of follower-like behavior, while expression of the L685W mutant supports, and even somewhat promotes chain-like invasion.

To further test the impact of KDM5B L685W population heterogeneity on collective invasion, we mixed leader lines expressing empty vector, wild-type KDM5B, and KDM5B L685W with unmodified leaders expressing mCherry at different proportions and monitored chain-like invasion in 3-D spheroids (Fig. S2.7). As both lines already express Dendra2, we quantified the fraction of chains led by red/magenta positive cells (mCherry, unmodified leaders) or green-only cells (KDM5B

modified lines) (Fig. S2.7A). We found that overexpression of wild-type KDM5B again suppressed chain formation in spheroids (Fig. S2.7D, 100% KDM5B WT) relative to empty vector, which was recovered proportionally as the fraction of unmodified leaders in the spheroid increased (Fig. S2.7D). Despite this suppression in total chains, the fraction that were led by a green-only, wild-type KDM5B overexpressing leader was roughly proportional to its abundance in the mixed spheroid (compare Fig. S2.7 panels C and D, 90% vs. 10% wild-type KDM5B). By contrast, leaders expressing L685W showed no suppression of chain-like behavior relative to vector-only cells (Fig S2.7D, 100% KDM5B L685W), and exhibited a similar frequency of green-only vs. magenta (unmodified leaders) in the lead position regardless of their relative proportions. These data support the idea that overexpression of wild-type KDM5B may suppress the emergence of cooperative behavior necessary to form a chain, such cells nevertheless retain the capacity to lead in a mixed population.

SNP analysis suggests distinct leader and follower cell lineages.

If high expression of wild-type KDM5B enforces the leader phenotype and/or suppresses the emergence of an alternate phenotype (e.g. followers), and this has an impact on collective behavior, then one might predict that there would be a selection against the expression of KDM5B L685W in leader cells and selection for KDM5B L685W in follower cells independently captured from H1299 spheroids relative to the parental population. Fortuitously, a common SNP (rs1141108, chr1:202715284, $G>A$) is located within the same exon as the KDM5B mutation, and the H1299 cell line was determined to be heterozygous for this SNP in our initial sequence analysis. Targeted deep resequencing of an amplicon including the region containing both the mutation (chr1:202715414) and the SNP (chr1:202715284) in genomic DNA thus enabled us to trace the relationship between allelic balance and the frequency of the KDM5B L685W mutation in each cell population (average depth=348,327 reads).

We first found that approximately one third of alleles in each population carried the *A* variant and two thirds carry the *G* variant at rs1141108, leading us to conclude that there are 3 copies of *KDM5B* and/or chromosome 1 in our strain of H1299 cells and that this ploidy is maintained across the three populations. The *ACTR3* mutation that gives rise to ARP3 K240R on chromosome 2 also occurs in approximately one third of reads from genomic DNA in leader cells (Fig. 2.1B). Focal copy number alterations in the genomic regions surrounding *ACTR3* and *KDM5B* have not been detected in H1299 cells by SNP copy number analyses (COSMIC Cell Lines Project: https://cancer.sanger.ac.uk/cell_lines). These data suggest that our strain of H1299 cells are functionally triploid. We further ascertained that the *KDM5B* variant (chr1:202715414: *A>C*) giving rise to the L685W mutation resides exclusively in *cis* with the SNP rs1141108 *G* allele (99.7% concordance across 267,414 total reads containing the mutation). The analysis further confirmed the proportions of *KDM5B* wild-type vs. mutant alleles shown in Fig. 2.1, including the more variable mutation frequency observed among different isolates of the parental population, the complete absence of the *KDM5B* L685W mutation (<0.00001%) in the leader cells and the very consistent ~30% in the follower cells (Fig. 2.6A). Additionally, the relative proportions of the wild-type and variant expressed at the mRNA level were largely reflective of the proportion at the genomic DNA level across all three populations, again suggesting that there was no allelic bias in expression of the wild-type versus mutant forms of *KDM5B* (Fig. 2.6A, S2.8).

The finding that the *KDM5B* L685W mutation was exclusive to the *G* allele while the *A* allele was exclusively wild-type allowed us to trace the relative proportions of the three genotypes (WT/*A*, WT/*G*, mutant/*G*) across the parental, leader, and follower populations. These data showed that whereas 4-12% of the parental population carries the *KDM5B* mutant/*G* allele, this allele was strongly selected against in the leader population, which exhibited essentially none of the *KDM5B* mutant/*G* alleles (Fig. 2.6B). Moreover, relative to the parental population, the *KDM5B* mutant/*G* allele was

nearly 2-fold enriched in the isolated followers (30%) ($p=0.008$; ANOVA plus Tukey's post-hoc correction) and approached the expected frequency (if each allele is independently sorted. Taken together these data suggest that whereas the parental population varies in the fraction of cells containing the mutant/SNP G allele, growth in 3-D culture and the isolation of leader cells selectively enrich for cells expressing only wild-type KDM5B (and against KDM5B L685W), and the isolation of followers selects for a population in which nearly every cell contains (and expresses) one copy of the mutant KDM5B L685W allele.

2.4 Discussion

The greatest threat to cancer patient mortality is the metastatic spread of tumor cells from the primary site [2]. Collective migration and invasion are major contributors to the dispersion of metastatic cancer cells [73, 75, 89, 330]. Collective invasion is typified by the coordinated movement of a group of cohesive cells, often including multiple heterogeneous cell populations with specialized functions. One well-studied example of collective invasion is that of chain-like invasion, in which specialized leader cells lead groups of cells termed follower cells, out of the tumor [73, 89, 288], with both populations playing important roles in the process of invasion. Until now, studies of the distinct populations within invasive chains have been limited by the inability to separate these populations. Development of the SaGA technique [288], enabled us to independently analyze leader and follower cells with different phenotypes to gain insight into population dynamics and the emergence of populations that differ in cell behavior. Our results identify a set of expressed mutations that define leader and follower cells, representing, to our knowledge, the first known instance of distinct mutations as contributors to the leader/follower phenotypes within collectively invading packs.

We confirmed the importance of the leader cell-enriched mutation ARP3 K240R to the invasive leader cell phenotype by introducing it into a population of non-invasive H1299 follower

cells. Both in pure spheroids and when mixed with 90% unmodified followers, ARP3 K240R-expressing followers displayed increased ability to invade and lead collective chains at both lower and higher protein expression levels. Rescue with wild-type ARP3 also conferred leader cell behavior, but only when expressed at supra-physiologic levels. One potential explanation is that ARP3 K240R increases the effective dosage of ARP3 protein, essentially recapitulating ARP3 overexpression even at low expression. Indeed, ARP3 K240R accumulated to higher levels than wild-type ARP3 when exogenously expressed from the same vector. The K240R mutation might interfere with ubiquitylation at K240, resulting in either decreased ARP3 turnover, or enhanced ARP3 activity. Ubiquitylation at K240 has previously been observed by mass spectroscopy in multiple human cell lines as well as mouse tissue, and K240R was predicted by SAPH-ire to have a high likelihood of functional consequence. ARP3 is a key subunit of the Arp2/3 complex that regulates intracellular actin dynamics in a number of processes, including lamellipodia protrusion during cell motility [315]. Indeed, we observed significantly reduced invasion in parental, leader, and follower cells upon ARP3 knockdown, supporting its importance for cell migration and invasion. Overexpression of Arp2/3 complex subunits including ARP2, ARP3, ARPC2, and ARPC5 has been shown to promote invasion in multiple cancer types including lung, colorectal, glioblastoma, and others [316-319, 331-333]. Our results now further indicate a role for ARP3 as contributing to tumor collective invasion by promoting the leader cell phenotype.

We also investigated the role of the follower cell enriched mutation, KDM5B L685W. Surprisingly, overexpression of wild-type KDM5B in leader cells suppressed invasive chains as well as the ability to lead invasive chains when mixed with followers. By contrast, overexpression of KDM5B L685W restored chain-like invasion. Interestingly, these data are seemingly at odds with published data showing that overexpression of wild-type KDM5B in cancer cells increases their migration and invasion in 2-D assays [261, 268, 271, 327-329], including in H1299 cells [269], and may be a reflection

of the differences in the assessment of invasive behavior in 2-D, which focuses on the behavior of single cells, and that of cells grown as spheroids and undergoing collective invasion. Collective invasion requires cooperation between cell phenotypes, and the ability of cells in the population to take on distinct roles upon suspension in 3-D. One interpretation of our data is that wild-type KDM5B may enforce the leader phenotype by tipping the balance away from cooperative behavior and that KDM5B L685W is somehow deficient in this regard. KDM5B acts as a transcriptional repressor, and its sustained expression in leader cells may act to repress the follower 'program' thus preventing the emergence of the alternate (follower) phenotype. Consistent with this idea, isolation of cells that take the lead position in 3-D invasive streams selects for those that lack the mutant allele, while the followers can express up to ~30% mutant KDM5B, exhibit reduced KDM5B expression, and have a skewed intracellular H3K4me2/3 to H3K4me1 ratio relative to leaders consistent with reduced/compromised intracellular KDM5 activity. Interestingly, KDM5B was recently identified as a key suppressor of transcriptomic heterogeneity. Whereas KDM5B overexpression enforces phenotypic stability of luminal breast cancer cells, its depletion/inhibition was shown to promote transcriptional plasticity and to facilitate breast cancer cells to overcome therapeutic resistance [279]. Reduced KDM5B expression/activity and the ensuing transcriptomic heterogeneity may also be a necessary step for the emergence of the specialized leader and follower phenotypes that define collectively invading chains. In this regard, it is interesting to note that leader cell cultures tend to be phenotypically stable over time and retain their highly invasive phenotype, whereas follower cultures exhibit a greater degree of phenotypic plasticity, with a tendency to spin off new invasive chains (and hence take on leader characteristics) over time [288], suggesting that at least some components of collective behavior are epigenetically driven.

Previous studies have suggested that cooperation between phenotypically and genetically distinct subpopulations derived from a common tumor cell population are necessary to promote

invasion and/or metastasis *in vivo*. For example, Calbo *et al.* [334] found that neuroendocrine and non-neuroendocrine-like clones derived from a common primary murine SCLC tumor were incapable of completing metastasis on their own, but once re-mixed became highly metastatic. Shared copy number variations confirmed a common clonal origin, but some genetic alterations (e.g. amplification of *Myc11*) were only observed in the neuroendocrine-like population. Likewise, Mateo *et al.* [335] found that co-culture of phenotypically distinct subpopulations isolated from the PC-3 prostate cancer cell line promoted invasion and metastasis over either population in isolation. Other studies have shown that the post-hoc admixture of cells engineered to undergo EMT or to express activated HRAS with the non-engineered control population was necessary to fully complete metastasis [336] or to promote collective motility of breast epithelial cells [337]. The necessity for heterogeneity in the transcriptional program, even when genetically driven, is an emerging theme, and is consistent with our observations here.

Emerging research regarding how followers cooperate with leaders to promote their invasive capabilities center on concepts such as contact inhibition of locomotion (CIL) and movement along a chemical gradient [73, 338]. CIL occurs when a migrating cell contacts another cell and begins forming protrusions opposite the site of contact to move in the opposite direction [118, 339, 340]. In this case, follower cells contact leader cells, forcing them to polarize and move in a forward direction. Follower cells also regulate signaling to leader cells, often through creation of chemical gradients that motivate leader cells to move in a particular direction [73, 115-117, 341, 342]. Indeed, we previously discovered that follower and leader cells engage in a symbiotic relationship, in which follower cells promote the survival and proliferation of leader cells, which in turn secrete VEGFA to promote the motility of follower cells [288].

Our sequencing data and the ability to discern the clonality of the distinct leader and follower subpopulations relative to that of the parental population from which they were derived allows us to

construct a model of the population origins (Fig. 2.6C). In a population of tumor cells which are already functionally triploid, *ACTR3* and *KDM5B* undergo mutation, each in separate cells. These mutational events are the beginning of the divergent paths that will develop two separate but cooperative subpopulations. Cells containing mutant *ACTR3* go on to form the highly invasive, slow proliferating leader cells while cells containing the *KDM5B* mutation ultimately become the follower population of invasion-deficient, rapidly proliferating cells. Expressing the leader specific *ACTR3* mutation in follower cells increases chain-like invasion. Likewise, expressing the follower specific *KDM5B* mutation in leader cells increases chain-like invasion. Our phenotypic data adds to this model the necessity for cells exhibiting both the leader phenotype and a follower phenotype to the cooperative behavior demonstrated in collective invasion.

The approach used herein, while innovative, does have certain limitations. First, the work focuses on the characterization of leader and followers isolated from a single NSCLC-derived cell line. How universal the identified mutations/genes might be in contributing to leader/follower behavior in other cell lines or cancer types is currently unknown. In addition, while our functional characterization supports a potential role for *ACTR3* and *KDM5B* in modulating collective invasion in this setting, these were only two of the phenotype-selective mutations identified. It is possible that no single alteration in isolation is sufficient to fully drive either phenotype, but rather the combinatorial effects of multiple genetic and epigenetic alterations contribute to collective behavior. Studies aimed at similar analyses of multiple cell lines, or collective invasion models and across cancer types may converge on critical “drivers” or pathways. Ultimately the identification of such key alterations may help in clinical decision-making by identifying predictive biomarkers or new therapeutic targets. Indeed, several molecular inhibitors of the KDM5 family are currently under development, particularly in the setting of combination therapy and therapy resistance [212, 213, 281, 283, 284, 343-345]. Chemical inhibition of *KDM5B* as an anti-metastatic modality has not yet been tested. Moreover, a

recently discovered Arp2/3 inhibitor CK-666 [346] has been shown to inhibit cell motility *in vivo* [347]. An alternative approach would be targeting other components of the Arp2/3 pathway. One such target, PLK4, has been implicated as a driver of cancer invasion and metastasis in part through its interaction with Arp2/3 subunits [348], and PLK4 inhibitors are currently under clinical investigation for patients with advanced solid tumors. Our identification of a panel of mutations delineating leader and follower cell phenotypes in a non-small cell lung cancer tumor population is an initial step toward elucidation of how heterogeneous genetic mutations contribute to cancer metastasis and how these vulnerabilities can be exploited to circumvent the development of metastasis.

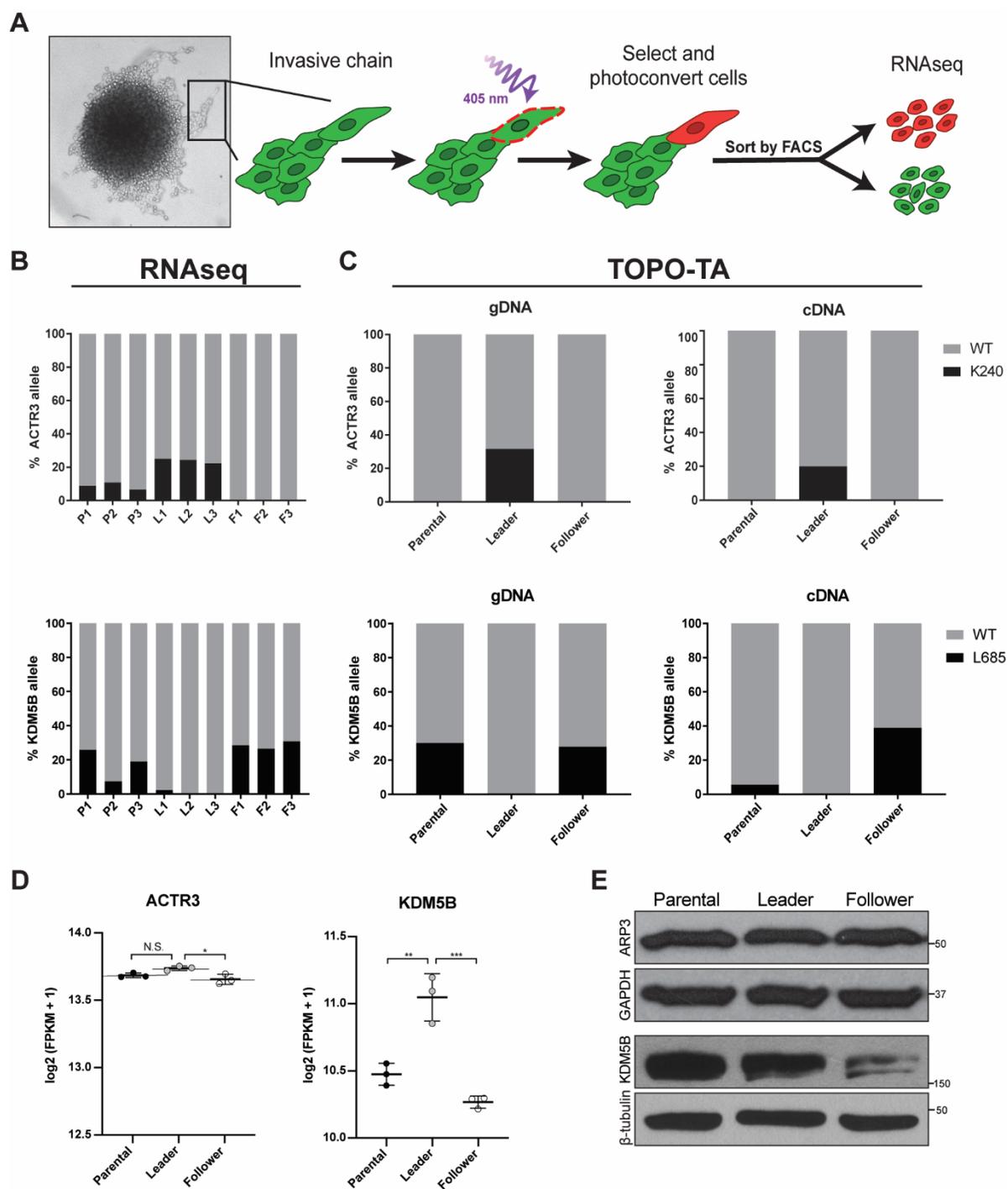


Figure 2.1. ARP3 K240R and KDM5B L685W are validated mutations in H1299 leader and follower cells. (A) Schematic of the SaGA protocol used to isolate leader and follower cell populations. (B) Variant allele frequencies for ACTR3 and KDM5B from RNA-sequencing of H1299

parental (P), leader (L) and follower (F) cells. $n=3$ separate populations per group. (C) TOPO-TA cloning and subsequent Sanger sequencing confirms the presence of ACTR3 K240R and KDM5B L685W mutations in both cDNA and genomic DNA (gDNA) from parental, leader and follower cells, respectively. $n=20$ colonies (ARP3 parental, follower gDNA and parental, leader cDNA and KDM5B parental gDNA); 19 colonies (ARP3 leader gDNA and KDM5B leader gDNA); 18 colonies (ARP3 follower cDNA and KDM5B follower gDNA and parental, follower cDNA); and 17 colonies (KDM5B leader cDNA) (association between the genotype and cell phenotype was determined by Fisher's exact test as follows: mutant vs. wild-type, leader versus follower ARP3 gDNA $p=0.008$, ARP3 cDNA $p=0.11$, KDM5B gDNA $p=0.02$, and KDM5B cDNA $p=0.008$). (D) mRNA expression (via RNA-seq) and (E) protein levels (via Western blot) of ACTR3 and KDM5B in H1299 parental, leader, and follower populations. $*p<0.05$, $**p<0.01$, $***p<0.001$ by one-way ANOVA with Tukey's post-test.

Table 2.1 RNA-seq reveals leader- and follower-specific gene mutations.							
	Gene symbol	Full name	Protein Function	Variant locus (GRCh37)	VAF (%)		
					Parental	Leaders	Followers
Leader-enriched	ACTR3	Actin-related protein 3 (ARP3)	Major component of Arp2/3 complex	chr2:114699797; A:G	8.76	23.4	0.08
	MCM5	Minichromosome maintenance complex component 5	Pre-replication complex during DNA replication	chr22:35809920; G:A	7.59	26.9	0.14
	MIPEP	Mitochondrial intermediate peptidase	Oxidative phosphorylation protein maturation	chr13:24413837; A:C	7.75	37.2	0.98
	NAE1	NEDD8 activating enzyme E1 subunit 1	Activation of neddylation pathway	chr16:66852492; T:C	26.2	58.9	0.13
	NUP93	Nucleoporin 93	Component of nuclear pore complex	chr16:56868312; G:A	25.9	57.8	0.29
	ZNF302	Zinc finger protein 302	Function has yet to be determined	chr19:35175335; G:C	4.71	34.6	0.47
Follower-enriched	CLEC11A	C-type lectin domain family 11, member A	Growth factor for hematopoietic progenitor cells	chr19:51228679; C:G	8.42	0	22.4
	KDM5B	Lysine Demethylase 5B	Demethylates lysine 4 of histone H3	chr1:202715414; A:C	17.5	0.74	28.1
	NDUFS1	NADH:Ubiquinone oxidoreductase core subunit S1	Core subunit of electron transport chain Complex I	chr2:207012514; C:A	10.4	0.28	22.8
	RERE	Arginine-glutamate dipeptide repeats	Possible role in controlling cell survival	chr1:8416225; G:C	38.7	0	65.4
	RNF115	Ring finger protein 115	E3 ubiquitin ligase	chr1:145686997; T:C	14.7	0.27	21.9
	SKA1	Spindle and kinetochore associated complex subunit 1	Chromosome segregation during mitosis	chr18:47902232; C:G	13.0	0	21.6
	TBP	TATA-Box binding protein	Involved in transcription initiation by RNA pol. II	chr6:170871308; G:T	12.4	0.36	21.7

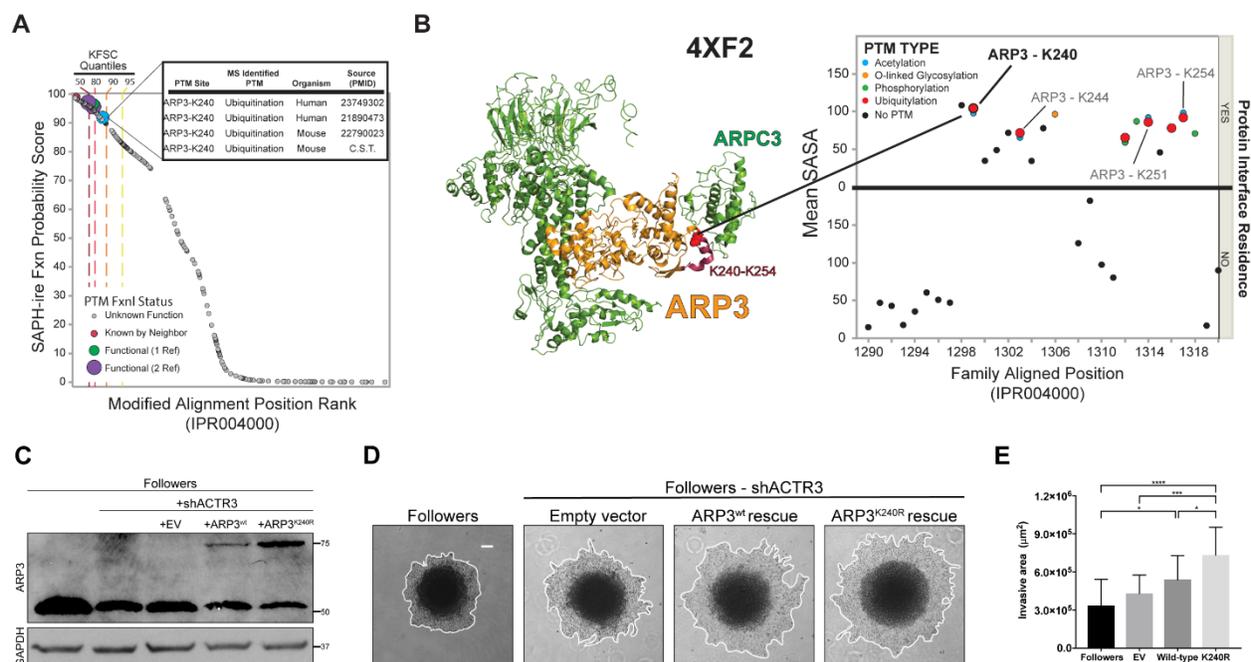


Figure 2.2. PTM hotspot analysis of ARP3 K240 suggests functional impact of the K240R mutation. (A) Plot of SAPH-ire probability score by rank for all modified alignment positions in the ARP protein family IPR004000. The ARP3 K240 ubiquitylation site is highly ranked along with other MAPs that contain PTMs with well-established function (4 or more supporting references), as indicated by known function source count (KFSC) quantiles. (Inset) Table of ARP3 K240 PTMs identified by mass spectrometry of human and mouse tissues, including literature sources. (B) Local PTM topology of the ARP3 family near ARP3 K240. PTM sites plotted by solvent accessible surface area (SASA) and proximity to the interface of a protein-protein interaction. Human ARP3 PTMs are labeled, revealing multiple ubiquitylation sites between K240-K254. (Left) Structure of Arp2/3 complex (PDB 4XF2) indicating ARP3 K240 (spheres) within the K240-K254 region of ARP3 (red). (C) Western blot showing exogenous (upper band) versus endogenous (lower band) ARP3 expression in unmodified followers, and shACTR3-followers rescued with either empty vector (EV), wild-type ARP3, or ARP3 K240R. Rescue constructs were mCherry-tagged to allow for visualization within invasive chains. (D) Images of invasion in Matrigel at 24hrs of spheroids comprised of unmodified

follower cells, or shACTR3 followers transfected with either empty vector, wild-type ARP3, or ARP3 K240R constructs. (E) Quantification of spheroid invasive area (mean \pm s.d., $n=17, 15, 16,$ and 17 spheroids per group, respectively, across $N=3$ experiments. $**p<0.01,$ $***p<0.001,$ $****p<0.0001$ by one-way ANOVA with Tukey's post-test). Scale bar: $100\ \mu\text{m}.$

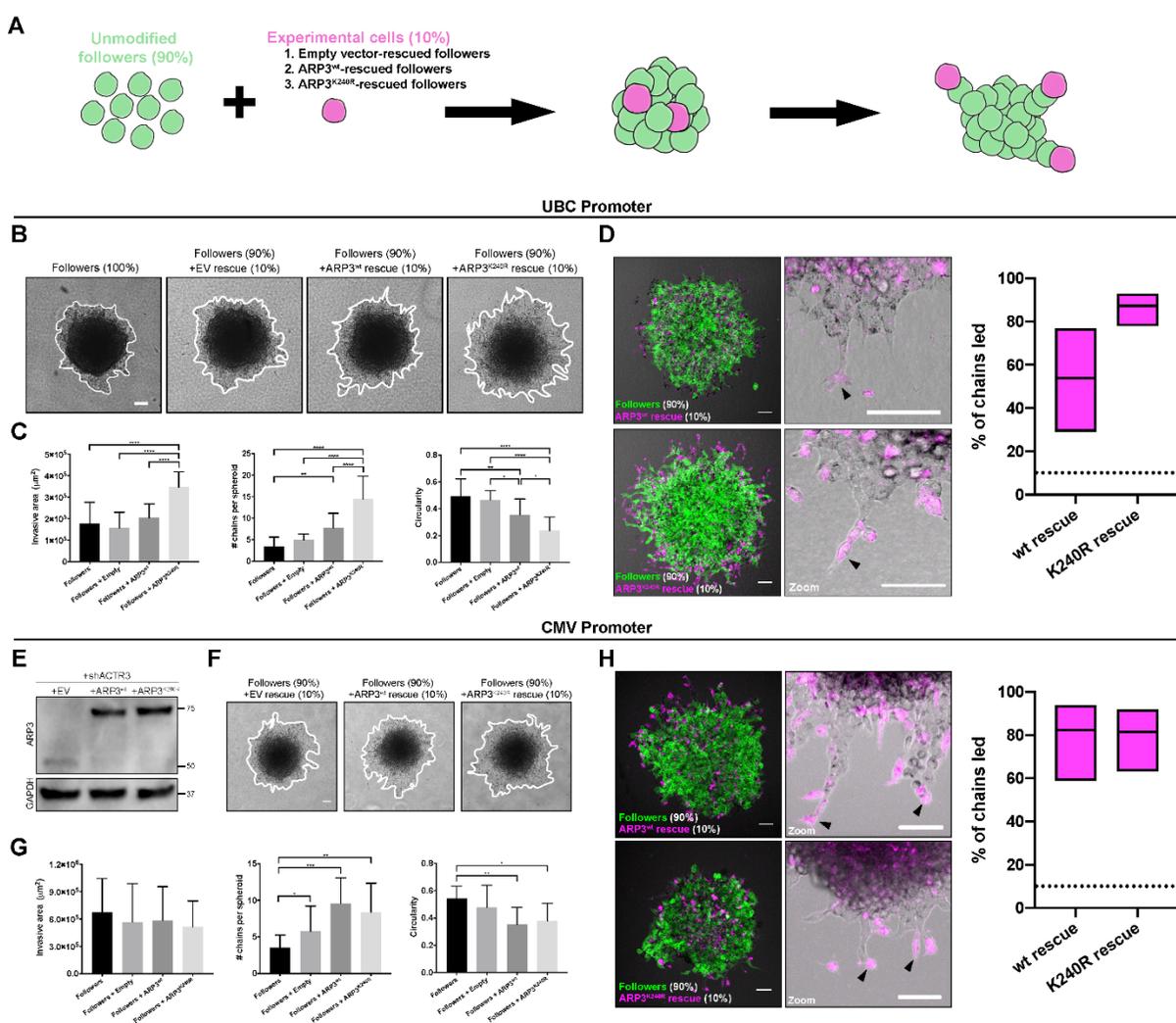


Figure 2.3. ARP3 K240R confers leader-like properties when expressed in follower cells. (A) Illustration of spheroid mixing experimental setup. Spheroids were comprised of either 100% unmodified followers, or 90% unmodified followers plus 10% of shACTR3 followers rescued with either empty vector, wild-type ARP3, or ARP3 K240R. (B) Invasion of mixed spheroids in Matrigel after 24 hr. Rescue constructs were expressed under control of the UBC promoter. Representative images shown for each condition. (C) Quantification of invasive area, circularity, and average number of chains per spheroid for each condition (mean \pm s.d., $n=14, 11, 18,$ and 16 spheroids for unmodified followers, EV rescue, wildtype ARP3 rescue, and ARP3 K240R rescue, respectively, across $N=3$

experiments. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$ by one-way ANOVA with Tukey's post-test). (D) Confocal fluorescence imaging of mixed spheroids, with unmodified followers shown in green and mCherry-ARP3 K240R-rescued cells or mCherry-wild-type ARP3-rescued cells shown in magenta. Black arrows indicate invasive chains being led by experimental rescue cells. Graphs show percentage of chains (mean \pm 95% confidence intervals) led by wild-type ARP3-rescued and ARP3 K240R-rescued followers. Dotted line denotes 10% of chains led, corresponding to the proportion of ARP3 rescued cells in the mixed spheroids. Scale bar: 100 μ m. (E) Western blot showing protein expression levels of ARP3 in followers after knockdown of endogenous ARP3 and rescue with empty vector, wild-type ARP3, or ARP3 K240R. Rescue constructs were expressed under control of the CMV promoter. (F-G) 24 hr invasion of mixed spheroids in Matrigel. Representative images (F) and quantification (G) of invasive area, circularity, and average numbers of chains per spheroid shown for each condition (mean \pm s.d., $n=12, 12, 12,$ and 11 spheroids per group, respectively, across $N=2$ experiments. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, by one-way ANOVA with Tukey's post-test). Scale bars: 100 μ m. (H) Confocal fluorescence imaging of mixed spheroids, with unmodified followers shown in green and mCherry-wild-type ARP3-rescued cells or mCherry-ARP3 K240R-rescued cells shown in magenta. Black arrows indicate invasive chains being led by experimental rescue cells. Graphs show the percentage of chains (mean \pm 95% confidence intervals) led by wild-type ARP3-rescued and ARP3 K240R-rescued followers. Dotted line denotes the proportion of ARP3 rescued cells in the mixed spheroids thus the 10% of chains expected to be led based on random chance.

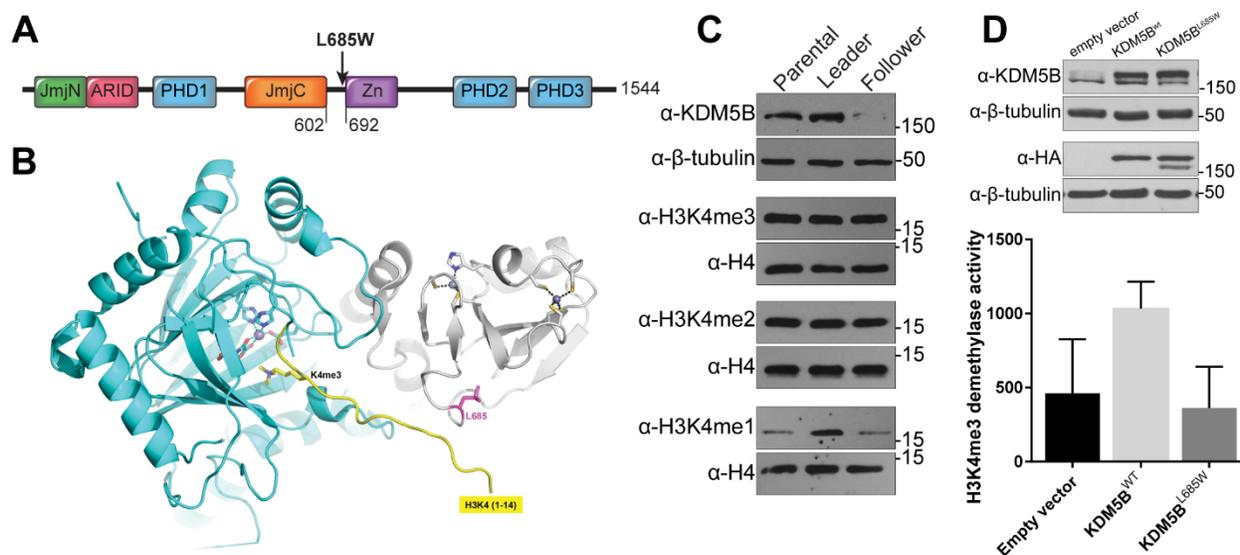


Figure 2.4. Characterization of the follower-enriched KDM5B L685W mutation. (A) Diagram of KDM5B protein domains and relative position of amino acid residue 685 to zinc finger domain. (B) Partial model of KDM5B L685W with H3 N-terminal tail (residues 1-14). The model was generated by superimposing a structure of KDM5B (PDB 5A1F) with that of KDM6A in complex with histone H3 peptide (PDB 3AVR), which placed a trimethylated H3K4 in the active site of KDM5B, near the Fe(II)-binding site, and the C-terminal tail of histone H3 near Leu685. (C) Western blot analysis of KDM5B expression and mono, di- and tri- methylated H3K4 in lysates from H1299 parental, leader, and follower cell populations. Loading controls β -tubulin and H4 are probed from the same blot as preceding bands. Experiment was repeated on three separate protein collections from cells at different passages ($N=3$) (D) *In vitro* H3K4 demethylation assay performed on lysates derived from three separate, biological replicates of parental H1299 cells expressing empty vector, wild-type KDM5B, or KDM5B L685W (mean \pm s.d. specific activity; $N=3$ biological replicates).

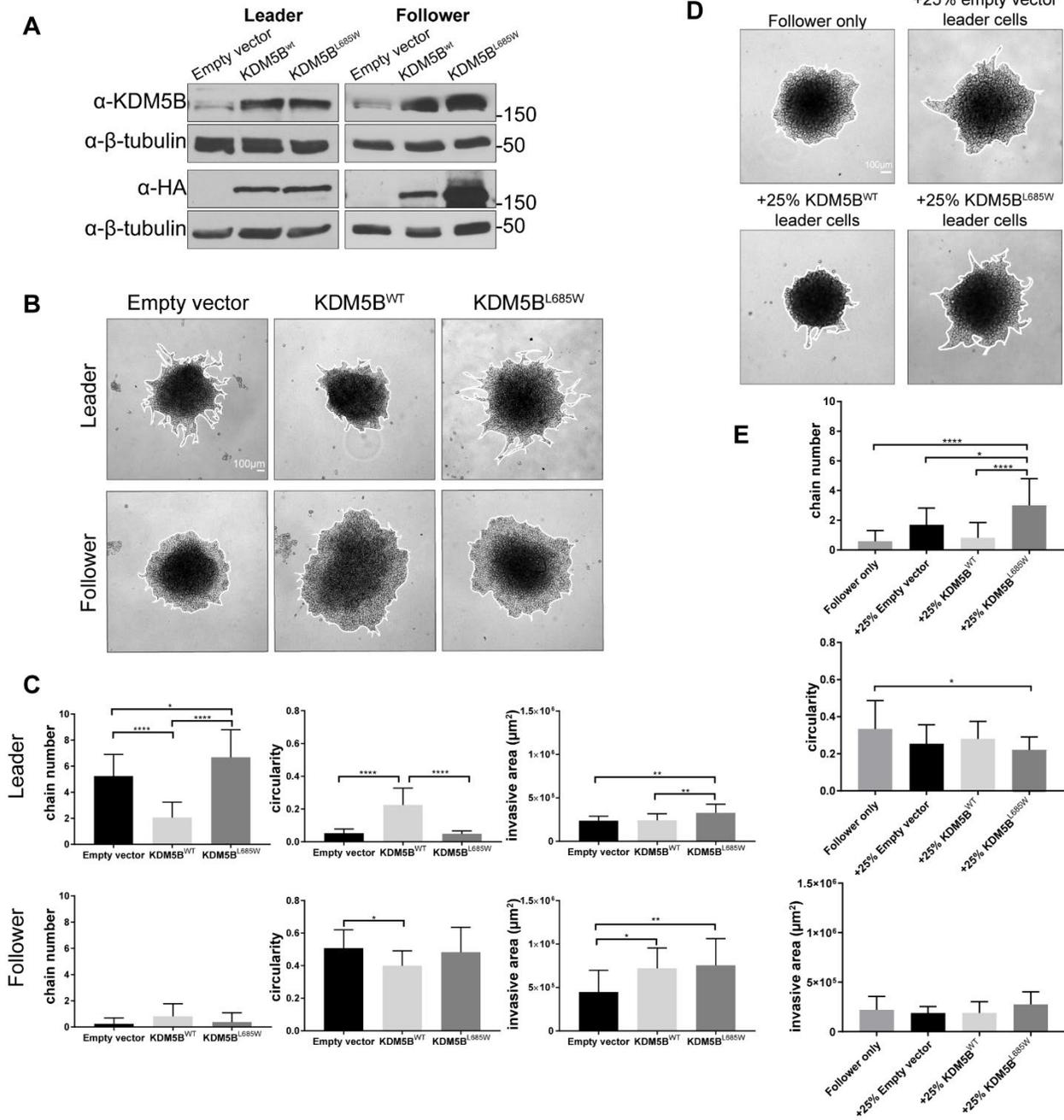


Figure 2.5. Wild-type KDM5B suppress and KDM5B L685W drives chain-like invasion. (A)

Western blot analysis demonstrating successful expression of HA-tagged wild-type and L685W KDM5B in leader and follower cell populations. (B) Images of invasion in Matrigel at 24 hr of leader and follower spheroids expressing empty vector, wild-type HA-KDM5B, or HA-KDM5B L685W.

(C) Quantification of invasive area, circularity, and chain number from spheroids depicted in (B)

(mean \pm s.d., $n=16$ spheroids (follower empty vector, leader KDM5B^{WT}, and leader KDM5B^{L685W}), $n=17$ spheroids (leader empty vector and follower KDM5B^{WT}), $n=18$ spheroids (follower KDM5B^{L685W}), across $N=3$ independent experiments, * $p<0.05$, ** $p<0.01$, *** $p<0.0001$ by one-way ANOVA with Tukey's post-test). (D) Images of invasion in Matrigel at 24 hr of spheroids composed of 75% follower cells and 25% leader cells expressing either empty vector, wild-type HA-KDM5B, or HA-KDM5B-L685W. (E) Quantification of invasive area, circularity, and chain number from spheroids depicted in (D) (mean \pm s.d., $n=16$ spheroids (+25% leader empty vector), $n=17$ spheroids (follower only and +25% leader KDM5B^{WT}), $n=18$ spheroids (+25% leader KDM5B^{L685W}), across three independent experiments, * $p<0.05$, *** $p<0.0001$ by one-way ANOVA with Tukey's post-test).

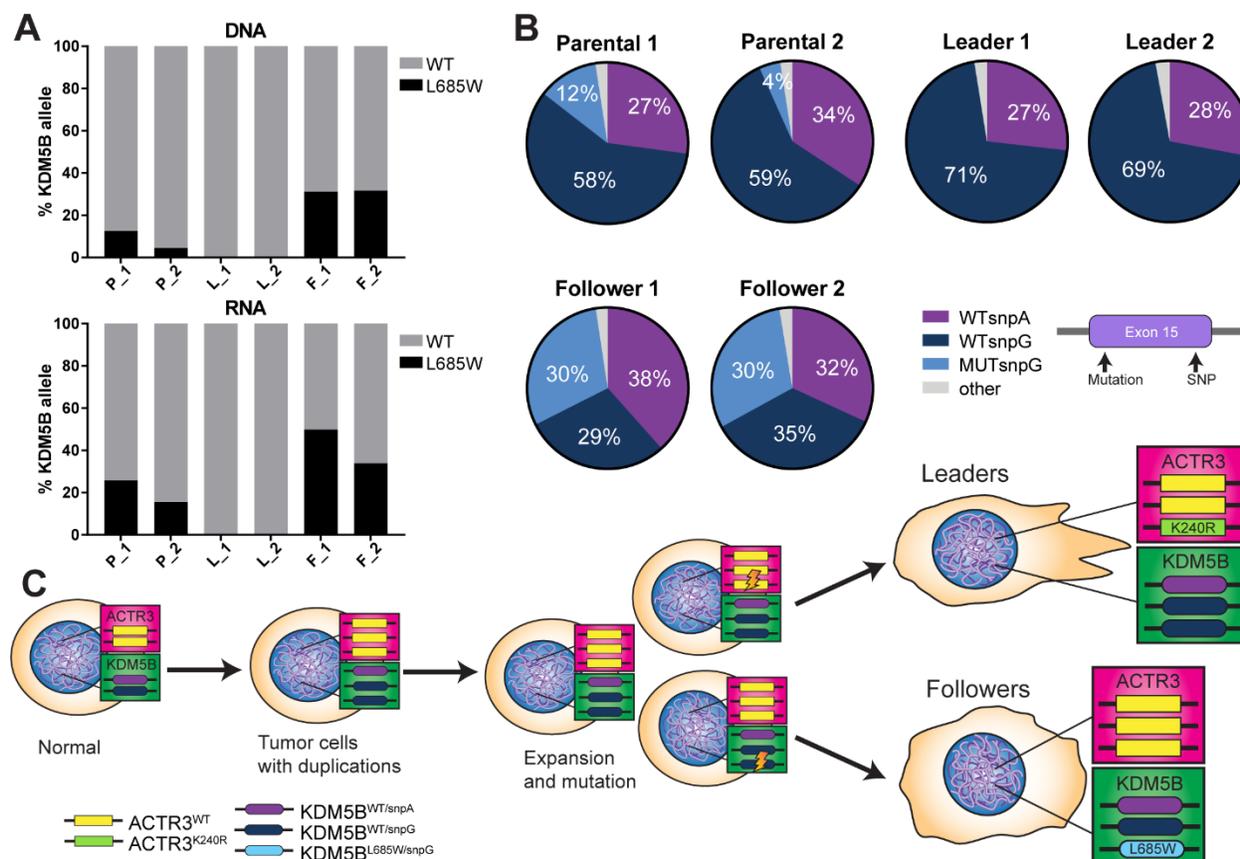


Figure 2.6. Leader and follower cells are derived from two separate populations defined by mutational profile. (A) Deep targeted resequencing across KDM5B exon 15 in genomic DNA (top) and RNA (bottom) isolated from parental, leader, and follower populations ($N=2$ independent isolates of DNA/RNA at separate passages of cells derived from a single phenotypic isolation. Average depth= 348,327 reads per sample). (B) Pie charts depicting proportions of KDM5B genotypes across parental, leader, and follower populations as determined from the deep amplicon sequencing shown in panel A. (C) Model of the potential history of leader and follower populations from parental cell population as inferred from the genetic profiles of KDM5B and ARP3.

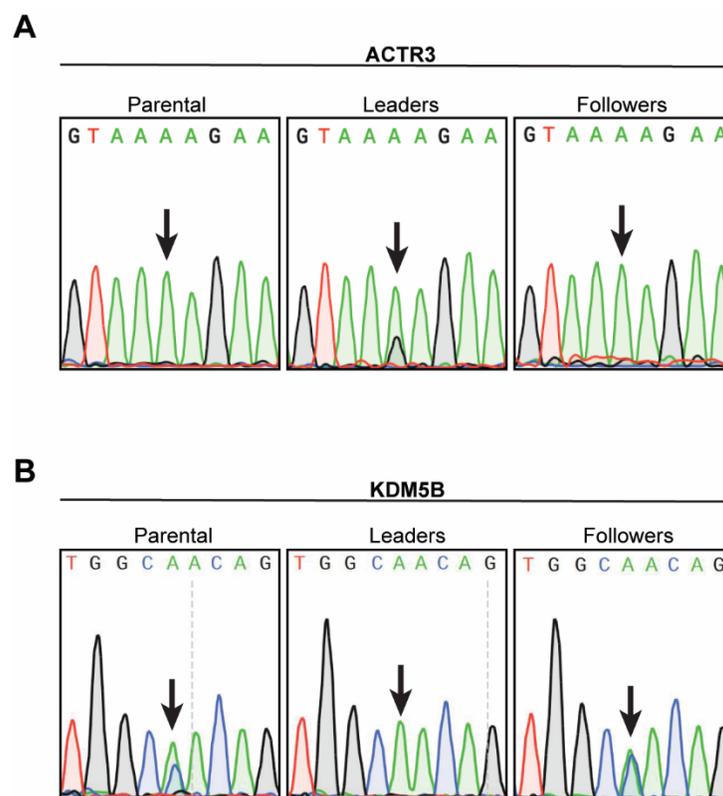


Figure S2.1. Confirmation of leader- and follower-enriched mutations. Sanger sequencing confirming leader-enriched ACTR3 mutation (A) and follower-enriched KDM5B mutation (B) in cDNA (shown) and genomic DNA isolated from H1299 parental, leader and follower populations. Black arrows indicate the bases of interest. (A) Only the wild-type A peak is seen in the parental and follower populations, while the leader population contains both A and G peaks. (B) Only the wild-type A peak is seen in the leader population, while the parental and follower populations contain both A and C peaks.

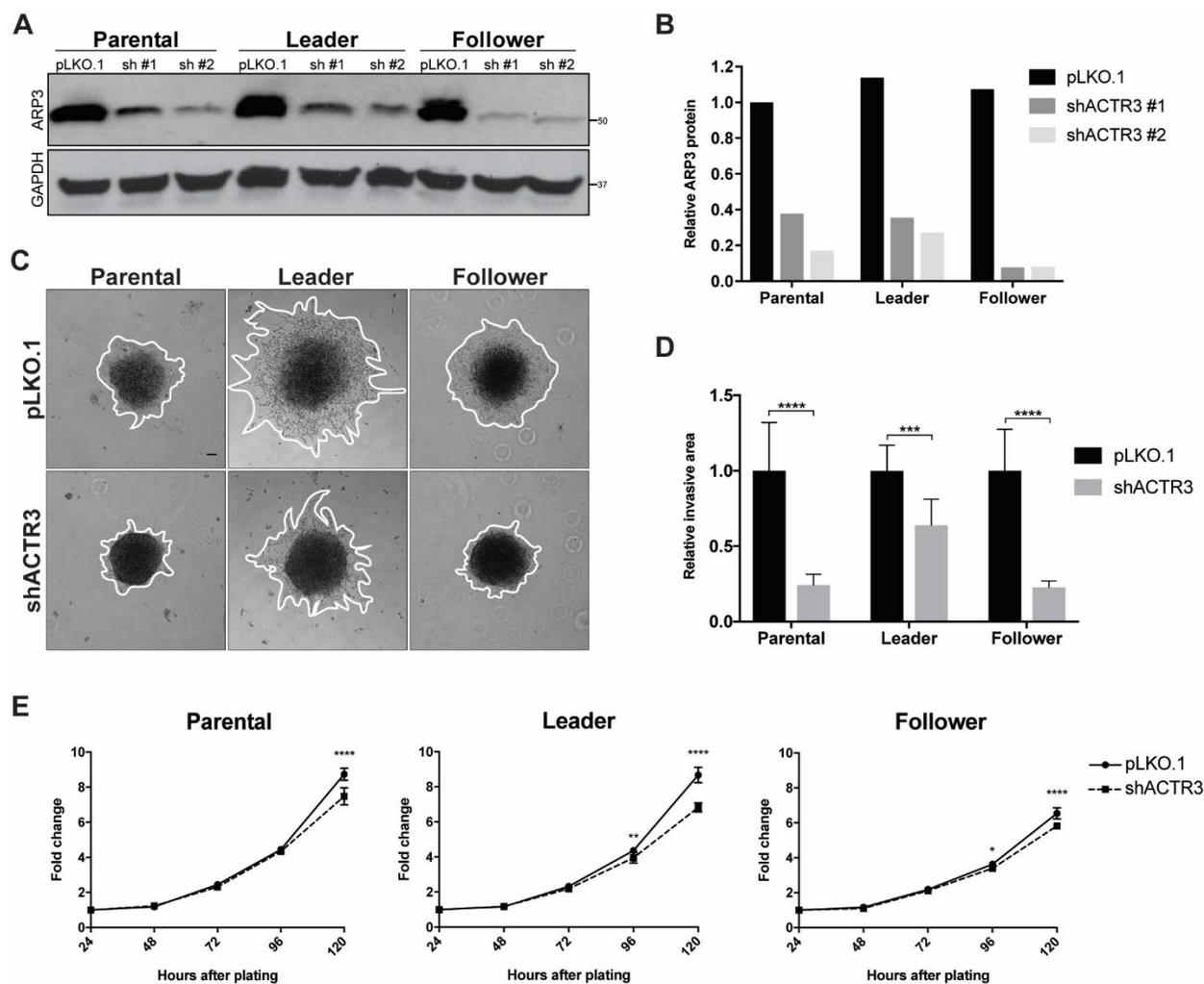


Figure S2.2. ARP3 knockdown inhibits 3-D invasion. (A) Western blot showing ARP3 protein levels in H1299 parental, leader and follower cells upon expression of empty pLKO.1 vector, ARP3 shRNA #1 (Millipore Sigma TRCN000029383), or ARP3 shRNA #2 (Millipore Sigma TRCN0000380403). (B) Western blot densitometry quantification, indicating 70-90% knockdown of ARP3 protein using either shRNA #1 or shRNA #2. (C) Representative images of 24-hour invasion of H1299 parental, leader, and follower spheroids expressing either empty pLKO.1 or shACTR3 #2. Scale bar = 100 μ m. (D) Quantification of relative 24-hour invasive area, normalized to pLKO.1 control for each group. (mean \pm s.d., n=5, 11, and 5 spheroids for parental, leader and follower lines,

respectively. $***p < 0.001$, $****p < 0.0001$ by two-way ANOVA with Sidak correction). (E) Growth rate of parental, leader, and follower lines expressing either empty pLKO.1 or shACTR3 #2. (mean+s.d., n=5 replicates per time point. $*p < 0.05$, $***p < 0.001$, $****p < 0.0001$ by two-way ANOVA with Šidák correction).

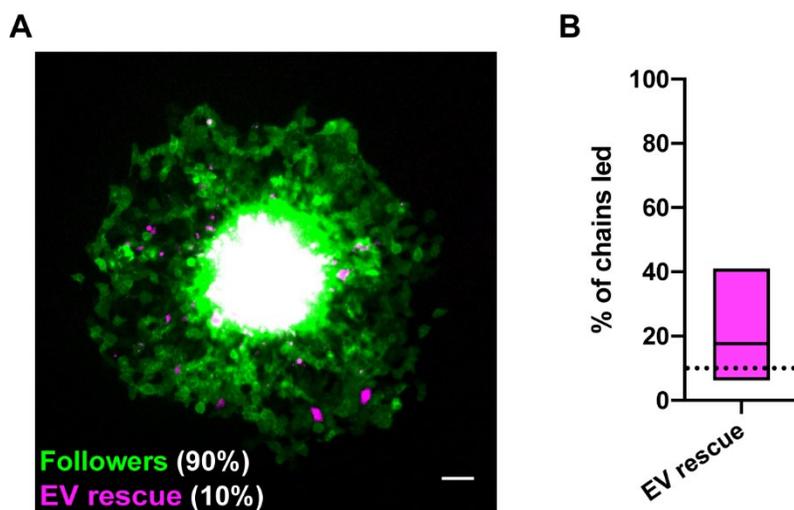


Figure S2.3. EV-rescued followers do not lead invasive chains. (A) Representative widefield fluorescence image of mixed spheroids, with unmodified followers shown in green and mCherry-empty vector-rescued cells shown in magenta. (B) Graph showing percentage of chains (with 95% confidence interval) led by empty vector-rescued followers. Dotted line denotes 10% of chains led, corresponding to the proportion of rescued cells in the mixed spheroids. Scale bar: 100 μm .

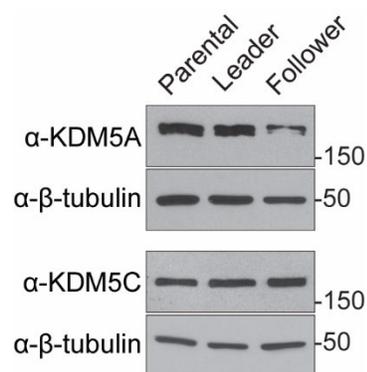


Figure S2.4. KDM5A but not KDM5C is downregulated in follower cells. Western blot analysis of KDM5A and KDM5C expression in lysates from H1299 parental, leader, and follower cell populations. Loading controls β -tubulin are probed from the same blot as preceding bands. Experiment was repeated on three separate protein collections from cells at different passages ($N=3$).

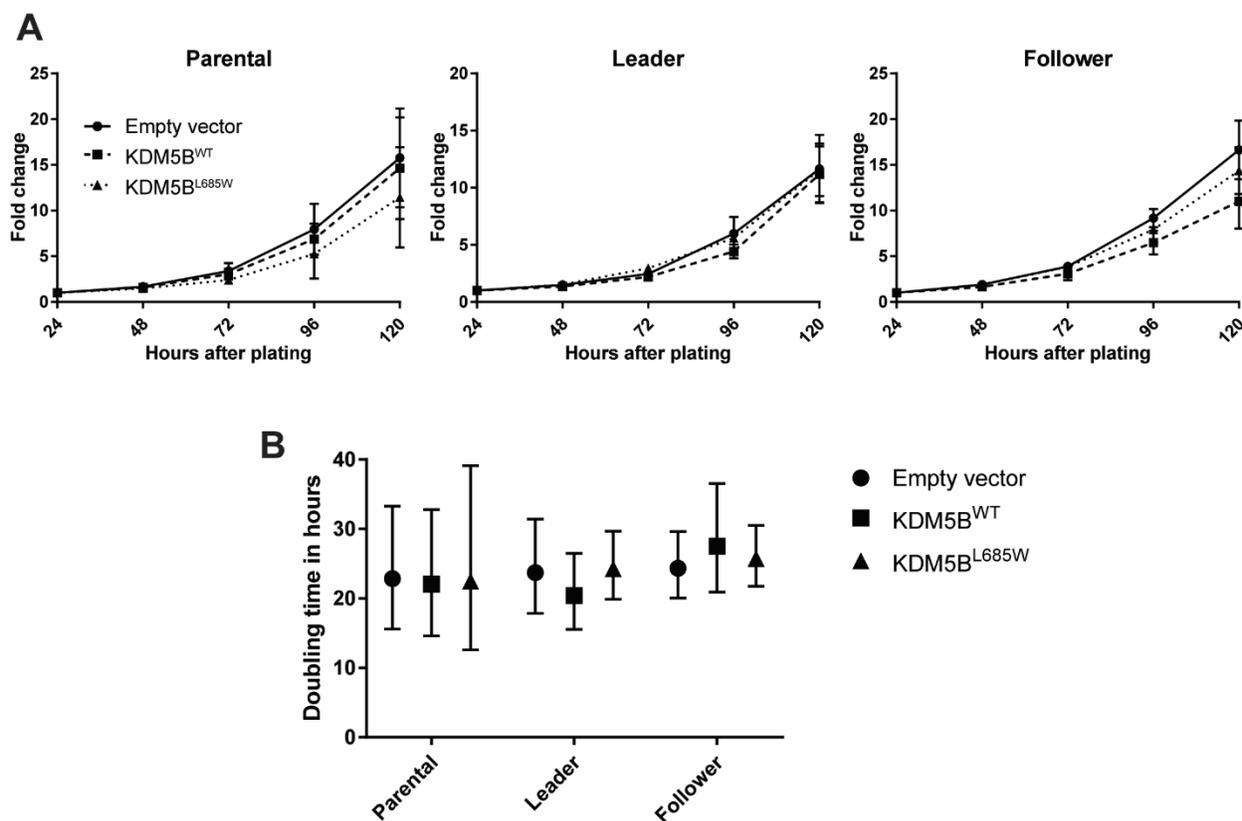


Figure S2.5 Overexpression of wild-type KDM5B or KDM5B L685W do not affect growth rate.

(A) Growth rate of parental, leader, and follower H1299 lines stably expression empty vector, wild-type HA-KDM5B, or HA-KDM5B-L685W (mean \pm s.d. of $N=3$ independent experiments at each time point). (B) Doubling time \pm 95% confidence interval for lines in (A) as calculated through the least squares fit of exponential growth of data in (A) in GraphPad Prism where doubling time equals $\ln(2)/\text{rate constant}$.

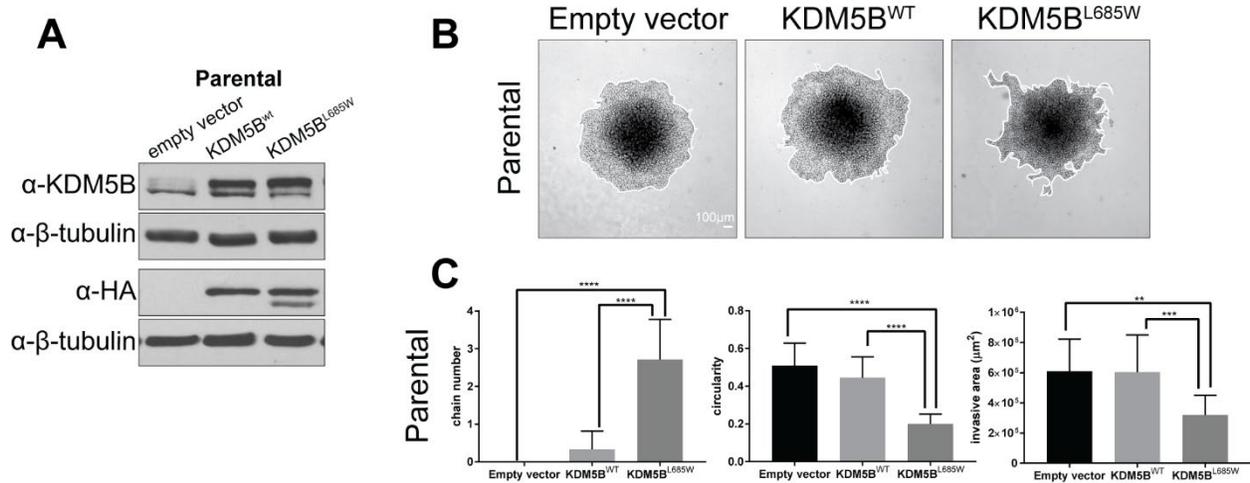


Figure S2.6. Wild-type KDM5B suppress and KDM5B L685W drives chain-like invasion in parental spheroids. (A) Western blot analysis demonstrating successful expression of HA-tagged wild-type and L685W KDM5B in parental populations. (B) Images of invasion in Matrigel at 24 hr of parental spheroids expressing empty vector, wild-type HA-KDM5B, or HA-KDM5B L685W. (C) Quantification of invasive area, circularity, and chain number from spheroids depicted in (B) (mean±s.d., $n=16$, 18, and 14 spheroids in empty vector, wild-type KDM5B, and KDM5B L685W respectively across $N=3$ experiments, * $p<0.05$, ** $p<0.01$, **** $p<0.0001$ by one-way ANOVA with Tukey's post-test).

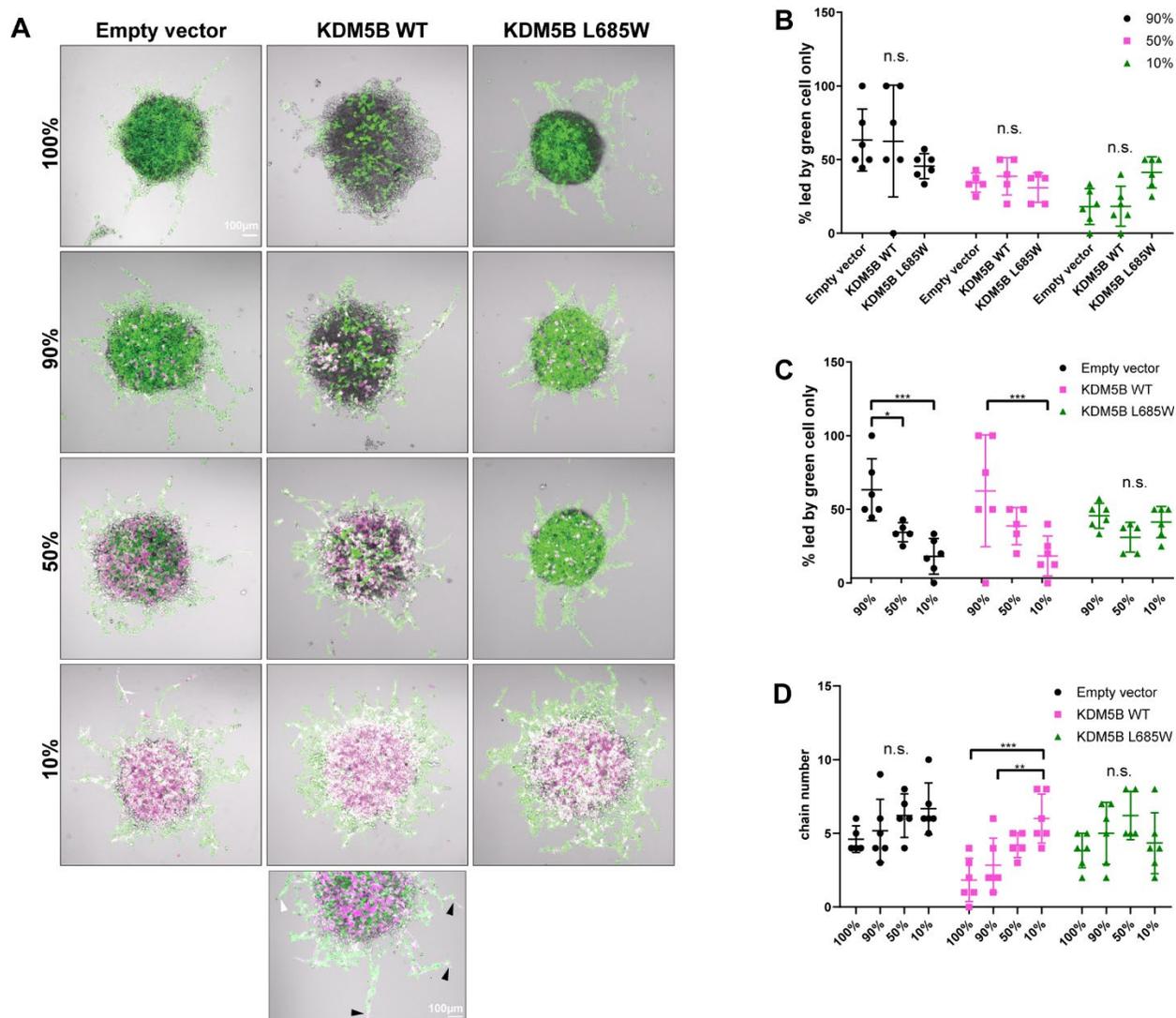


Figure S2.7. Leader cells overexpressing wild-type KDM5B but not KDM5B L685W exhibit diminished chain formation but retain leader activity in leader-only spheroids. (A) Representative confocal fluorescence imaging of spheroids in which, 10%, 50%, or 90% leaders stably expressing empty vector, wild-type HA-KDM5B, or HA-KDM5B L685W were mixed with unmodified leaders 24 hours after embedding in Matrigel. Both populations express Dendra2 (green) whereas the unmodified leaders also express mCherry (magenta) Bottom panel: Representative example of a spheroid containing 50% empty vector expressing and 50% unmodified leaders with

black arrows indicating magenta cell led chains and white arrows indicating green-only led chains. (B) Percent invasive chains led by green-only (vector/KDM5B/KDM5B L685W) transduced cells, grouped by KDM5B overexpression cell line (mean \pm s.d., $n=6$ spheroids for 90% and 10% mixes or $n=5$ spheroids for 50% mixes across $N=1$ biological replicate, n.s. not significant, $p>0.05$, by two-way ANOVA with Tukey's post-test). (C) Same data as in panel B except grouped by the fraction of green-only (vector/KDM5B/KDM5B L685W) transduced cells in the spheroid (mean \pm s.d., $n=6$ spheroids for 90% and 10% mixes or $n=5$ spheroids for 50% mixes across $N=1$ biological replicate, n.s. not significant, $*p<0.05$, $***p<0.001$ by two-way ANOVA with Tukey's post-test). (D) Average number of chains per spheroid grouped by percentage of the indicated trans-duced leader cell line (mean \pm s.d., $n=6$ spheroids for 90% and 10% mixes or $n=5$ spheroids for 50% mixes across $N=1$ biological replicate, n.s. not significant ($p>0.05$), $**p<0.01$, $***p<0.001$ by two-way ANOVA with Tukey's post-test).

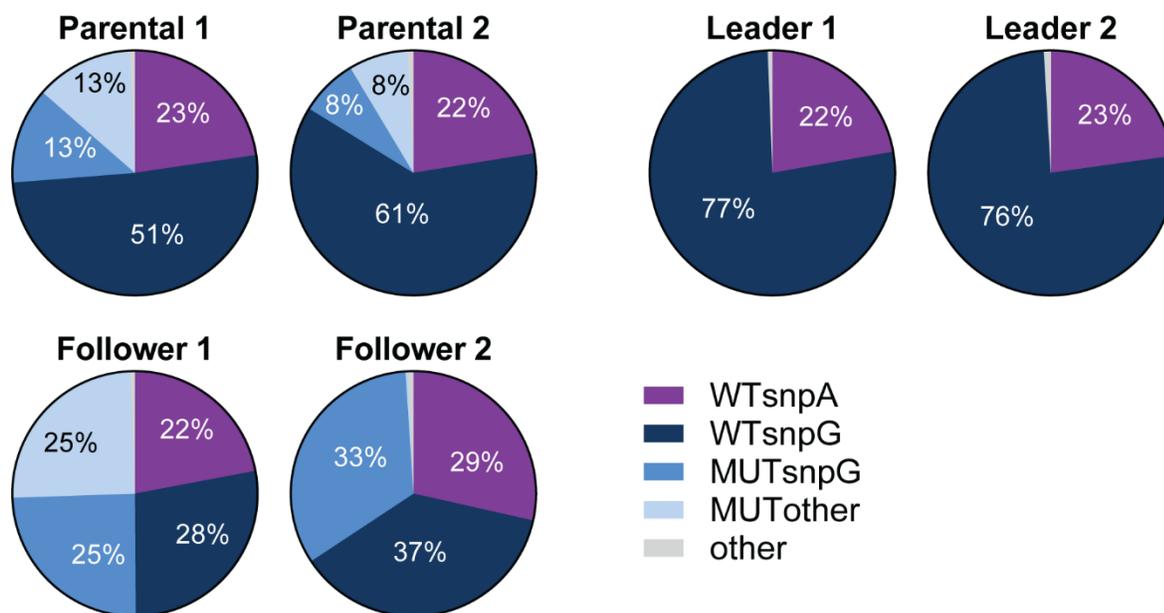


Figure S2.8. The relative proportions of the wild-type and variant mRNA expression reflects of the proportion at the genomic DNA level across all three populations. Pie charts depicting proportions of KDM5B genotypes across parental, leader, and follower populations as determined from the deep amplicon sequencing shown in Fig. 2.6.A ($N=2$ biological replicates per cell line).

Chapter 3

Investigating KDM5B as an anti-metastatic therapeutic target

A portion of the following text and Figure 1 are adapted from “Characterization of a Linked Jumonji Domain of the KDM5/JARID1 Family of Histone H3 Lysine 4 Demethylases” published by John R. Horton, Amanda Engstrom, Elizabeth L. Zoeller, Xu Liu, John R. Shanks, Xing Zhang, Margaret A. Johns, Paula M. Vertino, Haiyan Fu, and Xiaodong Cheng in *Journal of Biological Chemistry*,

2016 Feb. 5.

3.1 Introduction

Although metastatic cancer is a major cause of death in the United States, therapeutic targeting of this phenomenon has yet to be developed [1]. Currently much research is being conducted into the physiology and molecular biology governing how cancer cells leave the primary tumor and establish metastatic colonies at distant sites. Much progress has been made in understanding the proteins regulating the epithelial to mesenchymal transition (EMT), the first stage of metastasis and the process in which a cell loses its stationary epithelial properties and gains invasive and migratory properties like that of mesenchymal cells [45, 56, 59, 61]. However, the greater dynamics of invasion are now in the spotlight. Of particular note, collective invasion, in which multiple cells leave the primary tumor together as a cohesive group, has arisen as a significant contributor to successful metastasis [74, 82, 85, 87, 88, 349]. An important concept regarding collective invasion is that a group of cells can contain genetically heterogeneous cells with varying phenotypes (see Chapter 2) [73, 89, 90, 350]. One example is of leader and follower cells in collectively invading cell branches. Leaders are a specialized population of cells that lead the group of invading cells out, and followers are the trailing portion of the pack that follows the leaders [89, 288, 293, 294, 309, 350]. The distinct biology of each subpopulation could allow for more specific targeting of metastatic cancer.

Another rapidly expanding area of study in cancer biology is that of epigenetics. Epigenetic mechanisms contribute to the heritable regulation of gene expression without permanently altering the genetic code and govern much of the plasticity that defines cancer phenotypes [150, 156]. This plasticity allows for cell behavior to alter in response to changes in environment, as occurs throughout the process of metastasis or during the development of drug resistance [57, 157]. For these reasons, epigenetic modifiers are promising targets for chemotherapeutics. Traditionally, epigenetic inhibitors, such as inhibitors of DNA methyltransferases or histone deacetylases, have had quite broad effects across a whole class of enzymes and extensively disrupted transcription regulation [151]. More

recently, epigenetic drug development has focused on specificity of target proteins, pathway, cancer type, and biological effects [151, 159]. The burgeoning number of compounds aimed at inhibiting the lysine demethylase 5 (KDM5) family are one example of such renewed focus on specificity [209, 219, 282-286]. KDM5 family members have been implicated as drivers of oncogenesis, drug resistance, and invasion and metastasis [208, 209, 219]. Identifying which cancer types and cell behaviors exploit the KDM5 family is crucial to determining how to use these new inhibitors most effectively [287]. Here we explore how several newly developed KDM5 inhibitors affect cancer cells in different circumstances.

Furthermore, we recently identified the upregulation of KDM5B in an isolated leader cell population derived from a model of lung cancer chain-like collective invasion [288]. The complementary follower cell population derived from the same model is poorly invasive in isolation and expresses much lower amounts of KDM5B. Furthermore, a loss-of-function point mutation in KDM5B was found to be exclusive to the follower cell population (further described in Chapter 2). With KDM5B expression thus correlated with degree of invasiveness, we sought here to dissect the role of KDM5B in chain-like, collective invasion and to validate inhibition of the KDM5 family as a means to curtail cancer cell invasion.

3.2 Results

KDM5B and its closest relatives have been implicated in multiple cancer processes, such as drug resistance and oncogenesis, in several cancer types, including breast cancer [204, 216, 221, 254, 279, 351-354]. Furthermore, the road to developing effective inhibitors of KDM5s as an anti-cancer strategy has been an iterative process of whittling chemical structures down to those that complement the small differences in architecture of lysine demethylases. To understand how *in vitro* differences in the specificity of the KDM inhibitors influence cellular activity, we determined the impact of JIB-04

and GSK-J4 on the growth of two breast cancer cell lines, MDA-MB231 and MCF7. Overall, the breast cancer cells were ≈ 10 – 50 -fold more sensitive to JIB-04 treatment than to GSK-J4, with half-maximal growth-inhibitory concentrations (GI_{50}) in the 20 – 300 nM range for JIB-04 versus 1 – 3 μ M for GSK-J4 after 72 hours of treatment (Fig. 3.1A,B). Additionally, JIB-04 exhibited differential activity among the breast cancer cell lines. MCF7 cells were 14-fold more sensitive to the antiproliferative effects of JIB-04 than were MDA-MB231 cells, with $GI_{50} \approx 22$ nM (range = 16 – 25 nM, $n=3$) versus 300 nM (range 270 – 330 nM, $n=3$) (Fig. 3.1A,B). In contrast, GSK-J4-induced growth inhibition demonstrated relatively little cell type specificity, comparing MCF7 cells ($GI_{50} \approx 0.9$ – 1.0 μ M, $n=3$) versus MDA-MB231 cells ($GI_{50} \approx 2.9$ μ M, range = 1.3 – 5.2 μ M, $n=4$) (Fig. 5e). GSK-J5, a less active isomer of GSK-J4 [355] and negative control, showed little growth-inhibitory activity against either cell line ($GI_{50} > 40$ μ M). Thus, the relative selectivity of JIB-04 toward KDM5B versus KDM5C correlated with an increased cellular potency overall and a propensity for cell type specificity not observed with GSK-J4 (at least between these two cell lines) and may reflect dissimilarities in the underlying biology of the cell lines. For example, breast cancer cells may be more dependent on the KDM5 family versus KDM6, the preferred target of GSK-J1/4 *in vitro* [355, 356]. Cell type-specific differences in the levels or activity of KDM4 may also contribute because JIB-04 has nearly equivalent activity against KDM4 family enzymes *in vitro* and against cellular histone H3 trimethylated Lys-9 demethylase activity [357].

Having identified KDM5B as a promising target for inhibitors with antiproliferative effects, we next sought to expand inhibition of KDM5B to another model in lung cancer. Previously, leader cells and follower cells from collectively invading chains of non-small cell lung cancer (NSCLC) line H1299 were separately isolated (three follower and two leader populations) through fluorescence-activated cell sorting (FACS) after identification by activation of a photoconvertible dye [288]. These subpopulations were then expanded in 2-D culture for further study. Parental cells include cells

isolated from whole spheroids before isolating leaders or followers. RNA-seq analysis was performed on RNA isolated from parental H1299 cells isolated from spheroids from three different passages, three separately isolated populations of follower cells, and the two separately isolated populations of leader cells (including two passages of one of the leader populations post-expansion in 2-D culture). These studies revealed KDM5B to be overexpressed at the transcription level in leader cells (~150% of parental reads and ~175% of follower reads) (Fig. 2.1E, 3.2A). We verified the expression levels of KDM5B mRNA across parental, leader, and follower cells using quantitative reverse transcriptase PCR. Indeed, leader cells express more KDM5B mRNA than either parental or follower cells (again ~150% of average parental signal and ~175% of average follower signal when normalized to β -tubulin) (Fig. 3.2B). Furthermore, in Chapter 2, we found that leader cells express much more KDM5B protein than do follower cells by Western blot analysis. Although not the only gene upregulated in leaders versus followers or vice versa, KDM5B is of particular interest in studying leader cell behavior because KDM5B is correlated with invasive behavior across several cancer types [259, 268-273].

Given that KDM5B is more highly expressed in the more invasive leader cells, we sought to determine the role of KDM5B in chain-like invasion. KDM5B was independently stably knocked down in parental, leader, and follower H1299 lines with two short hairpin RNAs of different target sequences (Fig. 3.3A). We then ascertained the effects on chain-like collective invasion by imaging spheroids in 3-D culture (Fig. 3.3B,C). Despite similar growth rates (Fig. 3.4A,B) and consistent knockdown between both shRNAs at the protein level (Fig. 3.3A), the effects on invasion, including chain number, circularity, and invasive area, were significantly different between lines expressing the knockdown construct KDM5Bsh1 and those expressing KDM5Bsh2, across parental, leader, and follower cells (Fig. 3.3B,C). Lines expressing KDM5Bsh1 tended to have fewer chains, higher circularity, and less invasive area than those expressing KDM5Bsh2, particularly in the parental and leader lines. Even when mixing knockdown leader lines with wild-type followers at 25/75 to determine

the ability of KDM5B deficient leaders to lead a chain, the chain number was strikingly different between mixes with KDM5Bsh1 leaders and KDM5Bsh2 leaders (Fig. 3.3D,E).

We also stably transfected the knockdown leader lines with mCherry to determine if the chains generated were indeed led by KDM5B knockdown leader cells, in which case the cells at the end of a chain would fluoresce red, in a 25/75 mix with green only followers (Fig. 3.5A). The mixes with control leaders expressing a scramble control shRNA produced more chains than followers alone or either KDM5BshRNA1 or 2 as is consistent with the data in Figure 3 (Fig. 3.3D,E, 3.5A,B). However, no significant difference was observed in the percent of chains led by a red leader cell whether the mix contained scrambled control, KDM5BshRNA1, or KDM5BshRNA2 (Fig. 3.5A,B). These results indicate that chains were equally as likely to be led by one of the red leader cells regardless of whether the leader was a scramble shRNA control leader cell or if it was a KDM5Bsh knockdown leader cell. Leader cells in which KDM5B is knocked down are just as capable of leading chains as control but do suppress chain formation.

Our next approach in establishing the relationship between invasion phenotypes and KDM5B was to chemically inhibit the enzymatic activity of this demethylase. At this juncture, we moved from the partially selective inhibitors JIB-04 and GSK-J4 to treating cells with KDM5-C70, the cell permeable ethyl ester derivative of KDM5-C49, one of the earliest compounds developed to specifically target the KDM5 family [212, 284, 326, 343, 358]. To narrow our therapeutic window to the highest concentration for inhibition of KDM5 activity that still permits study of healthy cells during invasion, we first measured toxicity of KDM5-C70 in our H1299 parental, leader, and follower lines by cell growth assays (Fig. 3.6A). From these growth curves, we were able to determine that KDM5-C70 has no effect on cell growth up to doses of 150 μ M.

We next determined the effect of 5 μ M or 50 μ M of KDM5-C70 on KDM5 activity by assaying H3K4me3, the substrate for KDM5 family members, as collected from parental lines after 4 or 7 days

of treatment (Fig. 3.6B). H3K4me3 levels were greatly increased by C-70 treatment but changed little between the 5 μ M and 50 μ M treatments leading us to select 5 μ M as our optimal dose. No greater effect was observed after 7 days of treatment. To test the activity of the drug across all H1299 lines, we treated parental, leader, and follower cells with 5 μ M KDM5-C70 for 72 hours before collecting histones and assaying H3K4me3 levels (Fig. 3.6C). We found that H3K4me3 levels increased with KDM5-C70 treatment across all groups. We concluded the most effective treatment plan for studying the behavior of living cells under KDM5-C70 treatment is to treat with 5 μ M for 72 hours prior to further studies.

Having established a treatment protocol, we sought to determine the effect of KDM5 inhibition on invasive behavior by treating cells with KDM5-C70 before plating and embedding spheroids for 3-D invasion assays (Fig. 3.6C,D). However, no change in invasive area or circularity was noted between the untreated, DMSO, and KDM5-C70 treated sample groups (with the exception of some change in invasive area between untreated and DMSO controls in followers at 48 hours). Perhaps continued revision of the treatment protocol, for instance longer treatment times, might have a greater effect, such as changes in invasive activity, downstream of H3K4me3 levels becoming apparent. Another possibility is that off target effects could cancel out those exerted on the KDM5 family. KDM5-C70 has some lesser activity against the KDM4 and KDM6 family as well [212, 326].

Hunting for inhibitors that are both powerful and specific against the KDM5 family is a pursuit shared by many researchers and institutions. As a result, the number of new KDM5 inhibitors has multiplied in recent years. We employed three of these recently developed compounds—GDC-50 (N54 or NCGC00482457) (patent WO2016057924) [343], Dong-A-167 (NCGC00487054) (patent WO2016068580), and CPI-48 (NCGC00488278) [286]—to study the impact of KDM5B enzymatic activity on invasion. We first performed cell growth assays on parental H1299 cells to determine the effect of GDC-50, Dong-A-167, and CPI-48 on cell growth (Fig. 3.7A). As was seen with KDM5-

C70, there was no significant effect of any of these compounds on the growth of H1299 cells at doses up to 5 μ M of any compound or DMSO vehicle. We confirmed the activity of the compounds by assaying H3K4me3 after treating H1299 parental cells with 5 μ M GDC-50, Dong-A-167, CPI-48, or vehicle control for 72 hours (Fig. 3.7B). All three compounds greatly increased H3K4me3 over untreated or DMSO control samples.

We next treated parental, leader, and follower cells during spheroid formation and invasion in Matrigel with 5 μ M GDC-50, Dong-A-167, and CPI-48 (Fig. 3.7C,D). In our first experiment, we found that CPI-48, but not GDC-50 or Dong-A-167 decreased invasive area in parental, leader, and follower spheroids by 48 hours. Moreover, CPI-48 treatment resulted increased circularity in leader and follower cells at 24 hours over control spheroids. These results indicate that CPI-48 may decrease chain-like invasion and invasion as a whole and is a candidate for further study. We repeated our studies of invasion under CPI-48 treatment in triplicate (Fig. 3.8A,B). Although follower cells had a slight but statistically significant increase in circularity, overall, there was little impact of CPI-48 on invasive area across the parental, leader and follower populations. The genesis of the variation between independent experiments is unknown and thus the results remain inconclusive.

3.3 Discussion

Cancer has long plagued humankind, and for as nearly as long, people have attempted to stem this disease. Initial therapies were imprecise and often toxic. As scientists have progressively uncovered the distinct pathways driving cancers, recent decades have brought therapies targeted at specific cancer types and molecules that drive the disease. The ultimate goal is to find a target exploited by cancer cells but unimportant to healthy cells and to develop a molecule that interacts with only this target and prevents it from performing its pro-cancer role. The field of epigenetics is a fecund source

of therapeutic targets that has only recently reached peak harvest. Chemical inhibitors of the KDM5 family are some of the newest and most promising fruits of these studies.

A unifying theme of the work in this chapter is the need for specificity in targeting KDM5B and the KDM5 family, both in drug design and context of use. When designing chemical compounds to inhibit enzymatic activity of KDM5B, the strategy is quite complicated as members of the KDM5 family and other lysine demethylases have high similarity in structure [287, 326, 358]. The KDM5 proteins belong to a larger group of lysine demethylases that containing the catalytic Jumonji C domain (JmjC) and require Fe (II) ions and α -ketoglutarate as cofactors [208, 326, 359]. As these demethylases, including KDM5s, depend on these cofactors, most inhibitors of JmjC proteins have turned out to block the interactions with one or both of the cofactors. However, this approach is a double-edged sword as the binding sites for these molecules are near facsimiles across all JmjC proteins. For example, JIB-04 chelates iron at the catalytic site and also disrupts binding with the histone substrate across the Jumonji demethylase superfamily albeit with some differences in efficiency amongst individual enzymes [213, 357]. In our work in breast cancer cells, MCF7 cells were 14-fold more sensitive to JIB-04 as an antiproliferative agent than were MDA-MB231 cells. *In vitro*, JIB-04 was about 8-fold more potent at inhibiting the activity of KDM5B than KDM5C [213]. Meanwhile, GSK-J1, which is competitive with α -ketoglutarate and inhibits both the KDM5 and KDM6 family [355, 356], is conversely about 8-fold more potent against KDM5C than KDM5B [213]. GSK-J4, the cell permeable ethyl ester derivative of GSK-J1, was similarly active against MCF7 and MDA-MB231 cells. Differential effects on cytoplasmic targets of the KDM enzymes may also play a role because JIB-04 has been recently shown to inhibit translation and to sensitize cells to the growth inhibitory effects of mechanistic target of rapamycin (mTOR) inhibitors, possibly through its effects on KDM4A or KDM5A, both of which were found to associate with polysomes [360].

Context, such as cancer type, is another major consideration when selecting inhibitors of KDM5 enzymes. We obtained inconsistent results between our two targeting shRNAs with our KDM5B knockdown NSCLC lines despite both shRNAs efficiently diminishing the expression of KDM5B protein, either one or both could have off target effects on other proteins not detected in the assays utilized. These cryptic targets might also affect invasion. Even so, previous work using these same constructs has demonstrated that suppressing KDM5B expression decreases cell proliferation in breast cancer lines with high levels of endogenous KDM5B but had little effect on cell proliferation in breast cancer lines that already expressed low levels of endogenous KDM5B [282]. Moreover, in melanoma lines, knockdown of KDM5B using the same shRNA sequences was shown to increase cell proliferation [361], as is consistent with the known roles of KDM5B in melanoma as an anti-proliferative transcriptional regulator and marker of a slow-cycling subpopulation of cells required for continuous tumor growth [255, 276]. Such differences in response highlight how the background of the cells under study influences the impact of KDM5B modulation and how important a comprehensive study of KDM5B function across multiple processes and cancer types is necessary before use of inhibitors in the clinical setting.

As more specific inhibitors have become available, we have had the opportunity to try these compounds as inhibitors of invasive behaviors via KDM5B inhibition. KDM5-C49 and its cell permeable derivative KDM5-C70 blocks the α -ketoglutarate binding site in KDM5B (as well as in other KDM5s and to a lesser extent in KDM4s and KDM6s) [212]. In our work, treatment with 5 μ M of KDM5-C70 for over 72 hours failed to change how H1299 parental, leader, or follower cells invade in a 3-D assay. However, the limited published work on the effects of KDM5-C70 largely characterize its antiproliferative activity. KDM5-C70 reduced viability of multiple myeloma cells by 50% at about 20 μ M over 7 days [212]. In several breast cancer lines KDM5-C70 inhibited colony formation in several breast cancer cell lines by more than 70% when treated with 5 μ M for at least 7 days [343].

Yet, KDM5-C70 affected colony formation little in cell lines previously shown to be resistant to KDM5B knockdown [343] and did not induce cell death in HeLa cells at up to 10 μM after 24 hours [358]. These studies, including our own, vary in three major ways—the endpoint measured, the background of the cells in use, and the treatment regime. All three points must be addressed in preclinical stages of drug development before understanding how to test the efficacy of a compound in humans. The work depicted in this chapter plays a small part in gathering this information for our newest compounds.

GDC-50, Dong-A-167, and CPI-48 were developed in the last three years and have been used in even fewer published studies than KDM5-C70. GDC-50, Dong-A-167, and CPI-48 also compete with α -ketoglutarate at the active site of KDM5 enzymes [362-364]. Wu et al. 2018 found that at 10 μM all three compounds increased H3K4me3 levels after 3 days of treatment and activated interferon induced genes after 6 days of treatment [362]. GDC-50 is notable for its specificity for KDM5 proteins with an IC_{50} of 45, 56, and 55 nM for KDM5A, -B, and -C respectively whereas the strongest detected IC_{50} for other KDM enzymes tested (1A, 2B, 3B, 4C, 6A, 7B) was 4.1 μM for KDM4C [363]. Moreover, in breast cancer cells, GDC-50 induced higher levels of H3K4me3 but not H3K9me3 or H3K27me3, substrates for the KDM4 family and the KDM6 family, respectively [364]. Additionally, GDC-50 inhibited colony formation in MCF7 breast cancer cells at 10 μM over 12 days [364]. In the one functional study of CPI-48, Liu et al. 2019 found that treatment with 1 μM CPI-48 over 8 days in mouse melanoma cells greatly induced H3K4me3 levels and shifted cells to a more BRAF inhibitor-sensitive CD34+ state [361]. This newest generation of demethylase inhibitors show great promise in a variety of settings.

As the search for specificity at the molecular level for inhibition of KDM5 enzymes, either targeting the whole family or each member on its own, KDM5-C70, GDC-50, Dong-A-167, CPI-48, and their successors must have well defined protocols. These protocols must include precision for

use in specific cancer backgrounds with precise treatment courses and doses. KDM5B remains a promising target for repression of invasion and metastasis in multiple cancer types, but further studies both on the function of KDM5B and on the activities of inhibitor compounds are necessary before we can attain this goal.

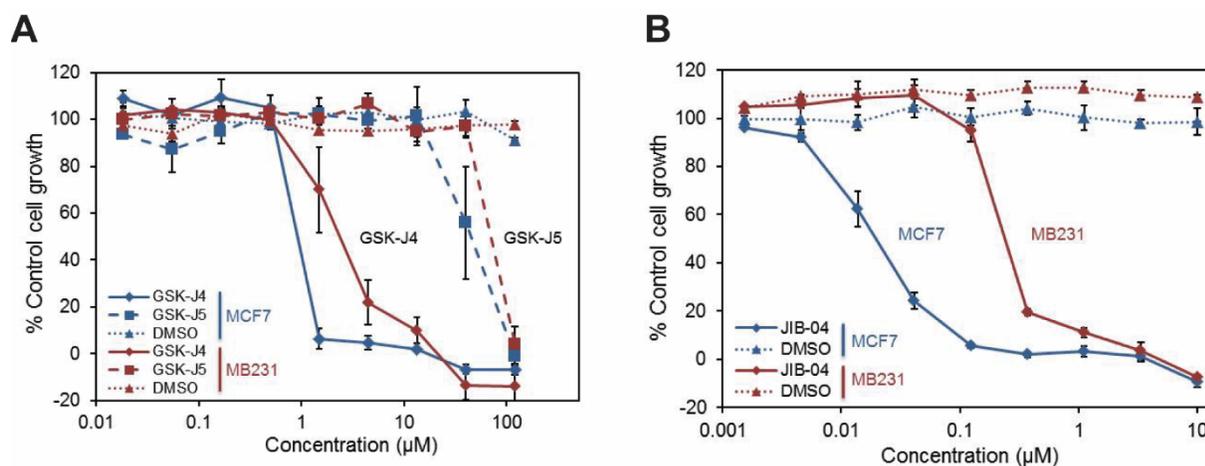


Figure 3.1 GSK-J4 and JIB-04 inhibit of breast cancer cell. Cell growth-inhibitory effects of GSK-J4, -J5 (A), and JIB-04 (B) in human breast cancer cell lines. GSK-J5 is a less active isomer of GSK-J4 [355] and is used as a negative control. MDA-MB231 (*red*) or MCF7 (*blue*) cells were treated with the indicated concentrations of JIB-04 or GSK-J4 (*solid line*) or GSK-J5 (*dashed line*) or vehicle equivalent (DMSO; *dotted line*), and the percentage of untreated control cell growth was determined after 72 hours by the sulforhodamine B assay. Data represent the means \pm s.e. (*error bars*) of three or four independent experiments performed in triplicate.

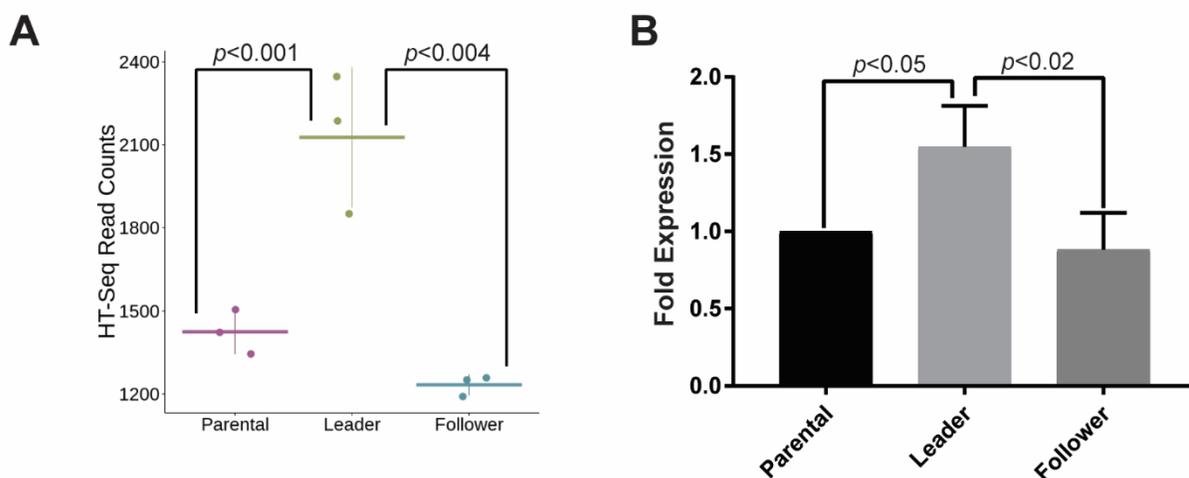


Figure 3.2 KDM5B is overexpressed in H1299 leader cells. (A) KDM5B transcription levels in parental, leader, and follower cells by RNA-seq. Parental cell and follower cell populations were derived from three separate isolations, and leader cell populations were derived from two separate isolations with two passages of one isolation being sequenced. HT-Seq Read counts are plotted for each experiment (dots) with mean count for each subpopulation represented by a horizontal bar (\pm s.d., vertical line). *P*-values were calculated by one-way ANOVA with Tukey's post-test. (B) KDM5B expression levels by quantitative reverse transcriptase PCR of RNA isolated from parental, leader, and follower cells. KDM5B expression levels were normalized to β -tubulin mRNA. (mean \pm s.d., $N=3$ independent experiments, 3 RNA isolations of three passages of parental, leader, and follower cells). *P*-values were calculated by one-way ANOVA with Tukey's post-test.

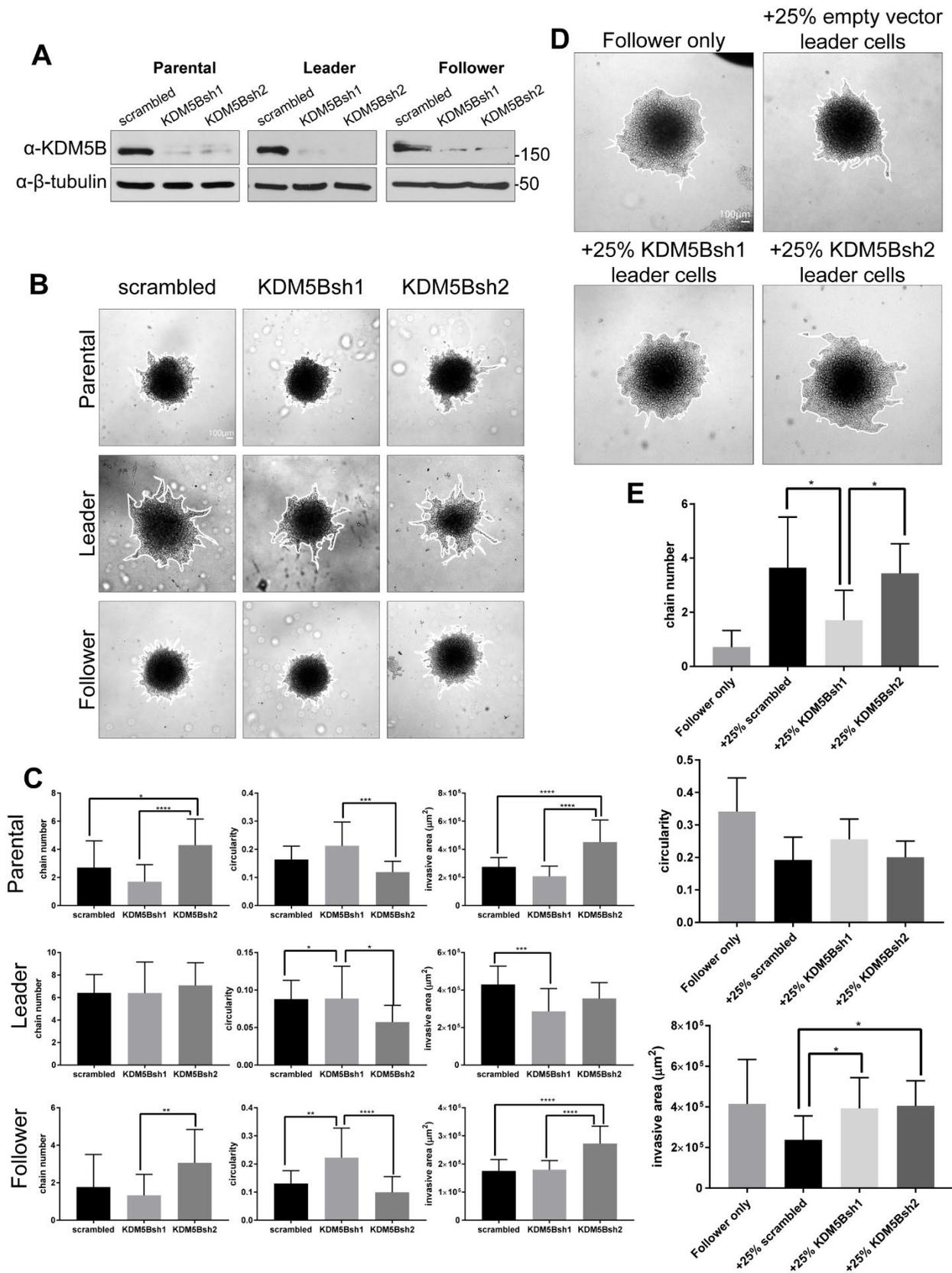


Figure 3.3 Knockdown of KDM5B produces varying results between shRNAs of different targeting sequences. (A) Western blot analysis demonstrating successful knockdown of KDM5B in parental, leader, and follower cell populations. (B) Images of invasion in Matrigel at 24 hours of parental, leader, and follower spheroids expressing scrambled shRNA, KDM5Bsh1, or KDM5Bsh2. (C) Quantification of invasive area, circularity, and chain number from spheroids depicted in (B) (mean \pm s.d., $n=13-18$ spheroids across $N=3$ experiments, $*p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$ by one-way ANOVA with Tukey's post-test). (D) Images of invasion in Matrigel at 24 hours of spheroids composed of 75% follower cells and 25% leader cells expressing either shRNA, KDM5Bsh1, or KDM5Bsh2. (E) Quantification of invasive area, circularity, and chain number from spheroids depicted in (D) (mean \pm s.d., $n=14-17$ spheroids across $N=3$ experiments, $*p<0.05$ by one-way ANOVA with Tukey's post-test).

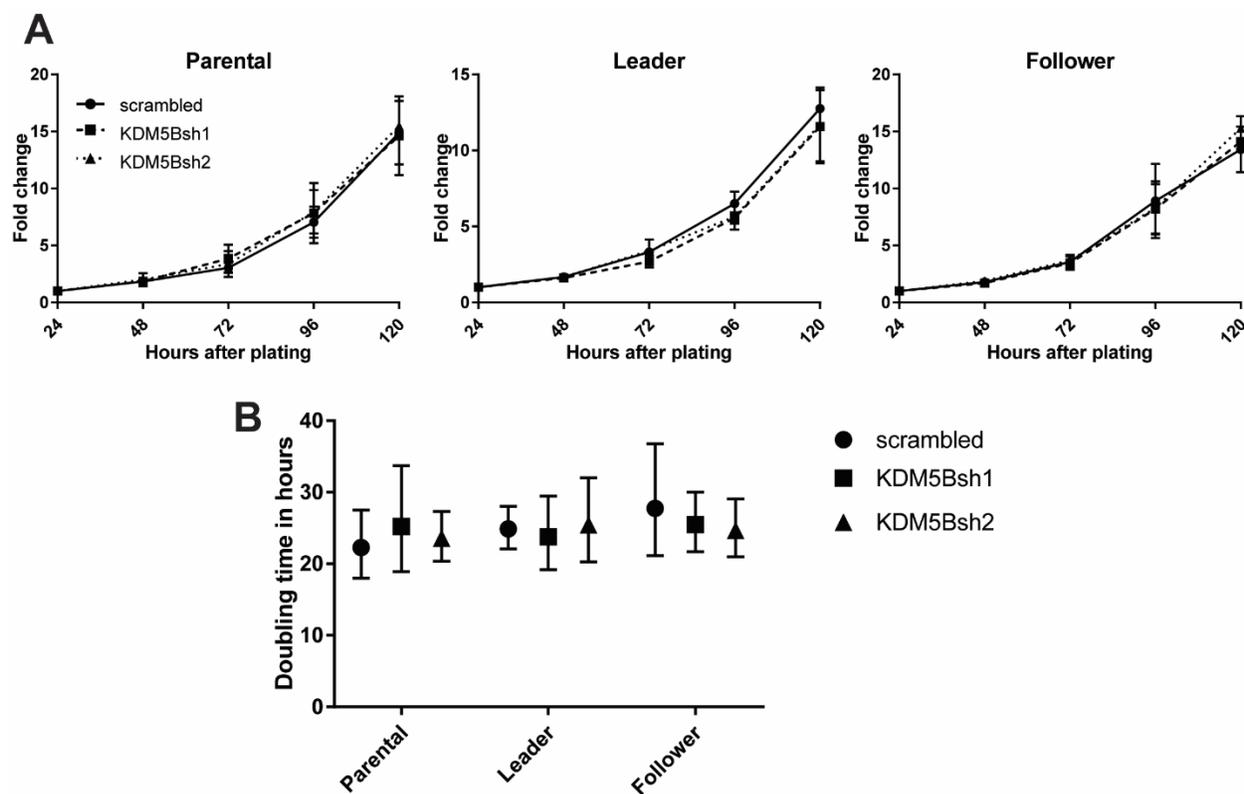


Figure 3.4 Knockdown of KDM5B does not alter growth rate. (A) Growth rate of parental, leader, and follower H1299 lines stably expression scrambled shRNA, KDM5Bsh1, or KDM5Bsh2 (mean \pm s.d. of $N=3$ independent experiments at each time point). (B) Doubling time $\pm 95\%$ confidence interval for lines in (A) as calculated through the least squares fit of exponential growth of data in (A) in GraphPad Prism where doubling time equals $\ln(2)/\text{rate constant}$.

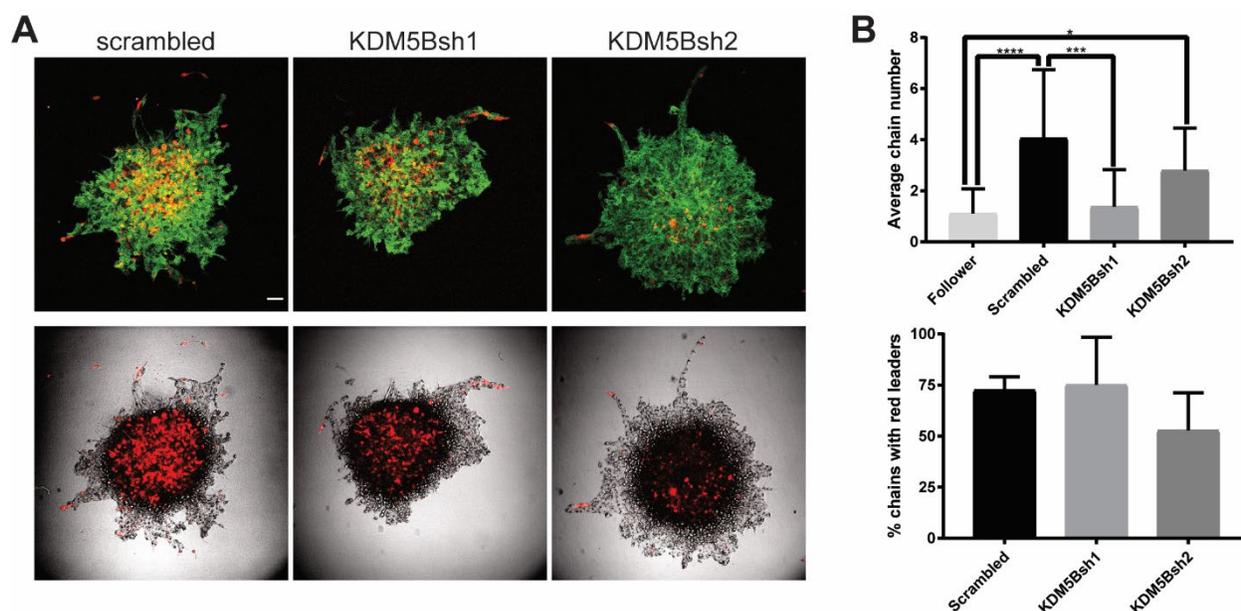


Figure 3.5 Knockdown of KDM5B does not change the percentage of chains led by introduced leader cells. (A) H1299 leader cells with stably expressing either scrambled shRNA, KDM5Bsh1, or KDM5Bsh2 and mCherry (red) were mixed at 25% with unmodified followers (remaining 75%) still stably expressing Dendra2 (green). Representative images at 24 hours invasion in Matrigel shown in (A). (B) Quantification of the number of chains (top) and percent of these chains positive for leader cells expressing mCherry. (bottom). Chain number is calculated by averaging number of chains per spheroid from $n=15-18$ spheroids across $N=3$ independent experiments. Percent of chains with red leader cells is the average of percents calculated from $N=3$ independent experiments.

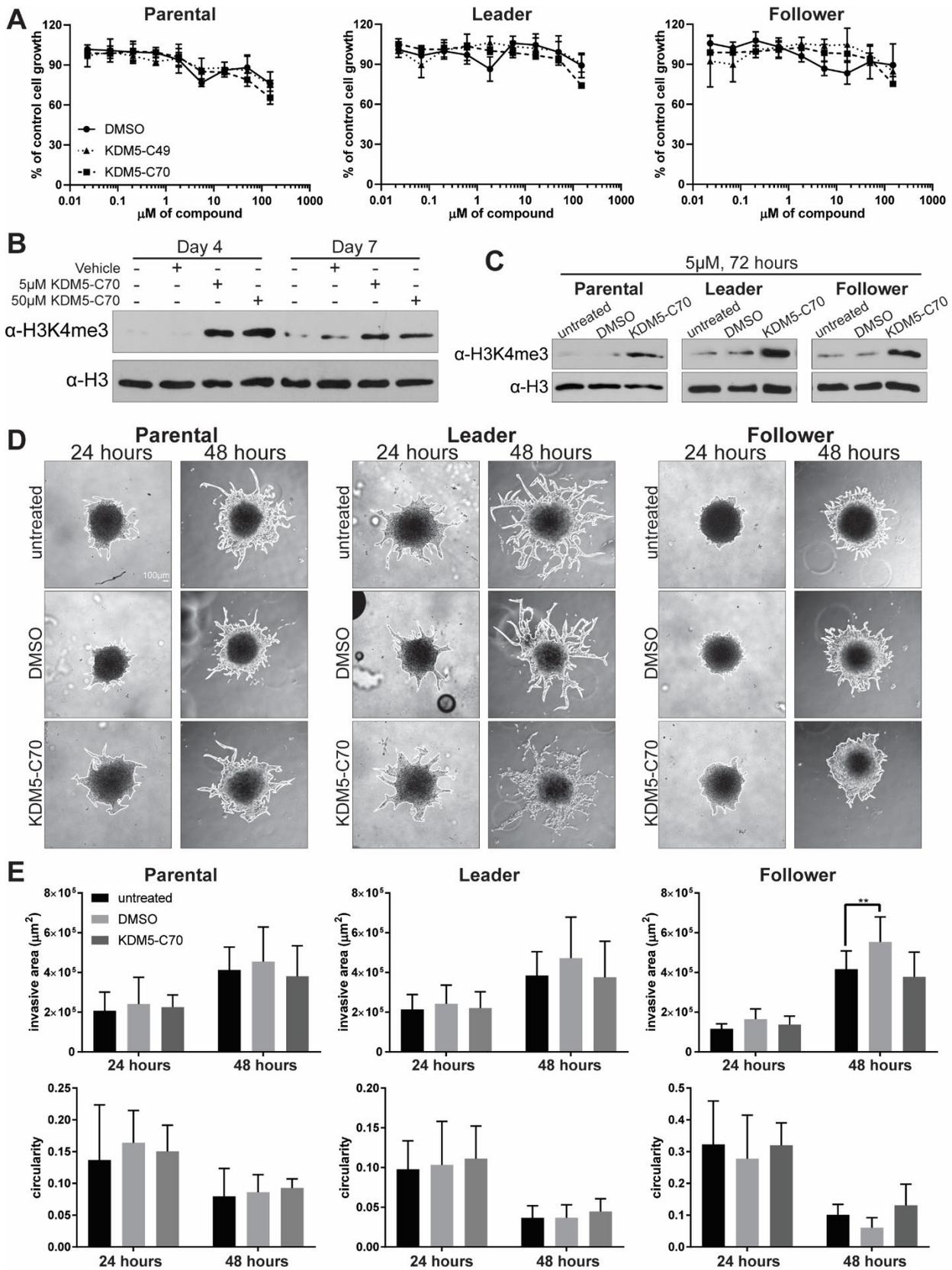


Figure 3.6 Treatment with KDM5 inhibitor KDM5-C70 fails to change invasive behavior. (A) Growth inhibitory effects of KDM5-C49 and KDM5-C70. Data represents mean \pm s.d. of $N=3$ independent experiments. (B) Western blot analysis in H1299 parental cells of H3K4me3 after four or seven days of treatment with either 5 or 50 μ M KDM5-C70 or vehicle equivalent to 50 μ M treatment. (C) Western blot analysis in H1299 parental, leader, and follower cells of H3K4me3 after 72 hours of treatment with 5 μ M of KDM5-C70. (D) Images of invasion in Matrigel 24 and 48 hours after embedding of parental, leader, and follower spheroids treated with 5 μ M KDM5-C70. (E) Quantification of invasive area and circularity from spheroids depicted in (D) (mean \pm s.d., $n=9$ spheroids in $N=3$ independent experiment, $**p<0.01$ by one-way ANOVA with Tukey's post-test).

Figure 3.7 Recently developed KDM5 inhibitors suggest invasion can be inhibited chemically.

(A) Growth inhibitory effects of GDC-50, Dong-A-167, and CPI-48. Data represents mean \pm s.d. of $N=3$ independent experiments. (B) Western blot analysis of H3K4me3 after 72-hour treatment with 5 μ M GDC-50, Dong-A-167, and CPI-48. (C) Images of invasion in Matrigel 24 and 48 hours after embedding of parental, leader, and follower spheroids treated with 5 μ M GDC-50, Dong-A-167, and CPI-48 for 72 hours during spheroid formation and during subsequent invasion after embedding in Matrigel. (D) Quantification of invasive area and circularity from spheroids depicted in (C) (mean \pm s.d., $n=3-6$ spheroids in one experiment, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$ by one-way ANOVA with Tukey's post-test).

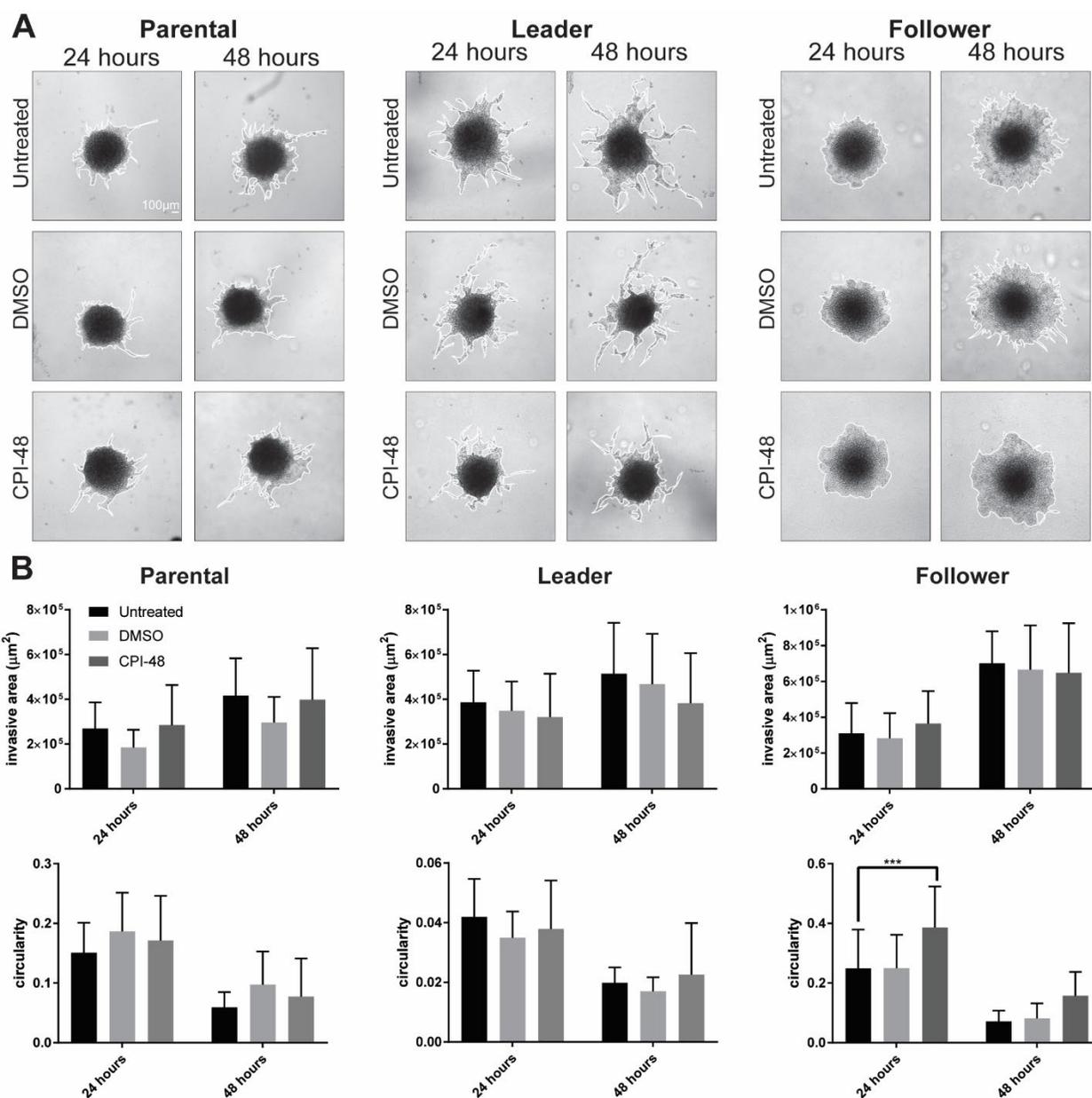


Figure 3.8 CPI-48 does not significantly alter invasion. (A) Images of invasion in Matrigel 24 and 48 hours after embedding of parental, leader, and follower spheroids treated with 5 μM CPI-48 for 72 hours during spheroid formation and during subsequent invasion after embedding in Matrigel. (B) Quantification of invasive area and circularity from spheroids depicted in (A) (mean \pm s.d., $n=3-6$ spheroids in one experiment, $*p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$ by one-way ANOVA with Tukey's post-test).

Chapter 4

Discussion

As metastatic cancer is a looming threat to lifespan and quality of life of all people, much scientific research is focused on abolishing this disease. Fundamental to accomplishing this goal is understanding how cancer cells leave the primary tumor so we may prevent their exodus and eventual colonization at distant, vital organs. The work presented in this dissertation focuses on defining the underlying differences between types of invasive cells and how targeting this difference could ultimately be an approach for metastasis prevention.

4.1 KDM5B as a mediator of collective invasion

4.1.1 Summary of findings

Increasingly, scientific evidence is indicating collective invasion, in which cells move as a cooperative, coordinated, and cohesive group, as a principal contributor to metastasis [82]. Early on in the history of metastasis research, animal models proved that clusters of cells are able to survive the passage through the vessels of the lung and are more likely to produce metastases than single cell when injected in the tail vein [83, 84]. From cancer mouse models to sequencing of human tumors, numerous studies have shown that metastatic lesions are often multiclonal, demonstrating an origin from more than one cancer cell [85, 87, 88, 349]. One study in a mouse breast cancer model even estimated that more than 97% of metastases were seeded by tumor cell clusters [88]. Important to collective invasion is genetic heterogeneity within tumors, both for detecting the presence of multiclonal metastases and for the efficiency of the movement of the group of invading cancer cells. Multiregional sequencing of primary and metastatic tumors has revealed that a significant proportion of mutations occur in distinct subclonal populations [4-7, 9, 14, 17]. When exposed to a new environmental challenge, as occurs during invasion and metastasis, the unique qualities of one or more subclones can permit these cells to thrive where the bulk of the tumor cells do not [29, 31-33]. Even

within a group of cells undergoing collective invasion, there may be heterogeneity essential to the function of the group as a cooperative.

Leader cells and follower cells are two functionally and phenotypically distinct groups which occur within some collectively invading groups, such as branches or chains. Leader cells occur at the front of the pack and lead the pack along the invasive edge, and follower cells invade behind leader cells, attached by cell-cell junctions [73, 74, 89]. Recent work by Konen et al. 2017 has permitted, for the first time, the study of leader and follower cells in isolation. By fluorescently labelling cells with a photoconvertible dye, leaders and followers could be selected by photoconversion of the dye in cells in either the leading or following position of an invasive chain and isolated by fluorescence-activated cell sorting [288]. Expansion of these populations enabled close observation of cell phenotypes and dissection of cell genotypes. Leaders grew slower, invaded more, and secreted VEGFA, which was necessary for pack formation [288]. In contrast, followers were much less invasive but proliferated at a much quicker rate [288].

By performing RNA-seq on leader, follower, and parental (the population from which leaders and followers were derived) cells, several mutations exclusive to either leader cells or follower cells emerged—i.e. each subpopulation had a specific mutational profile. Two likely gene candidates were identified based on their established functions: *ACTR3*, coding for actin-related protein 3, and *KDM5B*, lysine demethylase 5B, both of which are known mediators of invasion and migration. Our goal in Chapter 2 was to characterize the role of these candidate mutations in the specialized functions of leader and follower cells. While actin-related protein 3 (ARP3) and its leader specific mutation K240R are extensively covered in Chapter 2, the main focus of this dissertation as a whole is the role of lysine demethylase 5B (KDM5B) in collective invasion. As such, this discussion will also center on KDM5B and its follower specific mutation at L685W.

Sanger sequencing and TOPO-TA cloning (to confirm the relative proportions of each mutation in each population) validated the presence of *ACTR3* mutations in leader cells and *KDM5B* mutations. Sequencing of these clones revealed that *ACTR3* was mutated in about 30% of leader-derived clones, with almost no *ACTR3* mutations apparent in parental and follower population. *KDM5B* was mutated in approximately 30% of parental- and follower-derived clones.

To further understand how these mutations might be affecting leader and follower behavior respectively, we assessed the established roles of the wild-type proteins and studied the nature of these alterations. ARP3 is a key coordinator of lamellipodia protrusion during cell migration and is correlated with invasion, metastasis, and/or poor survival in gastric, colorectal, liver, and gallbladder cancers [315-319]. Residue K240 was found to be a target for ubiquitination, and mutation to arginine would likely obliterate any associated signaling with ubiquitination at this site, including protein degradation. Indeed, follower cells expressing ARP3 K240R were able to take on the role of leader cells, and overexpression of ARP3 mimicked the behavior of ARP3 K240R expressing lines, indicating that ARP3 K240R is likely a gain-of-function mutation.

KDM5B is a member of the lysine demethylase 5 family and as such, enzymatically removes di- and tri-methylation of histone 3 lysine 4 (H3K4me2 or H3K4me3 respectively) [202, 204]. H3K4me2 and H3K4me3 are most often indicative of a permissive state for transcription, and KDM5B is thus considered in most cases to be a repressor of transcription [208]. Furthermore, KDM5B is strongly and positively linked to invasion and metastasis in lung, stomach, liver, and breast cancer [259, 268-273]. Of note in these reports, the role of KDM5B in regulation of invasion was established largely in 2-D assays (such as scratch wound healing or transwell invasion/migration assays), in human samples of metastatic sites, or from just a few mouse models using tail vein injections. This study is the exception by examining how KDM5B affects collective invasion in a 3-D model.

In follower cells, but not leader cells, KDM5B is mutated at leucine 685 to tryptophan (L685W) in one third of reads. This change in bases substitutes a small, non-polar amino acid for a bulky one, which has the potential to disrupt KDM5B secondary structure and interactions with other proteins, including the histone substrate. Additionally, L685 is located just N-terminal of the C₅HC₂ zinc finger domain of KDM5B, which is indispensable for the demethylase activity of KDM5B both in *in vitro* studies and in cells [204, 213] suggesting that L685W may be a disruptive alteration. By modeling the known crystal structure of KDM5B with amino acids 1-14 of histone 3, we see that the histone tail travels very near L685W, and when binding a full-length histone 3, L685W could disrupt any further binding of substrate in this area. Fascinatingly, the end product of KDM5B demethylation, H3K4me₁, was present at higher levels in leader cells, which express the highest levels KDM5B and only wild-type KDM5B, as was consistent with the hypothesis that KDM5B L685W decreased the demethylase activity of KDM5B. However, all members of the KDM5 family demethylate histone 3 di- and tri-methylation [203-206]. In testing the *in vitro* enzymatic activity of KDM5B wild-type and KDM5B, we found, as a trend, that overexpression of wild-type KDM5B increased the cellular H3K4me₃ demethylase activity over empty vector control cells and overexpression of KDM5B L685W H3K4me₃ decrease demethylase activity slightly as compared to control cells. Thus, our data point to KDM5B being essential for the H3K4 demethylation activity of leader cells and suggest that the L685W mutation of KDM5B in followers likely decreases this activity and is a loss-of-function mutation.

In testing this hypothesis, we surprisingly found that spheroids from leader cells overexpressing wild-type KDM5B leader were much more circular and had significantly fewer chains than control leader spheroids. Given our original hypothesis that leaders require wild-type KDM5B to be leaders and reports from previous publications that KDM5B promotes invasion and migration in 2-D assays [259, 268-272], we had expected our leader cells which expressed even more KDM5B

to have an even more extreme leader phenotype, in which spheroids have large numbers of invasive chains and very irregular invasive borders. Moreover, when leaders overexpressed KDM5B L685W, these cells produced spheroids much like the control cells but with slightly more invasive chains on average. These results introduce the idea that heterogeneity is fundamental to creating an invasive chain. If cells all express a very high level of wild-type KDM5B, these cells may be so far tipped towards a leader phenotype that chains will not form without some cells taking on the functional role of followers.

Correspondingly, overexpression of wild-type KDM5B or KDM5B L685W in follower cells had almost no impact on how spheroids generated by these lines invaded. These results fit with our model that leaders and followers have distinct genotypes as the first half of Chapter 2 details how ARP3 K240R makes leaders from followers. Without the activating ARP3 K240R mutation, an increase in wild-type KDM5B expression is not sufficient to convert a follower to a leader phenotype. In spheroids mixed with unmodified followers and made to ascertain if these “tipped” leaders could still lead chains, our wild-type KDM5B overexpressing leaders still failed to create chains where as KDM5B L685W leaders built even more chains than control leaders. Perhaps because even at 25% of the population, the KDM5B L685W leader cells were able to better establish the unknown multiplex signaling pathway necessary for chain-like invasion.

Targeted deep resequencing and SNP analysis of KDM5B across parental, leader, and follower populations determined that the 30% of reads positive for KDM5B L685W in follower cells represented the entire population (one allele in 100% of cells within a triploid population of followers) and that KDM5B L685W was actively selected for within the follower population. Fortuitously, a common SNP was located within the same exon as KDM5B L685W and was heterozygous in the H1299 cell line. Sequencing around this exon revealed three genotypes for KDM5B: WT/snpA, WT/snpG, and MUT/snpG. Given that in each population the proportion of genotypes fit into ratios

based on thirds, our H1299 lines are likely triploid for KDM5B and/or chromosome 1, and so the 30% of follower reads positive for KDM5B L685W represents one allele present in all triploid follower cells. This result supports our conclusion from the previous 3-D invasion assays that follower cells do play a necessary role in creating invasive chains. Followers are a unique subpopulation within the parental population containing KDM5B L685W, and leaders are another rare distinct population containing ARP3 K240R.

While our initial supposition was that increasing expression of KDM5B in followers would induce these cells to form spheroids with more invasive chains, like the high KDM5B-expressing leader cells, we instead found that changes in KDM5B expression had little impact on how the follower cells invaded. We had also conjectured that KDM5B L685W leaders might be more circular and have fewer chains appearing much like pure follower spheroids, but overexpressing wild-type KDM5B decreased the ability of leaders to invade collectively in chains. Even so, our hypothesis that KDM5B L685W distinguishes follower cells from leader cells in a functionally significant way can still be accommodated.

An invasive chain inherently requires a mix of at least two cell types: those that lead and those that follow. The isolation process of each population enriches for cells of either a leader or follower phenotype and thus, enriches for the cells of the respective genotype or expression profile. Yet, even in leader purified populations, there is likely some variation in expression of key proteins, including KDM5B. More chains form in a leader purified population because more cells act as traditional leaders by directing attached groups of cells out of the spheroid [288], but chains may only be able to form because some of the purified leaders express a profile just a bit less like those at the end of these chains and may serve the role of follower. Increasing the expression of wild-type KDM5B in the already highly KDM5B-expressing leader cells seems to tip the balance of leader and follower phenotypes even further towards the expression profile of leader cells, thus losing the heterogeneity so

fundamental in building an invasive chain. This heterogeneity is maintained when KDM5B L685W, a probable loss of function mutation, is expressed in the leader cells.

4.1.2 Models of heterogeneity and cooperativity in invasion

Earlier studies of H1299 leaders and followers established that leader cells and follower cells have unique physiologies when cultured in isolation [288]. Leader cells developed highly invasive spheroids with many branches but divided slowly, often with mitotic defects. In contrast, follower cells grow rapidly but invade poorly [288]. When leaders were reintroduced to followers even at low levels, invasive chains formed with labelled leader cells predominantly at the forefront. Moreover, leader conditioned media increased the motility of follower cells [288]. Additionally, follower cell conditioned media alone prompted growth of leader cells [288]. Subsequent transcriptome profiling revealed that VEGF secretion by leader cell drove one half of this symbiotic relationship in which leader cells promote the movement of follower cells [288]. The pro-survival factors secreted by follower cells to promote leader cell growth are yet to be determined for the other half of this relationship [288]. Our work adds to the leader-follower model a genetic basis for the differences in cell behavior both together and separately and highlights the importance of heterogeneity as the underpinning to chain-like invasion in H1299 cells. This model of heterogeneity as a crucial component to invasive behaviors is not without precedence. Several studies have found that the presence of at least two populations are required to elicit specific invasive behaviors of one or both populations. While each study identifies a slightly different mechanism by which cells cooperate, these mechanisms may give us insight into how leader and followers are cooperating.

In our work, we cross-introduced SaGA-identified candidate mutations to leaders and followers to determine how these genes controlled collective invasion. An entire cadre of genes have previously been established as regulators of invasion and metastasis, and heterogeneous expression of

these mutations within a cell population may also regulate collective invasion. Tsuji et al. 2008 created two fluorescent cell lines from hamster cheek pouch carcinoma-1 cells; one induced to undergo EMT by expressing p12 and a green fluorescent protein and a control non-EMT induced line [336]. Xenografts of each line separately revealed that the p12 EMT cells were able to invade locally and were detected in the blood stream [336]. When injected into the tail vein, the non-EMT cells readily formed metastatic lesions, but the p12 EMT cells failed to form any metastases [336]. However, subcutaneous injection of a 1:1 mixture of p12 EMT cells and non-EMT control cells formed primary tumors with locally invasive fronts composed of both cell types and permitted intravasation of both cell types (i.e. both cell types were detected in the blood stream) but resulted in lung metastases composed only of non-EMT control cells [336]. In this model, both cells that had undergone EMT and those that had not (non-EMT) were derived from the same cell line and work together during the metastatic process.

As discussed in Chapter 1, heterogeneity is nearly ubiquitous in cancer cell populations. Our work establishes that inherently present heterogeneity is required for invasive behaviors. Given the frequency of genetic and epigenetic heterogeneity and invasion or metastasis across cancers, dependence of collective invasion on the presence of multiple subpopulations likely occurs in other cancerous milieus. Calbo et al. 2011 used a mouse model of small cell lung cancer (SCLC) to create cell lines, some of which exhibited phenotypically distinct subpopulations—a neuroendocrine (NE) type that grew as floating aggregates of small cells and expressed a neuroendocrine profile (high Ash1, Syp, and NCAM) and a non-neuroendocrine (non-NE) type that grew as large cells in a monolayer and expressed a profile absent of neuroendocrine markers but positive for mesenchymal markers, including vimentin, and CD44 [334]. After clonally isolating each cell type, admixture in 2-D culture revealed that the NE cells now attached to and spread along the larger non-NE cells as an adherent, multilayer culture [334]. Most interesting when injected subcutaneously into mice at a 4:1 ratio, the

admixture of NE and non-NE cells, despite an intermediate level of primary tumor growth, formed extensive metastasis to the liver whereas no metastasis was observed with injection of either single clone alone [334]. Moreover, the NE and non-NE cells had to be in close contact to cause metastases as injection of single clones in opposite flanks abrogated this effect [334]. However, the metastases appeared to be only of the NE-differentiated type as they lacked cells of the non-NE phenotype [334]. Array-CGH analysis demonstrated a shared background of genetic aberrations within NE and non-NE cells indicating a common clonal heritage but also revealed some significant differences between cell line pairs, in particular, amplification of *Myc11* repeated across NE cell lines [334]. The similar clonal origin of these genetically divergent, specialized populations is remarkably similar to the case of our H1299 leaders and followers. However, Calbo et al. 2011 did not explore the genetic backgrounds of these subclones past large genetic changes in copy number. Further study identified changes in cytokine secretion across NE and non-NE cells. NE conditioned media contained almost no cytokines, non-NE conditioned media had high levels of several cytokine, and mixture of NE and non-NE cells secreted IGFBP-3 and IGFBP-5 de novo [334]. This study demonstrated that several cell behaviors, including signaling and metastasis, are contingent on the presence of multiple subclonal populations in SCLC, analogous to our work with chain-like invasion in NSCLC.

Similarly, Mateo et al. 2014 discovered that clonal subpopulations generated from prostate cancer cell line PC-3 produced divergent phenotypes: a cancer stem cell (CSC) enriched, epithelial-like population called M cells and a CSC depleted, mesenchymal-like population called S cells [335]. Epithelial M cells readily formed spheroids (a sign of self-renewal potential), but mesenchymal S cells invaded at a higher rate through a transwell [335]. However, co-culture of M and S cells or culture of M cells in S cell conditioned media increased transwell invasion by M cells [335]. Mass spectrometry of conditioned media from both sets of cells revealed that the most overrepresented diffusible factor from the S conditioned media was the matricellular protein SPARC, and depletion of SPARC

abolished the effects of S conditioned media on M cells [335]. Similarly, co-implantation of M and S cells augmented tumor growth and M cell metastatic dissemination, which was lost when SPARC was depleted [335]. In addition to reinforcing the importance of heterogeneity in cooperative invasion, this study identifies a new molecule by which two cell types communicate to conduct cooperative migration. This study substantiates the importance of two subpopulations directing the behavior of a group of invading cells like our leaders and followers while providing a novel model for this cooperation.

Although we studied inherently occurring subpopulations, controlled introduction of heterogeneity can add a measure of precision in studying changes in cell behavior within interaction subpopulations. By assembling mosaic microtissues in 3-D culture of the breast epithelial line MCF10A, Liu et al. 2012 found that aggregates composed of one cell expressing constitutively active H-Ras^{V12} and several wild-type cells commonly produced single cell basal extrusions of the H-Ras^{V12} cell or motile multicellular protrusions led by the H-Ras^{V12} cell [337]. Such extrusions and protrusions were rare in homogeneous aggregates composed of only wild-type or only H-Ras^{V12} cells [337]. For the motile, multicellular protrusions to form, a rare cell with a unique characteristic, high H-Ras activity, had to be present but only amongst cells of lower H-Ras activity [337]. Our H1299 cells exhibit similar behavior in that leaders are a rare population within the follower-enriched parental population and that pushing leaders too far to the leader phenotype with overexpression of KDM5B (creating a population of cells universally expressing very high levels of KDM5B) stops chain-like invasion.

Important to our comprehension of the interaction between two subpopulations working together is to uncover the direct mechanism by which they modify the behavior of their invasive partners. Within a zebra fish melanoma model, Chapman et al. 2014 found that co-culturing or xenografting together a highly invasive melanoma line with a poorly invasive line caused the highly

invasive line to switch from a protease-independent mode of invasion to one requiring the MT1-MMP protease and to alter the extracellular matrix (ECM) by increasing type 1 collagen and fibronectin deposition [350]. Additionally, in xenografts, the poorly invasive melanoma cells greatly increased their invasion in the presence of the highly invasive cells, and both cell types were present at the invasive front [350]. By culturing the populations separately on either side of a transwell system, diffusible factors were identified as the cause of changes in the ECM [350]. Although the exact factors were not pinpointed, establishing a direct mechanism of altering invasion by remodeling the ECM makes this study unique among those reviewed here. Future studies of candidate mutations in our NSCLC leader and follower cell model would greatly benefit by establishing a direct mechanism as Chapman et al have achieved.

Taken together, the above studies provide evidence that even small amounts of heterogeneity in a population of cells can act synergistically to create cell behaviors not capable by either singular subclone. These studies are similar in their use of a single cell line, except for Chapman et al. 2014 who utilized two existing cell lines of the same cancer type, to demonstrate that two subpopulations with unique phenotypes act together to alter how the group as a whole invades the surrounding matrix or tissue [334-337, 350]. These studies also parallel our work as we too found subpopulations existing within an established cell line and that these subpopulations act synergistically. However, where the aforementioned reports focus on a mechanism through which cells cooperate, our work seeks to understand the basis of the differences. Moreover, we began with an intrinsically heterogeneous cell line and isolated cells during a naturally occurring process in which these cells demonstrated specialized roles (as opposed to engineering them by reintroducing single clones derived from a common cell line and selected for their varying phenotypes in 2-D culture [334, 335] or by altering expression of a specific protein in some of the cells [336, 337]) which is more likely to reflect the inherent biology of tumor cells. Having found that the mutation ARP3 K240R distinguishes leader

cells from followers and KDM5B L685W followers from leaders, we can now test the mechanism by which these populations contribute to cooperative invasion. The heterogeneity that exists even in populations of cells cultured from the same cell line or that has been introduced by a change in a single gene can be narrow but have great effects on cooperativity between subpopulations invading surrounding tissues (Fig. 4.1).

4.1.3 Future considerations

A clear limitation of our work is that we studied how naturally occurring mutations contribute to cooperative invasion in a single cell line. Since several studies utilizing different cell lines across multiple cancer types have shown that cooperation of subclonal populations aids invasive phenotypes, support for this phenomenon as more than a one-off is not lacking. Our study identifies how alterations in genetic background of naturally occurring co-existing population account for how these subpopulations can inherently activate cooperative invasion. We prove simply that cooperative invasion between two genetically distinct, yet related cell populations occurs naturally within a cell line and that this cooperation is attributable their heterogeneity. Should we focus more on which specific mutational drivers create leader and follower cells on a scale applicable to translation, we would need to perform a similar analysis across a much larger set of established cell lines or from patient samples. For example, we could use a panel of cell lines in which chain-like invasion occurs, subject them to SaGA and analyze the genetic differences between the leader and follower subpopulations. If a pattern of alterations in proteins or pathways is apparent, we could identify novel regulators of cooperative invasion. Indeed, although Calbo et al. 2011 did not focus on the genetic differences between subpopulations, they did find a consistent amplification of *Myc11* in NE cells derived from multiple tumors that was not found in matched non-NE cells. A more in-depth study at a finer scale could reveal more subtle but still significant genetic regulators of cooperative invasion [334].

We also found several additional candidate mutations that were follower or leader specific but only explored two. We began our work with the mutations most likely to affect invasion, but a profile as a whole may have more impact. We could use our leader and follower specific profiles as detailed in Table 2.1 to create leader or follower type cell lines in human bronchial epithelial cells. By converting these normal cells to possible leaders or followers, we could determine the relative contribution of each mutation to the phenotype. For example, if expressing three of the mutant proteins is required for the leader phenotype but expression of all the other mutant proteins are not, we could determine that only three of the mutations are driving the leader phenotype. Tsuji et al. (2008) and Liu et al. (2012) made cell subpopulations by overexpressing proteins of interest and found that mixing these populations altered invasive behavior of both groups [336, 337].

Previously, microarray transcriptome analysis comparing the transcripts of purified H1299 leader and follower populations uncovered vascular signaling as a major regulator of the H1299 leader and follower cell collective invasion [288]. Leader cells overexpress VEGF and Dll4 as compared to follower cells [288]. Concomitantly, followers overexpress the Notch receptor [288]. Modulation of this vascular pathway proved that these molecules regulate leader-follower collective invasion [288]. However, we have not connected our genetic alterations with this pathway or any other mechanism. The studies outlined above used a variety of methods to determine how signaling occurred across their two subpopulations. Calbo et al. 2011 initially subjected conditioned media collected from their NE cells alone, non-NE cells alone, or co-culture of both to a cytokine antibody array and found novel expression patterns from each condition [334]. In a follow-up study, Kwon et al. 2015 performed gene expression analysis via a microarray on NE cells that had been exposed to non-NE cell conditioned media to find that the ETS transcription factor *Pea3* became highly upregulated in the NE cells [365]. Mateo et al 2014 performed mass spectrometry on conditioned media from both M

cells and S cells on their own and identified SPARC as the most overrepresented diffusible factor S cells [335].

The analysis of conditioned media in these studies was a good strategy for identifying paracrine signaling between two cell types as a potential mediator of cooperative invasion. The cross-culturing of lines with conditioned media also helped to determine response to any paracrine signaling molecules. Initial studies on H1299 leaders and followers focused on expression differences between these lines when isolated and cultured separately, which is unique in that this data covers differences within the cells and not just in what factors are secreted [288]. However, this data only covers differences in steady state mRNA levels. Changes in location, activation of proteins, degradation, or modification of proteins are missing. Additionally, any synergistic effects are not measured as samples were not analyzed post-mixing the lines. To further understand how wild-type KDM5B and KDM5B L685W regulate collective invasion, mass spectrometry analysis of conditioned media from leaders and followers either overexpressing wild-type KDM5B or KDM5B L685W and from unmodified leaders, followers, and a co-culture of both should be conducted in hopes of finding a signaling network between leaders and followers, especially one that is KDM5B regulated. However, this approach would only work for a paracrine signaling network. To discover any juxtacrine signaling, in which cells must be in physical contact, perhaps gene expression analysis of mRNA following co-culture of fluorescently labelled leaders and followers after re-isolation by FACS might be the best option.

In ascertaining a mechanism of interaction in their cell lines, Chapman et al. 2014 chose a more targeted approach by predicting that matrix metalloproteinases were an important factor in invasive behaviors amongst their cell lines based on observations of the positions of each cell line in an invasive chain [350]. Their hypothesis was validated by differences in expression of proteases between cell lines and the effects of protease inhibitors on invasion in solo and co-cultures of the cell

lines [350]. By understanding the observed biology of our collective invasion patterns and carefully searching published literature on related subjects, we may to identify a candidate pathway for further investigation. In fact, KDM5B is known to suppress the transcription factor HOXA5 [265], which reduces VEGF [366], a signaling molecule overexpressed in H1299 leaders and required for collective chain formation [288]. By negatively regulating a negative regulator of VEGF, KDM5B overexpression could lead to overexpression of VEGF. If knockdown of HOXA5 mimics KDM5B overexpression or overexpression of HOXA5 abrogates the effects of KDM5B overexpression, we could connect the dots on the path of KDM5B to interaction between leaders and followers in a chain of cooperatively invading cells (Fig. 4.2).

The ultimate test for cooperative invasion is to employ an *in vitro* model of metastasis. Tsuji et al. 2008, Calbo et al. 2011, and Mateo et al. 2014 proved cooperativity in metastasis in cell lines through xenografts of each cell type individually and of a mix of both cell types [334-336]. Metastases were only possible with a combination of these lines and histology of locally invasive areas demonstrated a mix of both cell types [334-336]. Similarly, we can inject leaders and followers subcutaneously, together and separately, before measuring rate of metastases. Additionally, we can conduct this experiment while modulating KDM5B to understand how KDM5B regulates cooperativity.

4.2 KDM5B as a therapeutic target

4.2.1 Summary of findings

Metastatic disease is by far the most prolific killer following a patient's cancer diagnosis [1]. Nevertheless, development of therapeutics targeting metastases is lagging behind [367]. Current limitations include identifying targets at the right stage of the metastatic cascade, developing specific compounds, and finding the most effective time course for preclinical trials and eventually administration in patients [367]. One promising area for targeting metastasis is within epigenetics.

Epigenetic modifications as heritable yet reversible alterations of the genome and prime regulators of gene expression are key contributors to cancer cell plasticity [150, 156]. This plasticity permits cancer cells to respond quickly to changes in their environment as occur during drug treatment or invasion and metastasis [57, 157].

Development of compounds aimed at epigenetic enzymes has exploded in recent years. A particularly interesting set of epigenetic targets is the histone lysine demethylase family KDM5. KDM5 family members promote oncogenesis, drug resistance, and invasion and metastasis across many cancer types [208, 209, 219]. A major limitation for drug development in this field is creating specificity as members of the KDM5 family and related proteins containing the catalytic Jumonji domain have highly similar structures [287, 326, 358]. Even so, over the past decade considerable progress has been made in honing effective and specific molecules. We have had the opportunity to use a number of these molecules with varying specificities as anticancer agents in cell lines.

We were particularly interested in targeting KDM5B as it has the most proven connection to cancer formation and progression. Although upregulated in many cancer types, KDM5B was first identified as a gene overexpressed in breast cancer twenty years ago [254]. Since then KDM5B has also been shown to promote invasive behavior [249, 268] and chemoresistance [249, 275, 279, 352] in several cancer types, including breast cancer. Our earliest work with lysine demethylase inhibitors involved testing effects on cell proliferation in breast cancer cells after treatment with broad inhibitors JIB-04 and GSK-J4 [213]. MCF7 and MDA-MB231 were more sensitive to the growth inhibition effects of JIB-04 fold than to those of GSK-J4. Moreover, JIB-04 demonstrated much greater variance in activity between the two lines. JIB-04 inhibits histone binding of all members of the Jumonji demethylase superfamily with some variance in efficiency against individual enzymes [213, 357], but GSK-J4 inhibits just the KDM5 and KDM6 families [355, 356]. The relative sensitivities within cell lines to JIB-04 and between JIB-04 and GSK-J4 could be accounted for by differences in how well

each compound inhibits different Jumonji families and in the underlying biology of the cell lines, particularly since MCF7 cells are estrogen receptor positive and MDA-MB231 cells are not. Fascinatingly, KDM5A, KDM5B, and KDM5 inhibition were recently discovered to mediate anti-estrogen therapy in estrogen receptor positive breast cancer cells, including MCF7 cells [279].

KDM5B is also an interesting target in lung cancer as we found it to be upregulated by RNA-seq in a population of cells with a leader phenotype during collective invasion of H1299 NSCLC cells (Chapter 2). After confirming the upregulation of KDM5B with qPCR and Western blot analysis, we measured invasive capacity and behaviors in stable KDM5B knockdown lines of H1299 leader, follower, and parental isolates. Unfortunately, the most significant differences in invasion occurred between our two shRNA sequences targeting KDM5B. However, with one shRNA sequence, we did see a trend in decrease of chain-like invasion in the parental and leader cells as one would expect if KDM5B was a driving force in formation of and invasion by chains led by leader cells.

We next sought to inhibit the enzymatic activity of KDM5B to test the relationship between this enzyme and invasion. We first used compound KDM5-C70, one of the earliest compounds designed for high specificity of the KDM5 family [212, 284, 326, 343, 358]. After narrowing the therapeutic window through growth assays and histone methylation detection, we established a treatment protocol of 5 μ M KDM5-C70 over 72 hours before assaying collective invasion of a 3-D spheroid. However, despite an increase in H3K4me3 after KDM5-C70 treatment indicating that demethylase activity was inhibited, we noted no changes in invasive behavior across parental, leader, or follower lines. The possibilities remain that our treatment protocol was ineffective or that the off-target effects of KDM5-C70, which does inhibit all KDM5 proteins and KDM4 and KDM6 family members to some extent [212, 326], overwhelmed the effects on KDM5B. Additionally, a longer treatment period may be required to affect invasion.

Fortunately, we had access to the next generation of KDM5 inhibitors with even greater specificity. Developed in the last three years, GDC-50, Dong-A-167, and CPI-48 are included in few published studies. We tested these three compounds as we did with KDM5-C70 with growth assays and measuring histone methylation levels. As a result of these assays, we adopted a similar treatment protocol for initial studies: 5 μ M compound over 72 hours before measuring collective invasion. In our first experiment, we found that only CPI-48 increased circularity and thus may be decreasing chain like invasion. We conducted three more experiments to validate this data, but overall results failed to repeat. Perhaps the differences between the initial experiment and our repeats indicate that our treatment protocol needs a bit more tweaking to prevent invasion or that CPI-48 will not accomplish this task.

4.2.2 Future considerations

An obvious limitation to this study is that we were unable to chemically inhibit invasion by targeting KDM5B and fellow KDM5 family members, preventing our ability to fully understand the impact of KDM5B activity on invasion. A significant barrier is determining the proper treatment regimen. KDM5B as an epigenetic enzyme regulates gene expression. Even though H3K4me3 increased at 72 hours, the effects of modulating KDM5B may take longer as the target genes marked by H3K4me3 must restart transcription and exert their downstream effects on regulators of invasive behaviors. As well, the half-lives of our compounds are not established. A longer and/or repeated dosing schedule could yield more inhibition of invasion. Optimization of treatment protocol might be especially useful for CPI-48 as this compound was most likely to inhibit collective invasion.

Furthermore, specificity for KDM5 proteins remains an ongoing problem. Given that some compounds have varying degrees of affinity for other Jumonji proteins, the differences in efficacy may be attributable to the relative dependencies of a cell or phenotype on additional demethylases.

However, this complication may actually be beneficial. This effect could be abrogated by targeting all KDM5 proteins. Likewise, any disparity of cell response across different cell types could be dependent on the profile of lysine demethylase activity of each cell type as a whole, not just one enzyme. A deeper understanding of the biology of cancer types and histone lysine demethylases could allow us to exploit this possibility by targeting multiple demethylases at the same time.

Our work utilizes only cells in culture and, as with any study, would benefit from *in vivo* support. Several studies have shown that KDM5B regulates invasion and migration in 2-D assays [259, 268-272]. Some of these studies have also shown that cells overexpressing KDM5B are able to form metastasis after tail vein injection in mice [259, 271]. KDM5B has even been found to be overexpressed in human metastases as compared to non-metastatic tumors [270, 271]. Just because some cell types can invade locally does not mean they are capable of forming metastatic tumors [334-336]. In fact, experts in the field often debate about whether to target metastasis as a preventative measure at local invasion of the primary site or to treat the secondary, metastatic tumors [368]. As KDM5B can contribute to both actions, KDM5B may be a good anti-metastatic target. Future studies should include analysis of KDM5B inhibitors in xenograft models and tail vein injections in mice so we can study the impacts of KDM5B inhibition on both ends of the metastatic spectrum. Ultimately, the work presented in this dissertation aims to prevent deaths from metastatic cancer through understanding the process and identifying targets.

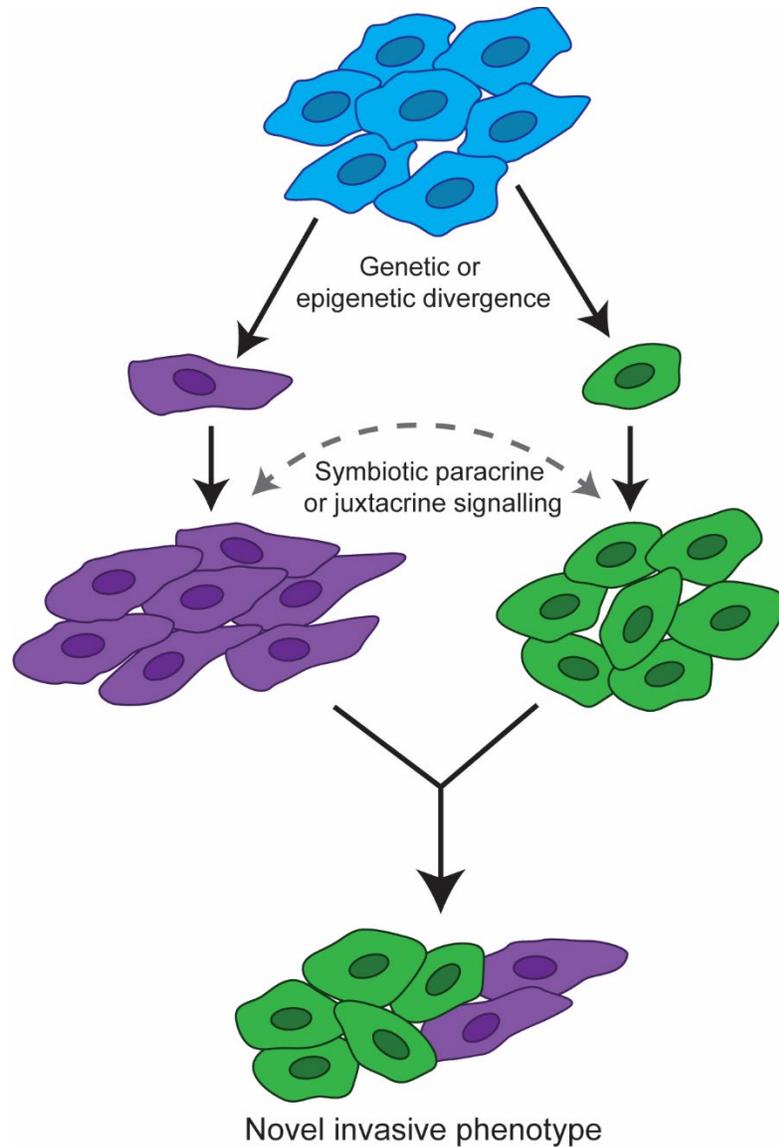


Figure 4.1 Model of invasive cooperativity within a heterogeneous population. Within a population of tumor cells, genetic or epigenetic alterations can give rise to subpopulations with divergent phenotypes. These new phenotypes may include producing or receiving paracrine signals that create a symbiosis between subpopulations. One possible result of this symbiosis is the ability to invade surrounding tissues that neither subpopulation would be capable of on its own.

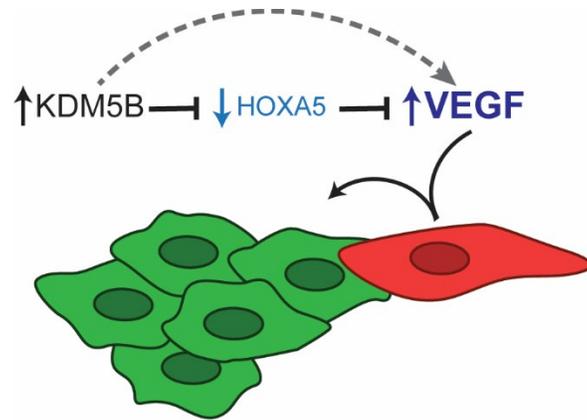


Figure 4.2. KDM5B may regulate leader behavior through HOXA5 and VEGF. KDM5B suppresses the transcription factor HOXA5. HOXA5 reduces expression of VEGF, which is molecule overexpressed in H1299 leaders and required for collective chain formation. Thus, upregulation of KDM5B could lead to upregulation of VEGF and promote leader behavior.

Chapter 5

Materials and methods

The following text is largely excerpted or adapted from “Genetic heterogeneity within collective invasion packs drives leader and follower cell phenotypes” under revision at *Journal of Cell Science* and

“Characterization of a Linked Jumonji Domain of the KDM5/JARID1 Family of Histone H3 Lysine 4 Demethylases” published by John R. Horton, Amanda Engstrom, Elizabeth L. Zoeller, Xu Liu, John R. Shanks, Xing Zhang, Margaret A. Johns, Paula M. Vertino, Haiyan Fu, and Xiaodong

Cheng in *Journal of Biological Chemistry*, 2016 Feb. 5.

Cell lines and culture conditions.

Leader and follower cells isolated via the SaGA technique from the H1299 human NSCLC cell line [288], as well as the parental H1299 cell line (ATCC, Manassas, VA), were cultured in RPMI-1640 media supplemented with 10% fetal bovine serum and 100 units mL⁻¹ of penicillin/streptomycin, and maintained at 37°C and 5% CO₂. H1299 cells had been transfected with the Dendra2 as previously described [288]. 293T cells were maintained in Dulbecco's Modified Eagle's Media (DMEM) supplemented with 10% fetal bovine serum and 100 units mL⁻¹ of penicillin/streptomycin at 37°C and 5% CO₂. MCF7 and MDA-MB231 breast cancer cell lines (ATCC, Manassas, VA) were maintained in Dulbecco's modified Eagle's medium plus 10% fetal bovine serum and 2mM glutamine and cultured in a humidified incubator at 37°C and 5% CO₂.

ARP3 knockdown and overexpression lines.

Lentivirus was prepared by seeding 2x10⁶ HEK293T cells in a 100 cm dish and co-transfecting with 5 µg transfer vector, 0.5 µg pMD2.G (Plasmid #12259, Addgene), 5 µg psPAX2 (Plasmid #12260, Addgene), and 1 µg of lentiviral vector. After 24 hours, media was replaced with 5 mL fresh complete media, then virus-containing media was collected after 24 hours. Media was centrifuged for 3 minutes at 1000 rpm, 4°C, and then supernatant was filtered through a 0.45 µm low protein-binding filter.

Target cells were seeded at 70% confluence in a 6-well plate one day prior to lentivirus collection. After virus collection, target cell media was replaced with 1 mL complete media plus 340 µL virus stock and 1.34 µL polybrene (10 mg mL⁻¹ stock), added dropwise. After 24 hours, media was replaced with 2 mL complete media. Selection antibiotics (shRNA: puromycin; ARP3 expression vectors: hygromycin) were added 48 hours after viral infection.

shACTR3 constructs were obtained from Millipore Sigma TRCN0000029383 and TRCN0000380403. mCherry-ARP3 lentiviral constructs were created by cloning ARP3-pmCherryC1 (a gift from Christien Merrifield; Addgene plasmid # 27682 [369]) into the pCDH-UBC-MCS-EF1 Hygro backbone. The UBC promoter was subsequently exchanged for a CMV promoter. The ARP3 K240R mutation was created using site-directed mutagenesis.

KDM5B knockdown and overexpression lines.

The pLKO-scrambled, pLKO-KDM5Bsh1, and pLKO-KDM5Bsh2 constructs were a kind gift from Dr. Qin Yan as previously described [282]. Lentivirus was prepared by seeding 9×10^5 HEK293T cells in a 6cm dish and co-transfecting with 0.1 μg VSV.G (viral envelope plasmid), 0.9 μg $\Delta\text{R8.2}$ (viral packaging plasmid), and 1 μg of lentiviral vector 24 hours post-seeding. After additional 24 hours, media was replaced with 6 mL fresh complete media, then virus-containing media was collected after further 24 hours. Media was centrifuged for 3 minutes at 1000 rpm, 4°C, and then supernatant was filtered through a 0.45 μm low protein-binding filter.

Target cells (H1299 parental, leader, and follower cell lines) were seeded at 1×10^5 cells per well in a 6-well plate one day prior to lentivirus collection. After virus collection, target cell media was replaced with 1 mL complete media plus 300 μL virus stock, and 10 $\mu\text{g mL}^{-1}$ polybrene (final), added dropwise. After 24 hours, media was replaced with 2 mL complete media. Cells were selected with 3 $\mu\text{g mL}^{-1}$ puromycin at 48 hours after viral infection.

For KDM5B knockdown lines stably express mCherry, H1299 leader cells with stably expressing either scrambled control RNA, KDM5Bsh1, or KDM5Bsh2 were transiently transfected with 2 μg mCherry hygro vector to 1×10^5 cells per well of 6-well plate with Lipofectamine 3000 (ThermoFisher Scientific Inc.) After 48 hours, plasmid expressing cells were selected with 500 $\mu\text{g mL}^{-1}$ hygromycin.

The HA-KDM5B construct was a kind gift from Dr. Qin Yan [282]. The puromycin selection marker was exchanged for a hygromycin cassette, and the KDM5B L685W mutation was introduced by site directed mutagenesis. H1299 cells were transfected with 2 μg plasmid vector to 1×10^5 cells per well of 6-well plate with Lipofectamine 3000 (ThermoFisher Scientific Inc., Carlsbad, CA, USA). After 48 hours, transduced cells were selected with 500 $\mu\text{g mL}^{-1}$ hygromycin.

RNA-sequencing and variant calling.

RNA-sequencing was performed in triplicate on H1299 parental, leader and follower cells. For parental cells, three different passages were used. For follower cells, three separately-isolated populations were used. For leader cells, two separately-isolated populations were used: one passage of one population, and two passages of the other. RNA library preparation and sequencing were carried out by the Emory Integrated Genomics core and Omega Bio-Tek, Inc. using the TruSeq Stranded mRNA kit, followed by quantification using a Quantus Fluoremeter (Promega, Madison, WI, USA) and integrity assessment using an Agilent 2200 TapeStation instrument. Sequencing was performed using a HiSeq2500 instrument (Illumina, Inc., San Diego, CA, USA), with 50M total sequencing reads generated per sample using the PE100 run format.

Data processing and statistical analyses were performed by the Emory Biostatistics and Bioinformatics Shared Resource. Raw paired-end fastq reads were assessed for quality and contamination using FastQC [370] and trimmed with Trimmomatic v0.32 [371]. Quality filtered reads were mapped against human reference genome hg19 using STAR aligner v2.3.0 [372]. Picard tools v1.111 (<http://broadinstitute.github.io/picard>) was used to assess post-alignment QC and to remove PCR duplicates. With an average yield of 30M post-filtered reads, 88% of the reads mapped uniquely with 78% of reads covered in the coding and UTR region. Genomic variants from RNA-seq were called using SamTools v0.1.19 mpileup [373] with Varscan v2.3.6 [374] and functionally annotated

using ANNOVAR [375]. A filtering criterion was applied requiring that reported variant had $\geq 6X$ read depth coverage, $\geq 2X$ supporting alternate reads in all samples. Variants associated with intronic, intergenic or synonymous changes, pseudo genes, non-coding RNAs, or sex chromosomes were excluded. Variants were filtered if they were known in dbSNP but not present in COSMIC database. This resulted in a total of 6240 variants. A pairwise two-sided independent t-test was done to compare cell populations mean variant allele frequencies between the groups.

Data Deposition

RNAseq data generated as part of this project have been deposited in the NCBI Sequence Read Archive (SRA) under accession # PRJNA542374.

Western blotting.

Total cellular protein expression was assessed via Western blotting as previously described [376]. For histone detection, lifted cells were washed twice with PBS and lysed with Triton extract buffer (PBS containing 0.5% Triton X 100, 2 mM phenylmethylsulfonyl fluoride (PMSF), and 0.02% sodium azide) 10 minutes on ice. Samples were collected by centrifugation for at 4°C for 10 minutes at 2000 rpm. Histones were then acid extracted in 0.2 M HCl by rotation overnight at 4°C. Samples were then electrophoresed on a 6.5% SDS-PAGE gel and subject to Western blotting.

Reagents and antibodies.

Primary antibodies for Western blot: ARP3 antibodies (Santa Cruz, cat. no. sc-48344; Proteintech, cat. no. 13822-1-AP) were used at 1:1000. KDM5B antibody (Sigma, cat. no. HPA 027179) was used at 1:500. KDM5A antibody (Cell Signaling, cat. no. 3876S) was used at 1:1000 in 5% milk. KDM5C antibody (Active Motif, cat. no. 39229) was used at 1:2000 in 5% milk. GAPDH

antibody (Cell Signaling, cat. no. 2118) was used at 1:30,000. Beta-tubulin antibody was used (Sigma, cat. no. T4026) at 1:5000. HA antibody (Biolegend cat. no. 901501) was used at 1:1000. H3K4me3 and H3K4me2 antibodies (Millipore cat. nos. 07-473 and 07-030 respectively) were used at 1:10,000. H3K4me1 antibody (Cell Signaling cat. no. 5326T) was used at 1:5000. H4 antibody (Millipore cat. no. 05-858R) was used at 1:20,000. Horseradish peroxidase-conjugated secondary antibodies (Jackson ImmunoResearch) were used at 1:10,000 for Western blot.

JIB-04 (Sigma-Aldrich SML0808), GSK-J4 (Sigma-Aldrich SML0701), and GSK-J5 (Cayman 12074) were dissolved in 100% DMSO to 20 and 40 mM, respectively. KDM5-C49 (Xcess Biosciences M60191) and KDM5-C70 (Xcess Biosciences M60192) were dissolved in 100% DMSO to 50 mM. GDC-50, Dong-A-167, and CPI-48 were gifted to us at 10 mM in 100% DMSO by Dr. Mark Henderson and the National Center for Advancing Translational Sciences Chemical Genomics Center (NCGC). GDC-50, Dong-A-167, and CPI-48 were synthesized respectively as in patent WO2016057924 and Horton et al., patent WO2016068580, and Liang et al [286, 364]. Stock solutions were stored at -20 °C.

3-D invasion assays, spheroid microscopy and image analysis.

Spheroids were generated as previously described [376] and embedded in 2 mg mL⁻¹ Matrigel (BD Biosciences) diluted in complete media. Images were taken at 0, 24, and in some cases 48 hours post-embedding at 4x using either an Olympus IX51 or CKX41 microscope.

For mixed spheroid experiments, cells were plated together in low-adhesion wells in the indicated ratios with 3000 total cells per spheroid and embedded as previously described. After 24 hours, spheroids were imaged using a Leica SP8 inverted confocal microscope. Invasive area and spheroid circularity were measured using ImageJ as previously described [288].

For invasion assays utilizing chemical inhibitors, H1299 parental, leader, and follower cells were plated at 3×10^4 cells/well in a 6-well plate and treated with 5 μ M KDM5-C70 or vehicle equivalent for 72 hours post-adhesion. Cells were then lifted and plated for spheroid formation as described above in media containing 5 μ M KDM5-C70 or equivalent vehicle for 72 hours. Spheroids were then embedded in Matrigel as above. Both Matrigel and media in dish contained 5 μ M KDM5-C70 or vehicle equivalent. Spheroids treated with GDC-50, Dong-A-167, and CPI-48 were plated and embedded as above in 5 μ M final in media during spheroid formation and in Matrigel and media during invasion.

Target validation.

DNA and RNA were isolated from H1299 parental, leader and follower cells using DNeasy Blood & Tissue Kit and the RNeasy Mini Kit (Qiagen Sciences, Germantown, MD, USA), respectively. Isolation of samples occurred in two independent biological replicates. RNA was reverse transcribed with M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) to generate cDNA. Primers were designed and used to amplify regions surrounding each locus of interest subject to Sanger sequencing at GENEWIZ (South Plainfield, NJ, USA). Primer sequences are listed in Table S1. PCR products were cloned into the pCR4-TOPO TA vector by TOPO-TA cloning (ThermoFisher Scientific 450071, Carlsbad, CA, USA) and transformed into bacteria. Fifty individual colonies for each gene in each cell type were and re-streaked on a new ampicillin plate. After 24 hours twenty colonies each were sent to GENEWIZ for Sanger sequencing.

Targeted deep resequencing and analysis.

Targeted region of KDM5B surrounding exon 15 was PCR amplified using primer sequences listed in Table S1. Amplified products were purified by SPRI beads (KAPA Biosystems, Cape Town,

South Africa). Libraries were then created with custom TruSeq compatible adapters and KAPA Hyper Prep Kit (KAPA Biosystems, Cape Town, South Africa). Quality for each library was checked using an Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Libraries were sequence on an Illumina HiSeq 4000 (Illumina, Inc., San Diego, CA, USA) using 150 bp paired-end sequencing at NYU Genome Technology Center. Quality trimmed reads were mapped to the human genome (GRCh37) using Bowtie 2 [377]. Variant allele frequencies for the KDM5B SNP (rs1141108) were quantified as the fractions # of reads (average depth= 348,327 reads for allele per line) at the corresponding genomic position exhibiting the *A* or *G* allele at position chr1: 202715284 using R packages Rsamtools, ShortRead, GenomicAlignments, and BSgenome.Hsapiens.UCSC.hg19 [378-382]. Differences in KDM5B variant allele frequencies were based on analysis of variance with Tukey's post-hoc correction using the R functions 'aov' and 'TukeyHSD', respectively.

Histone demethylase activity assay.

Direct quantification of H3K4me3 demethylase activity, cells were lysed by sonication in PBS (3x 4 sec) in the presence of protease inhibitors. Equal amounts of protein were incubated with a histone H3K4me3 substrate for 2 hours at 37°C before specific immune-detection of the H3K4me2 product using Epigenase™ JARID Demethylase Activity/Inhibition Assay Kit (Epigentek P-3083, Farmingdale, NY, USA). Detailed reaction conditions were as previously described [357, 383].

Statistical analysis.

Two-tailed, unpaired Student's t-test was used to assess statistical significance between any two conditions. Ordinary one-way ANOVA with Tukey's multiple comparisons test was used for experiments in which three or more conditions were being compared. Fisher's exact test was used to compare observed versus expected values for invasive chains; odds ratios were calculated via the

Baptista Pike method. Significance was concluded for p values <0.05 . Error bars represent mean \pm s.d. Confidence intervals of proportions in the mixed spheroid experiments were calculated via the Wilson/Brown method. Fisher's exact test was used to test the association of the identified mutations and the phenotypes (e.g. mutant vs. wild-type, leader versus follower) in TOPO-TA cloning experiments.

Cell proliferation assay.

H1299 cells expressing constructs for either overexpression or knockdown of KDM5B were seeded at 1×10^3 cells/well in six wells each of the interior wells of 96 well plate and grown in culture conditions described above until fixed at 24 hours, 48 hours, 72 hours, 96 hours, and 120 hours post-seeding. For drug treatment experiments, H1299 cells were plated at 2×10^3 cells/well, and MCF7 and MDA-MB231 cell lines 5×10^3 cells/well grown in triplicate wells in cell conditions as described above. After adhering to the wells for 24 hours, cells were treated for 72 hours in the presence of a compound or vehicle until fixation. One plate of adherent cells (plate Day0) at 24 hours was also fixed on the day of drug treatment. Cells were fixed with $100 \mu\text{L}$ /well cold 10% TCA for 1 hour at 4°C , washed three times in ultrapure water, and dried. Cells were then stained with $100 \mu\text{L}$ /well 0.4% sulforhodamine B in 1% acetic acid for 30 minutes, washed three times in 1% acetic acid, and dried. Dye was dissolved in $200 \mu\text{L}$ /well 10mM Tris base (pH10.5) 30 min. Optical density (OD) was measured by spectrophotometer at 510nm. Fold-change in OD follows: $(\text{OD}_{\text{X hours}}) / (\text{OD}_{24 \text{ hours}})$. The percentage of control cell growth was calculated as $(\text{Treatment group OD}_{72 \text{ hours}} - \text{OD}_{\text{Day 0}}) / (\text{Untreated group OD}_{72 \text{ hours}} - \text{OD}_{\text{Day 0}})$.

Quantitative Reverse Transcriptase PCR.

RNA was isolated with RNeasy Mini-Kit (Qiagen Sciences) and reverse transcribed with M-MLV Reverse Transcriptase (Invitrogen). The cDNA was amplified with primers specific to KDM5B and β tubulin using IQ SYBR Green Supermix (Bio-Rad) in real-time PCR. Primer sequences as follows: KDM5B forward GTGCGCTGTACTGTGAAGGA, KDM5B reverse GGCTGGCTTGAGTACCAGTC, β -tubulin forward CTTCGGCCAGATCTTCAGAC, and β -tubulin reverse AGAGAGTGGGTCAGCTGGAA.

References

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2019*. CA Cancer J Clin, 2019. **69**(1): p. 7-34.
2. Mehlen, P. and A. Puisieux, *Metastasis: a question of life or death*. Nat Rev Cancer, 2006. **6**(6): p. 449-58.
3. Torre, L.A., R.L. Siegel, and A. Jemal, *Lung Cancer Statistics*. Adv Exp Med Biol, 2016. **893**: p. 1-19.
4. McGranahan, N. and C. Swanton, *Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future*. Cell, 2017. **168**(4): p. 613-628.
5. Johnson, B.E., T. Mazor, C. Hong, M. Barnes, K. Aihara, C.Y. McLean, S.D. Fouse, S. Yamamoto, H. Ueda, K. Tatsuno, et al., *Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma*. Science, 2014. **343**(6167): p. 189-193.
6. de Bruin, E.C., N. McGranahan, R. Mitter, M. Salm, D.C. Wedge, L. Yates, M. Jamal-Hanjani, S. Shafi, N. Murugaesu, A.J. Rowan, et al., *Spatial and temporal diversity in genomic instability processes defines lung cancer evolution*. Science, 2014. **346**(6206): p. 251-6.
7. Zhang, J., J. Fujimoto, J. Zhang, D.C. Wedge, X. Song, J. Zhang, S. Seth, C.W. Chow, Y. Cao, C. Gumbs, et al., *Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing*. Science, 2014. **346**(6206): p. 256-9.
8. Sottoriva, A., H. Kang, Z. Ma, T.A. Graham, M.P. Salomon, J. Zhao, P. Marjoram, K. Siegmund, M.F. Press, D. Shibata, et al., *A Big Bang model of human colorectal tumor growth*. Nat Genet, 2015. **47**(3): p. 209-16.
9. Gerlinger, M., S. Horswell, J. Larkin, A.J. Rowan, M.P. Salm, I. Varela, R. Fisher, N. McGranahan, N. Matthews, C.R. Santos, et al., *Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing*. Nat Genet, 2014. **46**(3): p. 225-233.

10. Stoecklein, N.H., S.B. Hosch, M. Bezler, F. Stern, C.H. Hartmann, C. Vay, A. Siegmund, P. Scheunemann, P. Schurr, W.T. Knoefel, et al., *Direct genetic analysis of single disseminated cancer cells for prediction of outcome and therapy selection in esophageal cancer*. *Cancer Cell*, 2008. **13**(5): p. 441-53.
11. Anderson, K., C. Lutz, F.W. van Delft, C.M. Bateman, Y. Guo, S.M. Colman, H. Kempfski, A.V. Moorman, I. Tittley, J. Swansbury, et al., *Genetic variegation of clonal architecture and propagating cells in leukaemia*. *Nature*, 2011. **469**(7330): p. 356-61.
12. Potter, N.E., L. Ermini, E. Papaemmanuil, G. Cazzaniga, G. Vijayaraghavan, I. Tittley, A. Ford, P. Campbell, L. Kearney, and M. Greaves, *Single-cell mutational profiling and clonal phylogeny in cancer*. *Genome Research*, 2013. **23**(12): p. 2115-2125.
13. Meyer, M., J. Reimand, X. Lan, R. Head, X. Zhu, M. Kushida, J. Bayani, J.C. Pressey, A.C. Lionel, I.D. Clarke, et al., *Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity*. *Proc Natl Acad Sci U S A*, 2015. **112**(3): p. 851-856.
14. Wang, Y., J. Waters, M.L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, et al., *Clonal evolution in breast cancer revealed by single nucleus genome sequencing*. *Nature*, 2014. **512**(7513): p. 155-60.
15. Patel, A.P., I. Tirosh, J.J. Trombetta, A.K. Shalek, S.M. Gillespie, H. Wakimoto, D.P. Cahill, B.V. Nahed, W.T. Curry, R.L. Martuza, et al., *Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma*. *Science*, 2014. **344**(6190): p. 1396-1401.
16. Ramskold, D., S. Luo, Y.C. Wang, R. Li, Q. Deng, O.R. Faridani, G.A. Daniels, I. Khrebtukova, J.F. Loring, L.C. Laurent, et al., *Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells*. *Nat Biotechnol*, 2012. **30**(8): p. 777-82.

17. Navin, N., J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, et al., *Tumour evolution inferred by single-cell sequencing*. Nature, 2011. **472**(7341): p. 90-4.
18. Seshan, V.E. and R. Shen, *FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing*. Nucleic Acids Research, 2016. **44**(16): p. e131-e131.
19. Fischer, A., I. Vazquez-Garcia, C.J.R. Illingworth, and V. Mustonen, *High-definition reconstruction of clonal composition in cancer*. Cell Rep, 2014. **7**(5): p. 1740-1752.
20. Jiang, Y., Y. Qiu, A.J. Minn, and N.R. Zhang, *Assessing intratumor heterogeneity and tracking longitudinal and spatial clonal evolutionary history by next-generation sequencing*. Proc Natl Acad Sci U S A, 2016. **113**(37): p. E5528-E5537.
21. Deshwar, A.G., S. Vembu, C.K. Yung, G.H. Jang, L. Stein, and Q.J.G.B. Morris, *PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors*. Genome Biology, 2015. **16**(1): p. 35.
22. Roth, A., A. McPherson, E. Laks, J. Biele, D. Yap, A. Wan, M.A. Smith, C.B. Nielsen, J.N. McAlpine, S. Aparicio, et al., *Clonal genotype and population structure inference from single-cell tumor sequencing*. Nature Methods, 2016. **13**: p. 573.
23. Ha, G., A. Roth, J. Khattra, J. Ho, D. Yap, L.M. Prentice, N. Melnyk, A. McPherson, A. Bashashati, E. Laks, et al., *TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data*. Genome Research, 2014. **24**(11): p. 1881-1893.
24. Greaves, M., *Evolutionary determinants of cancer*. Cancer Discov, 2015. **5**(8): p. 806-20.
25. Nowell, P.C., *The clonal evolution of tumor cell populations*. Science, 1976. **194**(4260): p. 23-28.
26. Williams, M.J., B. Werner, C.P. Barnes, T.A. Graham, and A. Sottoriva, *Identification of neutral tumor evolution across cancer types*. Nature Genetics, 2016. **48**: p. 238.

27. Ling, S., Z. Hu, Z. Yang, F. Yang, Y. Li, P. Lin, K. Chen, L. Dong, L. Cao, Y. Tao, et al., *Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution*. Proc Natl Acad Sci U S A, 2015. **112**(47): p. E6496-E6505.
28. Driessens, G., B. Beck, A. Caauwe, B.D. Simons, and C. Blanpain, *Defining the mode of tumour growth by clonal analysis*. Nature, 2012. **488**: p. 527.
29. McGranahan, N., F. Favero, E.C. de Bruin, N.J. Birkbak, Z. Szallasi, and C. Swanton, *Clonal status of actionable driver events and the timing of mutational processes in cancer evolution*. Science Translational Medicine, 2015. **7**(283): p. 283ra54-283ra54.
30. Nadeu, F., J. Delgado, C. Royo, T. Baumann, T. Stankovic, M. Pinyol, P. Jares, A. Navarro, D. Martin-Garcia, S. Bea, et al., *Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia*. Blood, 2016. **127**(17): p. 2122-30.
31. Merlo, L.M. and C.C. Maley, *The role of genetic diversity in cancer*. J Clin Invest, 2010. **120**(2): p. 401-3.
32. Gerlinger, M., N. McGranahan, S.M. Dewhurst, R.A. Burrell, I. Tomlinson, and C. Swanton, *Cancer: evolution within a lifetime*. Annu Rev Genet, 2014. **48**: p. 215-36.
33. Burrell, R.A., N. McGranahan, J. Bartek, and C. Swanton, *The causes and consequences of genetic heterogeneity in cancer evolution*. Nature, 2013. **501**(7467): p. 338-45.
34. Landau, D.A., S.L. Carter, P. Stojanov, A. McKenna, K. Stevenson, M.S. Lawrence, C. Sougnez, C. Stewart, A. Sivachenko, L. Wang, et al., *Evolution and impact of subclonal mutations in chronic lymphocytic leukemia*. Cell, 2013. **152**(4): p. 714-26.
35. Szerlip, N.J., A. Pedraza, D. Chakravarty, M. Azim, J. McGuire, Y. Fang, T. Ozawa, E.C. Holland, J.T. Huse, S. Jhanwar, et al., *Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response*. Proc Natl Acad Sci U S A, 2012. **109**(8): p. 3041-6.

36. Inda, M.M., R. Bonavia, A. Mukasa, Y. Narita, D.W. Sah, S. Vandenberg, C. Brennan, T.G. Johns, R. Bachoo, P. Hadwiger, et al., *Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma*. *Genes Dev*, 2010. **24**(16): p. 1731-45.
37. Keats, J.J., M. Chesi, J.B. Egan, V.M. Garbitt, S.E. Palmer, E. Braggio, S. Van Wier, P.R. Blackburn, A.S. Baker, A. Dispenzieri, et al., *Clonal competition with alternating dominance in multiple myeloma*. *Blood*, 2012. **120**(5): p. 1067-76.
38. Turajlic, S. and C. Swanton, *Metastasis as an evolutionary process*. *Science*, 2016. **352**(6282): p. 169-75.
39. Goto, T., Y. Hirotsu, K. Amemiya, H. Mochizuki, and M. Omata, *Understanding Intratumor Heterogeneity and Evolution in NSCLC and Potential New Therapeutic Approach*. *Cancers (Basel)*, 2018. **10**(7).
40. Campbell, P.J., S. Yachida, L.J. Mudie, P.J. Stephens, E.D. Pleasance, L.A. Stebbings, L.A. Morsberger, C. Latimer, S. McLaren, M.L. Lin, et al., *The patterns and dynamics of genomic instability in metastatic pancreatic cancer*. *Nature*, 2010. **467**(7319): p. 1109-13.
41. Sun, R., Z. Hu, A. Sottoriva, T.A. Graham, A. Harpak, Z. Ma, J.M. Fischer, D. Shibata, and C. Curtis, *Between-region genetic divergence reflects the mode and tempo of tumor evolution*. *Nat Genet*, 2017. **49**(7): p. 1015-1024.
42. Fidler, I., *The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited*. *Nature Reviews Cancer*, 2003. **3**(6): p. 453-458.
43. Gupta, G.P. and J. Massague, *Cancer metastasis: building a framework*. *Cell*, 2006. **127**(4): p. 679-95.
44. Talmadge, J.E. and I.J. Fidler, *AACR centennial series: the biology of cancer metastasis: historical perspective*. *Cancer Res*, 2010. **70**(14): p. 5649-69.

45. Lambert, A.W., D.R. Pattabiraman, and R.A. Weinberg, *Emerging Biological Principles of Metastasis*. Cell, 2017. **168**(4): p. 670-691.
46. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites*. Nat Rev Cancer, 2002. **2**(8): p. 563-72.
47. Valastyan, S. and R.A. Weinberg, *Tumor metastasis: molecular insights and evolving paradigms*. Cell, 2011. **147**(2): p. 275-92.
48. Harris, A.L., *Hypoxia — a key regulatory factor in tumour growth*. Nature Reviews Cancer, 2002. **2**: p. 38.
49. Popper, H.H., *Progression and metastasis of lung cancer*. Cancer Metastasis Rev, 2016. **35**(1): p. 75-91.
50. Liu, Y.L., J.M. Yu, X.R. Song, X.W. Wang, L.G. Xing, and B.B. Gao, *Regulation of the chemokine receptor CXCR4 and metastasis by hypoxia-inducible factor in non small cell lung cancer cell lines*. Cancer Biol Ther, 2006. **5**(10): p. 1320-6.
51. Paliwal, S., R.C. Kovi, B. Nath, Y.W. Chen, B.C. Lewis, and S.R. Grossman, *The alternative reading frame tumor suppressor antagonizes hypoxia-induced cancer cell migration via interaction with the COOH-terminal binding protein corepressor*. Cancer Res, 2007. **67**(19): p. 9322-9.
52. Monnier, Y., P. Farmer, G. Bieler, N. Imaizumi, T. Sengstag, G.C. Alghisi, J.C. Stehle, L. Ciaroni, S. Andrejevic-Blant, R. Moeckli, et al., *CYR61 and alphaVbeta5 integrin cooperate to promote invasion and metastasis of tumors growing in preirradiated stroma*. Cancer Res, 2008. **68**(18): p. 7323-31.
53. Wei, L., X.R. Song, J.J. Sun, X.W. Wang, L. Xie, and L.Y. Lv, *Lysyl oxidase may play a critical role in hypoxia-induced NSCLC cells invasion and migration*. Cancer Biother Radiopharm, 2012. **27**(10): p. 672-7.

54. Kim, B., E.J. Sohn, J.H. Jung, E.A. Shin, O.H. You, J. Im, and S.H. Kim, *Inhibition of ZNF746 suppresses invasion and epithelial to mesenchymal transition in H460 non-small cell lung cancer cells*. *Oncol Rep*, 2014. **31**(1): p. 73-8.
55. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression*. *Nat Rev Cancer*, 2002. **2**(6): p. 442-54.
56. Lee, J.M., S. Dedhar, R. Kalluri, and E.W. Thompson, *The epithelial-mesenchymal transition: new insights in signaling, development, and disease*. *J Cell Biol*, 2006. **172**(7): p. 973-81.
57. Tam, W.L. and R.A. Weinberg, *The epigenetics of epithelial-mesenchymal plasticity in cancer*. *Nat Med*, 2013. **19**(11): p. 1438-49.
58. Savagner, P., *Epithelial-mesenchymal transitions: from cell plasticity to concept elasticity*. *Curr Top Dev Biol*, 2015. **112**: p. 273-300.
59. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition*. *J Clin Invest*, 2009. **119**(6): p. 1420-8.
60. Nieto, M.A., R.Y. Huang, R.A. Jackson, and J.P. Thiery, *EMT: 2016*. *Cell*, 2016. **166**(1): p. 21-45.
61. Thiery, J.P., H. Acloque, R.Y. Huang, and M.A. Nieto, *Epithelial-mesenchymal transitions in development and disease*. *Cell*, 2009. **139**(5): p. 871-90.
62. Brabletz, T., *EMT and MET in metastasis: where are the cancer stem cells?* *Cancer Cell*, 2012. **22**(6): p. 699-701.
63. Ocana, O.H., R. Corcoles, A. Fabra, G. Moreno-Bueno, H. Acloque, S. Vega, A. Barrallo-Gimeno, A. Cano, and M.A. Nieto, *Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer *Prrx1**. *Cancer Cell*, 2012. **22**(6): p. 709-24.

64. Korpai, M., B.J. Ell, F.M. Buffa, T. Ibrahim, M.A. Blanco, T. Celia-Terrassa, L. Mercatali, Z. Khan, H. Goodarzi, Y. Hua, et al., *Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization*. Nat Med, 2011. **17**(9): p. 1101-8.
65. Chao, Y.L., C.R. Shepard, and A. Wells, *Breast carcinoma cells re-express E-cadherin during mesenchymal to epithelial reverting transition*. Mol Cancer, 2010. **9**: p. 179.
66. Christiansen, J.J. and A.K. Rajasekaran, *Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis*. Cancer Res, 2006. **66**(17): p. 8319-26.
67. Tsai, J.H., J.L. Donaher, D.A. Murphy, S. Chau, and J. Yang, *Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis*. Cancer Cell, 2012. **22**(6): p. 725-36.
68. Jolly, M.K., K.E. Ware, S. Gilja, J.A. Somarelli, and H. Levine, *EMT and MET: necessary or permissive for metastasis?* Mol Oncol, 2017. **11**(7): p. 755-769.
69. Headley, M.B., A. Bins, A. Nip, E.W. Roberts, M.R. Looney, A. Gerard, and M.F. Krummel, *Visualization of immediate immune responses to pioneer metastatic cells in the lung*. Nature, 2016. **531**(7595): p. 513-7.
70. Reymond, N., B.B. d'Agua, and A.J. Ridley, *Crossing the endothelial barrier during metastasis*. Nat Rev Cancer, 2013. **13**(12): p. 858-70.
71. Luzzi, K.J., I.C. MacDonald, E.E. Schmidt, N. Kerkvliet, V.L. Morris, A.F. Chambers, and A.C. Groom, *Multistep Nature of Metastatic Inefficiency: Dormancy of Solitary Cells after Successful Extravasation and Limited Survival of Early Micrometastases*. The American Journal of Pathology, 1998. **153**(3): p. 865-873.
72. del Pozo Martin, Y., D. Park, A. Ramachandran, L. Ombrato, F. Calvo, P. Chakravarty, B. Spencer-Dene, S. Derzsi, Caroline S. Hill, E. Sahai, et al., *Mesenchymal Cancer Cell-Stroma*

- Crosstalk Promotes Niche Activation, Epithelial Reversion, and Metastatic Colonization.* Cell Reports, 2015. **13**(11): p. 2456-2469.
73. Mayor, R. and S. Etienne-Manneville, *The front and rear of collective cell migration.* Nat Rev Mol Cell Biol, 2016. **17**(2): p. 97-109.
74. Wang, X., A. Enomoto, N. Asai, T. Kato, and M. Takahashi, *Collective invasion of cancer: Perspectives from pathology and development.* Pathology International, 2016. **66**(4): p. 183-192.
75. Friedl, P. and D. Gilmour, *Collective cell migration in morphogenesis, regeneration and cancer.* Nature reviews. Molecular cell biology, 2009. **10**(july): p. 445-457.
76. Theveneau, E. and R. Mayor, *Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration.* Dev Biol, 2012. **366**(1): p. 34-54.
77. Ghashghaei, H.T., C. Lai, and E.S. Anton, *Neuronal migration in the adult brain: are we there yet?* Nat Rev Neurosci, 2007. **8**(2): p. 141-51.
78. Rorth, P., *Collective cell migration.* Annu Rev Cell Dev Biol, 2009. **25**: p. 407-29.
79. Andrew, D.J. and A.J. Ewald, *Morphogenesis of epithelial tubes: Insights into tube formation, elongation, and elaboration.* Developmental Biology, 2010. **341**(1): p. 34-55.
80. Montell, D.J., *Morphogenetic Cell Movements: Diversity from Modular Mechanical Properties.* Science, 2008. **322**(5907): p. 1502-1505.
81. Trepat, X. and J.J. Fredberg, *Plithotaxis and emergent dynamics in collective cellular migration.* Trends in Cell Biology, 2011. **21**(11): p. 638-646.
82. Cheung, K.J. and A.J. Ewald, *A collective route to metastasis: Seeding by tumor cell clusters.* Science, 2016. **352**(6282): p. 167-169.
83. Zeidman, I. and J.M. Buss, *Transpulmonary Passage of Tumor Cell Emboli.* Cancer Res, 1952. **12**(10): p. 731-733.

84. Liotta, L.A., J. Kleinerman, and G.M. Saldel, *The Significance of Hematogenous Tumor Cell Clumps in the Metastatic Process*. Cancer Research, 1976. **36**(3): p. 889-894.
85. Gudem, G., P. Van Loo, B. Kremeyer, L.B. Alexandrov, J.M.C. Tubio, E. Papaemmanuil, D.S. Brewer, H.M.L. Kallio, G. Hognas, M. Annala, et al., *The evolutionary history of lethal metastatic prostate cancer*. Nature, 2015. **520**(7547): p. 353-357.
86. Aceto, N., A. Bardia, D.T. Miyamoto, M.C. Donaldson, B.S. Wittner, J.A. Spencer, M. Yu, A. Pely, A. Engstrom, H. Zhu, et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis*. Cell, 2014. **158**(5): p. 1110-1122.
87. Maddipati, R. and B.Z. Stanger, *Pancreatic Cancer Metastases Harbor Evidence of Polyclonality*. Cancer Discovery, 2015. **5**(10): p. 1086-1097.
88. Cheung, K.J., V. Padmanaban, V. Silvestri, K. Schipper, J.D. Cohen, A.N. Fairchild, M.A. Gorin, J.E. Verdone, K.J. Pienta, J.S. Bader, et al., *Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters*. Proc Natl Acad Sci U S A, 2016. **113**(7): p. E854-63.
89. Haeger, A., K. Wolf, M.M. Zegers, and P. Friedl, *Collective cell migration: guidance principles and hierarchies*. Trends Cell Biol, 2015. **25**(9): p. 556-66.
90. Khalil, A.A. and P. Friedl, *Determinants of leader cells in collective cell migration*. Integrative Biology, 2010. **2**(11-12): p. 568-574.
91. Boettiger, D., *Mechanical control of integrin-mediated adhesion and signaling*. Curr Opin Cell Biol, 2012. **24**(5): p. 592-9.
92. Dabiri, B.E., H. Lee, and K.K. Parker, *A potential role for integrin signaling in mechano-electrical feedback*. Prog Biophys Mol Biol, 2012. **110**(2-3): p. 196-203.

93. Alexander, S., G.E. Koehl, M. Hirschberg, E.K. Geissler, and P. Friedl, *Dynamic imaging of cancer growth and invasion: a modified skin-fold chamber model*. *Histochem Cell Biol*, 2008. **130**(6): p. 1147-54.
94. Ventre, M., F. Causa, and P.A. Netti, *Determinants of cell-material crosstalk at the interface: towards engineering of cell instructive materials*. *J R Soc Interface*, 2012. **9**(74): p. 2017-32.
95. Etienne-Manneville, S. and A. Hall, *Integrin-mediated activation of Cdc42 controls cell polarity in migrating astrocytes through PKC ζ* . *Cell*, 2001. **106**(4): p. 489-98.
96. Gaggioli, C., S. Hooper, C. Hidalgo-Carcedo, R. Grosse, J.F. Marshall, K. Harrington, and E. Sahai, *Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells*. *Nat Cell Biol*, 2007. **9**(12): p. 1392-400.
97. Peng, H., Y.M. Ong, W.A. Shah, P.C. Holland, and S. Carbonetto, *Integrins regulate centrosome integrity and astrocyte polarization following a wound*. *Dev Neurobiol*, 2013. **73**(5): p. 333-53.
98. Priya, M.K., G. Sahu, D.R. Soto-Pantoja, N. Goldy, A.M. Sundaresan, V. Jadhav, T.R. Barathkumar, U. Saran, B.M. Jaffar Ali, D.D. Roberts, et al., *Tipping off endothelial tubes: nitric oxide drives tip cells*. *Angiogenesis*, 2015. **18**(2): p. 175-89.
99. Ispanovic, E., D. Serio, and T.L. Haas, *Cdc42 and RhoA have opposing roles in regulating membrane type 1-matrix metalloproteinase localization and matrix metalloproteinase-2 activation*. *Am J Physiol Cell Physiol*, 2008. **295**(3): p. C600-10.
100. Tan, W., T.R. Palmby, J. Gavard, P. Amornphimoltham, Y. Zheng, and J.S. Gutkind, *An essential role for Rac1 in endothelial cell function and vascular development*. *Faseb j*, 2008. **22**(6): p. 1829-38.
101. Braren, R., H. Hu, Y.H. Kim, H.E. Beggs, L.F. Reichardt, and R. Wang, *Endothelial FAK is essential for vascular network stability, cell survival, and lamellipodial formation*. *J Cell Biol*, 2006. **172**(1): p. 151-62.

102. Amerongen, G.P.v.N., P. Koolwijk, A. Versteilen, and V.W.M.v. Hinsbergh, *Involvement of RhoA/Rho Kinase Signaling in VEGF-Induced Endothelial Cell Migration and Angiogenesis In Vitro*. *Arterioscler Thromb Vasc Biol*, 2003. **23**(2): p. 211-217.
103. Lee, J.G. and E.P. Kay, *FGF-2-Induced Wound Healing in Corneal Endothelial Cells Requires Cdc42 Activation and Rho Inactivation through the Phosphatidylinositol 3-Kinase Pathway*. *Investigative Ophthalmology & Visual Science*, 2006. **47**(4): p. 1376-1386.
104. Pintucci, G., D. Moscatelli, F. Saponara, P.R. Biernacki, F.G. Baumann, C. Bizekis, A.C. Galloway, C. Basilico, and P. Mignatti, *Lack of ERK activation and cell migration in FGF-2-deficient endothelial cells*. *Faseb j*, 2002. **16**(6): p. 598-600.
105. El-Sibai, M., P. Nalbant, H. Pang, R.J. Flinn, C. Sarmiento, F. Macaluso, M. Cammer, J.S. Condeelis, K.M. Hahn, and J.M. Backer, *Cdc42 is required for EGF-stimulated protrusion and motility in MTLn3 carcinoma cells*. *J Cell Sci*, 2007. **120**(Pt 19): p. 3465-74.
106. Dise, R.S., M.R. Frey, R.H. Whitehead, and D.B. Polk, *Epidermal growth factor stimulates Rac activation through Src and phosphatidylinositol 3-kinase to promote colonic epithelial cell migration*. *Am J Physiol Gastrointest Liver Physiol*, 2008. **294**(1): p. G276-85.
107. Han, J., L. Li, J. Hu, L. Yu, Y. Zheng, J. Guo, X. Zheng, P. Yi, and Y. Zhou, *Epidermal growth factor stimulates human trophoblast cell migration through Rho A and Rho C activation*. *Endocrinology*, 2010. **151**(4): p. 1732-42.
108. Wilsbacher, J.L., S.L. Moores, J.S.J.C.C. Brugge, and Signaling, *An active form of Vav1 induces migration of mammary epithelial cells by stimulating secretion of an epidermal growth factor receptor ligand*. *Cell Communication and Signaling*, 2006. **4**(1): p. 5.
109. Kurokawa, K., R.E. Itoh, H. Yoshizaki, Y.O. Nakamura, and M. Matsuda, *Coactivation of Rac1 and Cdc42 at lamellipodia and membrane ruffles induced by epidermal growth factor*. *Mol Biol Cell*, 2004. **15**(3): p. 1003-10.

110. Kolsch, V., P.G. Charest, and R.A. Firtel, *The regulation of cell motility and chemotaxis by phospholipid signaling*. J Cell Sci, 2008. **121**(Pt 5): p. 551-9.
111. Barber, M.A. and H.C. Welch, *PI3K and RAC signalling in leukocyte and cancer cell migration*. Bull Cancer, 2006. **93**(5): p. E44-52.
112. Noguera-Troise, I., C. Daly, N.J. Papadopoulos, S. Coetzee, P. Boland, N.W. Gale, H.C. Lin, G.D. Yancopoulos, and G. Thurston, *Blockade of Dll4 inhibits tumour growth by promoting non-productive angiogenesis*. Nature, 2006. **444**(7122): p. 1032-7.
113. Hellstrom, M., L.K. Phng, J.J. Hofmann, E. Wallgard, L. Coultas, P. Lindblom, J. Alva, A.K. Nilsson, L. Karlsson, N. Gaiano, et al., *Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis*. Nature, 2007. **445**(7129): p. 776-80.
114. Cai, D. and D.J. Montell, *Diverse and dynamic sources and sinks in gradient formation and directed migration*. Curr Opin Cell Biol, 2014. **30**: p. 91-8.
115. Theveneau, E., L. Marchant, S. Kuriyama, M. Gull, B. Moepps, M. Parsons, and R. Mayor, *Collective chemotaxis requires contact-dependent cell polarity*. Dev Cell, 2010. **19**(1): p. 39-53.
116. Malet-Engra, G., W. Yu, A. Oldani, J. Rey-Barroso, N.S. Gov, G. Scita, and L. Dupre, *Collective cell motility promotes chemotactic prowess and resistance to chemorepulsion*. Curr Biol, 2015. **25**(2): p. 242-250.
117. Muinonen-Martin, A.J., O. Susanto, Q. Zhang, E. Smethurst, W.J. Faller, D.M. Veltman, G. Kalna, C. Lindsay, D.C. Bennett, O.J. Sansom, et al., *Melanoma cells break down LPA to establish local gradients that drive chemotactic dispersal*. PLoS Biol, 2014. **12**(10): p. e1001966.
118. Mayor, R. and C. Carmona-Fontaine, *Keeping in touch with contact inhibition of locomotion*. Trends Cell Biol, 2010. **20**(6): p. 319-28.
119. Becker, S.F., R. Mayor, and J. Kashef, *Cadherin-11 mediates contact inhibition of locomotion during Xenopus neural crest cell migration*. PLoS One, 2013. **8**(12): p. e85717.

120. Barriga, E.H., P.H. Maxwell, A.E. Reyes, and R. Mayor, *The hypoxia factor Hif-1alpha controls neural crest chemotaxis and epithelial to mesenchymal transition*. J Cell Biol, 2013. **201**(5): p. 759-76.
121. Astin, J.W., J. Batson, S. Kadir, J. Charlet, R.A. Persad, D. Gillatt, J.D. Oxley, and C.D. Nobes, *Competition amongst Eph receptors regulates contact inhibition of locomotion and invasiveness in prostate cancer cells*. Nat Cell Biol, 2010. **12**(12): p. 1194-204.
122. Batson, J., L. Maccarthy-Morrogh, A. Archer, H. Tanton, and C.D. Nobes, *EphA receptors regulate prostate cancer cell dissemination through Vav2-RhoA mediated cell-cell repulsion*. Biol Open, 2014. **3**(6): p. 453-62.
123. Villar-Cervino, V., M. Molano-Mazon, T. Catchpole, M. Valdeolmillos, M. Henkemeyer, L.M. Martinez, V. Borrell, and O. Marin, *Contact repulsion controls the dispersion and final distribution of Cajal-Retzius cells*. Neuron, 2013. **77**(3): p. 457-71.
124. Carmona-Fontaine, C., H. Matthews, and R. Mayor, *Directional cell migration in vivo: Wnt at the crest*. Cell Adh Migr, 2008. **2**(4): p. 240-2.
125. Mayor, R. and E. Theveneau, *The role of the non-canonical Wnt-planar cell polarity pathway in neural crest migration*. Biochem J, 2014. **457**(1): p. 19-26.
126. Shnitsar, I. and A. Borchers, *PTK7 recruits dsb to regulate neural crest migration*. Development, 2008. **135**(24): p. 4015-24.
127. Matthews, H.K., L. Marchant, C. Carmona-Fontaine, S. Kuriyama, J. Larrain, M.R. Holt, M. Parsons, and R. Mayor, *Directional migration of neural crest cells in vivo is regulated by Syndecan-4/Rac1 and non-canonical Wnt signaling/RhoA*. Development, 2008. **135**(10): p. 1771-80.
128. Moore, R., E. Theveneau, S. Pozzi, P. Alexandre, J. Richardson, A. Merks, M. Parsons, J. Kashef, C. Linker, and R. Mayor, *Par3 controls neural crest migration by promoting microtubule catastrophe during contact inhibition of locomotion*. Development, 2013. **140**(23): p. 4763-75.

129. Theveneau, E. and R. Mayor, *Integrating chemotaxis and contact-inhibition during collective cell migration: Small GTPases at work*. *Small GTPases*, 2010. **1**(2): p. 113-117.
130. Desai, R.A., S.B. Gopal, S. Chen, and C.S. Chen, *Contact inhibition of locomotion probabilities drive solitary versus collective cell migration*. *J R Soc Interface*, 2013. **10**(88): p. 20130717.
131. Ulmer, B., C. Hagenlocher, S. Schmalholz, S. Kurz, A. Schweickert, A. Kohl, L. Roth, D. Sela-Donenfeld, and M. Blum, *Calponin 2 acts as an effector of noncanonical Wnt-mediated cell polarization during neural crest cell migration*. *Cell Rep*, 2013. **3**(3): p. 615-21.
132. Davis, J.R., A. Luchici, F. Mosis, J. Thackery, J.A. Salazar, Y. Mao, G.A. Dunn, T. Betz, M. Miodownik, and B.M. Stramer, *Inter-cellular forces orchestrate contact inhibition of locomotion*. *Cell*, 2015. **161**(2): p. 361-73.
133. Krause, M. and A. Gautreau, *Steering cell migration: lamellipodium dynamics and the regulation of directional persistence*. *Nat Rev Mol Cell Biol*, 2014. **15**(9): p. 577-90.
134. Theveneau, E., B. Steventon, E. Scarpa, S. Garcia, X. Trepas, A. Streit, and R. Mayor, *Chase-and-run between adjacent cell populations promotes directional collective migration*. *Nat Cell Biol*, 2013. **15**(7): p. 763-72.
135. Scarpa, E., A. Szabo, A. Bibonne, E. Theveneau, M. Parsons, and R. Mayor, *Cadherin Switch during EMT in Neural Crest Cells Leads to Contact Inhibition of Locomotion via Repolarization of Forces*. *Dev Cell*, 2015. **34**(4): p. 421-34.
136. Martin, K., M. Vilela, N.L. Jeon, G. Danuser, and O. Pertz, *A growth factor-induced, spatially organizing cytoskeletal module enables rapid and persistent fibroblast migration*. *Dev Cell*, 2014. **30**(6): p. 701-16.
137. Yamada, S. and W.J. Nelson, *Localized zones of Rho and Rac activities drive initiation and expansion of epithelial cell-cell adhesion*. *J Cell Biol*, 2007. **178**(3): p. 517-27.

138. Caussinus, E., J. Colombelli, and M. Affolter, *Tip-cell migration controls stalk-cell intercalation during Drosophila tracheal tube elongation*. *Curr Biol*, 2008. **18**(22): p. 1727-34.
139. Reffay, M., M.C. Parrini, O. Cochet-Escartin, B. Ladoux, A. Buguin, S. Coscoy, F. Amblard, J. Camonis, and P. Silberzan, *Interplay of RhoA and mechanical forces in collective cell migration driven by leader cells*. *Nat Cell Biol*, 2014. **16**(3): p. 217-23.
140. Li, L., R. Hartley, B. Reiss, Y. Sun, J. Pu, D. Wu, F. Lin, T. Hoang, S. Yamada, J. Jiang, et al., *E-cadherin plays an essential role in collective directional migration of large epithelial sheets*. *Cell Mol Life Sci*, 2012. **69**(16): p. 2779-89.
141. Ladoux, B. and A. Nicolas, *Physically based principles of cell adhesion mechanosensitivity in tissues*. *Rep Prog Phys*, 2012. **75**(11): p. 116601.
142. Pruitt, B.L., A.R. Dunn, W.I. Weis, and W.J. Nelson, *Mechano-transduction: from molecules to tissues*. *PLoS Biol*, 2014. **12**(11): p. e1001996.
143. Yonemura, S., Y. Wada, T. Watanabe, A. Nagafuchi, and M. Shibata, *alpha-Catenin as a tension transducer that induces adherens junction development*. *Nat Cell Biol*, 2010. **12**(6): p. 533-42.
144. Yonemura, S., *A mechanism of mechanotransduction at the cell-cell interface: emergence of alpha-catenin as the center of a force-balancing mechanism for morphogenesis in multicellular organisms*. *Bioessays*, 2011. **33**(10): p. 732-6.
145. Leckband, D.E. and J. de Rooij, *Cadherin adhesion and mechanotransduction*. *Annu Rev Cell Dev Biol*, 2014. **30**: p. 291-315.
146. Trepap, X., M.R. Wasserman, T.E. Angelini, E. Millet, D.A. Weitz, J.P. Butler, and J.J. Fredberg, *Physical forces during collective cell migration*. *Nature Physics*, 2009. **5**: p. 426.
147. Lange, J.R. and B. Fabry, *Cell and tissue mechanics in cell migration*. *Exp Cell Res*, 2013. **319**(16): p. 2418-23.
148. Gov, N., *Cell mechanics: moving under peer pressure*. *Nat Mater*, 2011. **10**(6): p. 412-4.

149. Tambe, D.T., C.C. Hardin, T.E. Angelini, K. Rajendran, C.Y. Park, X. Serra-Picamal, E.H. Zhou, M.H. Zaman, J.P. Butler, D.A. Weitz, et al., *Collective cell guidance by cooperative intercellular forces*. Nat Mater, 2011. **10**(6): p. 469-75.
150. Dawson, M.A. and T. Kouzarides, *Cancer epigenetics: from mechanism to therapy*. Cell, 2012. **150**(1): p. 12-27.
151. Jones, P.A., J.P. Issa, and S. Baylin, *Targeting the cancer epigenome for therapy*. Nat Rev Genet, 2016. **17**(10): p. 630-41.
152. Weinstein, J.N., E.A. Collisson, G.B. Mills, K.R. Shaw, B.A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, and J.M. Stuart, *The Cancer Genome Atlas Pan-Cancer analysis project*. Nat Genet, 2013. **45**(10): p. 1113-20.
153. You, J.S. and P.A. Jones, *Cancer genetics and epigenetics: two sides of the same coin?* Cancer Cell, 2012. **22**(1): p. 9-20.
154. Kundaje, A., W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M.J. Ziller, et al., *Integrative analysis of 111 reference human epigenomes*. Nature, 2015. **518**(7539): p. 317-30.
155. Beck, S., B.E. Bernstein, R.M. Campbell, J.F. Costello, D. Dhanak, J.R. Ecker, J.M. Greally, J.P. Issa, P.W. Laird, K. Polyak, et al., *A blueprint for an international cancer epigenome consortium. A report from the AACR Cancer Epigenome Task Force*. Cancer Res, 2012. **72**(24): p. 6319-24.
156. Flavahan, W.A., E. Gaskell, and B.E. Bernstein, *Epigenetic plasticity and the hallmarks of cancer*. Science, 2017. **357**(6348).
157. Easwaran, H., H.C. Tsai, and S.B. Baylin, *Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance*. Mol Cell, 2014. **54**(5): p. 716-27.
158. Tessarz, P. and T. Kouzarides, *Histone core modifications regulating nucleosome structure and dynamics*. Nat Rev Mol Cell Biol, 2014. **15**(11): p. 703-8.

159. Mohammad, H.P., O. Barbash, and C.L. Creasy, *Targeting epigenetic modifications in cancer therapy: erasing the roadmap to cancer*. Nat Med, 2019. **25**(3): p. 403-418.
160. Bird, A., *DNA methylation patterns and epigenetic memory*. Genes Dev, 2002. **16**(1): p. 6-21.
161. Robertson, K.D., *DNA methylation and human disease*. Nature Reviews Genetics, 2005. **6**: p. 597.
162. Wu, H. and Y. Zhang, *Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation*. Genes Dev, 2011. **25**(23): p. 2436-52.
163. Kriaucionis, S. and N. Heintz, *The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain*. Science, 2009. **324**(5929): p. 929-930.
164. Tahiliani, M., K.P. Koh, Y. Shen, W.A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L.M. Iyer, D.R. Liu, L. Aravind, et al., *Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1*. Science, 2009. **324**(5929): p. 930-935.
165. Wu, X. and Y. Zhang, *TET-mediated active DNA demethylation: mechanism, function and beyond*. Nat Rev Genet, 2017. **18**(9): p. 517-534.
166. Ito, S., L. Shen, Q. Dai, S.C. Wu, L.B. Collins, J.A. Swenberg, C. He, and Y. Zhang, *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine*. Science, 2011. **333**(6047): p. 1300-3.
167. He, Y.F., B.Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, et al., *Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA*. Science, 2011. **333**(6047): p. 1303-7.
168. Weber, A.R., C. Krawczyk, A.B. Robertson, A. Kusnierczyk, C.B. Vagbo, D. Schuermann, A. Klungland, and P. Schar, *Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism*. Nat Commun, 2016. **7**: p. 10806.

169. Maiti, A. and A.C. Drohat, *Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites*. J Biol Chem, 2011. **286**(41): p. 35334-8.
170. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond*. Nat Rev Genet, 2012. **13**(7): p. 484-92.
171. Jones, P.A. and S.M. Taylor, *Cellular differentiation, cytidine analogs and DNA methylation*. Cell, 1980. **20**(1): p. 85-93.
172. Yang, X., H. Han, D.D. De Carvalho, F.D. Lay, P.A. Jones, and G. Liang, *Gene body methylation can alter gene expression and is a therapeutic target in cancer*. Cancer Cell, 2014. **26**(4): p. 577-90.
173. Issa, J.P., G. Garcia-Manero, F.J. Giles, R. Mannari, D. Thomas, S. Faderl, E. Bayar, J. Lyons, C.S. Rosenfeld, J. Cortes, et al., *Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies*. Blood, 2004. **103**(5): p. 1635-40.
174. Fenaux, P., G.J. Mufti, E. Hellstrom-Lindberg, V. Santini, C. Finelli, A. Giagounidis, R. Schoch, N. Gattermann, G. Sanz, A. List, et al., *Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study*. Lancet Oncol, 2009. **10**(3): p. 223-32.
175. Lubbert, M., S. Suciu, A. Hagemeyer, B. Ruter, U. Platzbecker, A. Giagounidis, D. Selleslag, B. Labar, U. Germing, H.R. Salih, et al., *Decitabine improves progression-free survival in older high-risk MDS patients with multiple autosomal monosomies: results of a subgroup analysis of the randomized phase III study 06011 of the EORTC Leukemia Cooperative Group and German MDS Study Group*. Ann Hematol, 2016. **95**(2): p. 191-9.

176. Oki, Y., J. Jelinek, L. Shen, H.M. Kantarjian, and J.P. Issa, *Induction of hypomethylation and molecular response after decitabine therapy in patients with chronic myelomonocytic leukemia*. *Blood*, 2008. **111**(4): p. 2382-4.
177. Tsai, H.C., H. Li, L. Van Neste, Y. Cai, C. Robert, F.V. Rassool, J.J. Shin, K.M. Harbom, R. Beatty, E. Pappou, et al., *Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells*. *Cancer Cell*, 2012. **21**(3): p. 430-46.
178. Prebet, T., S.D. Gore, B. Esterni, C. Gardin, R. Itzykson, S. Thepot, F. Dreyfus, O.B. Rauzy, C. Recher, L. Ades, et al., *Outcome of high-risk myelodysplastic syndrome after azacitidine treatment failure*. *J Clin Oncol*, 2011. **29**(24): p. 3322-7.
179. Stewart, D.J., J.P. Issa, R. Kurzrock, M.I. Nunez, J. Jelinek, D. Hong, Y. Oki, Z. Guo, S. Gupta, and Wistuba, II, *Decitabine effect on tumor global DNA methylation and other parameters in a phase I trial in refractory solid tumors and lymphomas*. *Clin Cancer Res*, 2009. **15**(11): p. 3881-8.
180. Kouzarides, T., *Chromatin Modifications and Their Function*. *Cell*, 2007. **128**(4): p. 693-705.
181. Tan, M., H. Luo, S. Lee, F. Jin, Jeong S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, et al., *Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification*. *Cell*, 2011. **146**(6): p. 1016-1028.
182. Falkenberg, K.J. and R.W. Johnstone, *Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders*. *Nat Rev Drug Discov*, 2014. **13**(9): p. 673-91.
183. Morin, R.D., N.A. Johnson, T.M. Severson, A.J. Mungall, J. An, R. Goya, J.E. Paul, M. Boyle, B.W. Woolcock, F. Kuchenbauer, et al., *Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin*. *Nat Genet*, 2010. **42**(2): p. 181-5.
184. McCabe, M.T., H.M. Ott, G. Ganji, S. Korenchuk, C. Thompson, G.S. Van Aller, Y. Liu, A.P. Graves, A. Della Pietra, 3rd, E. Diaz, et al., *EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations*. *Nature*, 2012. **492**(7427): p. 108-12.

185. Bodor, C., V. Grossmann, N. Popov, J. Okosun, C. O'Riain, K. Tan, J. Marzec, S. Araf, J. Wang, A.M. Lee, et al., *EZH2 mutations are frequent and represent an early event in follicular lymphoma*. Blood, 2013. **122**(18): p. 3165-8.
186. Kleer, C.G., Q. Cao, S. Varambally, R. Shen, I. Ota, S.A. Tomlins, D. Ghosh, R.G. Sewalt, A.P. Otte, D.F. Hayes, et al., *EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells*. Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11606-11.
187. Varambally, S., S.M. Dhanasekaran, M. Zhou, T.R. Barrette, C. Kumar-Sinha, M.G. Sanda, D. Ghosh, K.J. Pienta, R.G. Sewalt, A.P. Otte, et al., *The polycomb group protein EZH2 is involved in progression of prostate cancer*. Nature, 2002. **419**(6907): p. 624-9.
188. Yap, D.B., J. Chu, T. Berg, M. Schapira, S.W. Cheng, A. Moradian, R.D. Morin, A.J. Mungall, B. Meissner, M. Boyle, et al., *Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation*. Blood, 2011. **117**(8): p. 2451-9.
189. Wigle, T.J., S.K. Knutson, L. Jin, K.W. Kuntz, R.M. Pollock, V.M. Richon, R.A. Copeland, and M.P. Scott, *The Y641C mutation of EZH2 alters substrate specificity for histone H3 lysine 27 methylation states*. FEBS Lett, 2011. **585**(19): p. 3011-4.
190. Knutson, S.K., S. Kawano, Y. Minoshima, N.M. Warholic, K.C. Huang, Y. Xiao, T. Kadowaki, M. Uesugi, G. Kuznetsov, N. Kumar, et al., *Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma*. Mol Cancer Ther, 2014. **13**(4): p. 842-54.
191. Vaswani, R.G., V.S. Gehling, L.A. Dakin, A.S. Cook, C.G. Nasveschuk, M. Duplessis, P. Iyer, S. Balasubramanian, F. Zhao, A.C. Good, et al., *Identification of (R)-N-((4-Methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-2-methyl-1-(1-(2,2,2-trifluoroethyl)piperidin-4-yl)ethyl)-1H-indole-3-carboxamide (CPI-1205), a Potent and Selective Inhibitor of Histone Methyltransferase*

- EZH2, Suitable for Phase I Clinical Trials for B-Cell Lymphomas*. J Med Chem, 2016. **59**(21): p. 9928-9941.
192. Kung, P.P., P. Bingham, A. Brooun, M. Collins, Y.L. Deng, D. Dinh, C. Fan, K.S. Gajiwala, R. Grantner, H.J. Gukasyan, et al., *Optimization of Orally Bioavailable Enhancer of Zeste Homolog 2 (EZH2) Inhibitors Using Ligand and Property-Based Design Strategies: Identification of Development Candidate (R)-5,8-Dichloro-7-(methoxy(oxetan-3-yl)methyl)-2-((4-methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-3,4-dihydroisoquinolin-1(2H)-one (PF-06821497)*. J Med Chem, 2018. **61**(3): p. 650-665.
193. Sankar, S., E.R. Theisen, J. Bearss, T. Mulvihill, L.M. Hoffman, V. Sorna, M.C. Beckerle, S. Sharma, and S.L. Lessnick, *Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth*. Clin Cancer Res, 2014. **20**(17): p. 4584-97.
194. Sorna, V., E.R. Theisen, B. Stephens, S.L. Warner, D.J. Bearss, H. Vankayalapati, and S. Sharma, *High-throughput virtual screening identifies novel N'-(1-phenylethylidene)-benzohydrazides as potent, specific, and reversible LSD1 inhibitors*. J Med Chem, 2013. **56**(23): p. 9496-508.
195. Mould, D.P., A.E. McGonagle, D.H. Wiseman, E.L. Williams, and A.M. Jordan, *Reversible inhibitors of LSD1 as therapeutic agents in acute myeloid leukemia: clinical significance and progress to date*. Med Res Rev, 2015. **35**(3): p. 586-618.
196. Kuhn, M.W., E. Song, Z. Feng, A. Sinha, C.W. Chen, A.J. Deshpande, M. Cusan, N. Farnoud, A. Mupo, C. Grove, et al., *Targeting Chromatin Regulators Inhibits Leukemogenic Gene Expression in NPM1 Mutant Leukemia*. Cancer Discov, 2016. **6**(10): p. 1166-1181.
197. Falini, B., C. Mecucci, E. Tiacci, M. Alcalay, R. Rosati, L. Pasqualucci, R. La Starza, D. Diverio, E. Colombo, A. Santucci, et al., *Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype*. N Engl J Med, 2005. **352**(3): p. 254-66.

198. Heijnen, W.T., J. De Fruyt, A.I. Wierdsma, P. Sienaert, and T.K. Birkenhager, *Efficacy of Tranylcypromine in Bipolar Depression: A Systematic Review*. *J Clin Psychopharmacol*, 2015. **35**(6): p. 700-5.
199. Baker, G.B., R.T. Coutts, K.F. McKenna, and R.L. Sherry-McKenna, *Insights into the mechanisms of action of the MAO inhibitors phenelzine and tranylcypromine: a review*. *J Psychiatry Neurosci*, 1992. **17**(5): p. 206-14.
200. Harris, W.J., X. Huang, J.T. Lynch, G.J. Spencer, J.R. Hitchin, Y. Li, F. Ciceri, J.G. Blaser, B.F. Greystoke, A.M. Jordan, et al., *The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells*. *Cancer Cell*, 2012. **21**(4): p. 473-87.
201. Daigle, S.R., E.J. Olhava, C.A. Therkelsen, A. Basavapathruni, L. Jin, P.A. Boriack-Sjodin, C.J. Allain, C.R. Klaus, A. Raimondi, M.P. Scott, et al., *Potent inhibition of DOT1L as treatment of MLL-fusion leukemia*. *Blood*, 2013. **122**(6): p. 1017-25.
202. Xhabija, B. and B.L. Kidder, *KDM5B is a master regulator of the H3K4-methylome in stem cells, development and cancer*. *Semin Cancer Biol*, 2018.
203. Klose, R.J., Q. Yan, Z. Tothova, K. Yamane, H. Erdjument-Bromage, P. Tempst, D.G. Gilliland, Y. Zhang, and W.G. Kaelin, *The Retinoblastoma Binding Protein RBP2 Is an H3K4 Demethylase*. *Cell*, 2007. **128**(5): p. 889-900.
204. Yamane, K., K. Tateishi, R.J. Klose, J. Fang, L.A. Fabrizio, H. Erdjument-Bromage, J. Taylor-Papadimitriou, P. Tempst, and Y. Zhang, *PLU-1 Is an H3K4 Demethylase Involved in Transcriptional Repression and Breast Cancer Cell Proliferation*. *Molecular Cell*, 2007. **25**(6): p. 801-812.
205. Christensen, J., K. Agger, P.A.C. Cloos, D. Pasini, S. Rose, L. Sennels, J. Rappsilber, K.H. Hansen, A.E. Salcini, and K. Helin, *RBP2 Belongs to a Family of Demethylases, Specific for Tri- and Dimethylated Lysine 4 on Histone 3*. *Cell*, 2007. **128**(6): p. 1063-1076.

206. Tahiliani, M., P. Mei, R. Fang, T. Leonor, M. Rutenberg, F. Shimizu, J. Li, A. Rao, and Y. Shi, *The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation*. Nature, 2007. **447**: p. 601.
207. Zhang, Y., H. Yang, X. Guo, N. Rong, Y. Song, Y. Xu, W. Lan, X. Zhang, M. Liu, Y. Xu, et al., *The PHD1 finger of KDM5B recognizes unmodified H3K4 during the demethylation of histone H3K4me2/3 by KDM5B*. Protein Cell, 2014. **5**(11): p. 837-50.
208. Plch, J., J. Hrabeta, and T. Eckschlager, *KDM5 demethylases and their role in cancer cell chemoresistance*. Int J Cancer, 2019. **144**(2): p. 221-231.
209. Harmeyer, K.M., N.D. Facompre, M. Herlyn, and D. Basu, *JARID1 Histone Demethylases: Emerging Targets in Cancer*. Trends Cancer, 2017. **3**(10): p. 713-725.
210. Pilka, E.S., T. James, and J.H. Lisztwan, *Structural definitions of Jumonji family demethylase selectivity*. Drug Discov Today, 2015. **20**(6): p. 743-9.
211. Secombe, J. and R.N. Eisenman, *The function and regulation of the JARID1 family of histone H3 lysine 4 demethylases: the Myc connection*. Cell Cycle, 2007. **6**(11): p. 1324-8.
212. Johansson, C., S. Velupillai, A. Tumber, A. Szykowska, E.S. Hookway, R.P. Nowak, C. Strain-Damerell, C. Gileadi, M. Philpott, N. Burgess-Brown, et al., *Structural analysis of human KDM5B guides histone demethylase inhibitor development*. Nat Chem Biol, 2016. **12**(7): p. 539-45.
213. Horton, J.R., A. Engstrom, E.L. Zoeller, X. Liu, J.R. Shanks, X. Zhang, M.A. Johns, P.M. Vertino, H. Fu, and X. Cheng, *Characterization of a Linked Jumonji Domain of the KDM5/JARID1 Family of Histone H3 Lysine 4 Demethylases*. J Biol Chem, 2016. **291**(6): p. 2631-46.
214. Klein, B.J., L. Piao, Y. Xi, H. Rincon-Arano, S.B. Rothbart, D. Peng, H. Wen, C. Larson, X. Zhang, X. Zheng, et al., *The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers*. Cell Rep, 2014. **6**(2): p. 325-35.

215. Torres, I.O., K.M. Kuchenbecker, C.I. Nnadi, R.J. Fletterick, M.J. Kelly, and D.G. Fujimori, *Histone demethylase KDM5A is regulated by its reader domain through a positive-feedback mechanism*. Nat Commun, 2015. **6**: p. 6204.
216. Catchpole, S., B. Spencer-Dene, D. Hall, S. Santangelo, I. Rosewell, M. Guenatri, R. Beatson, A.G. Scibetta, J.M. Burchell, and J. Taylor-Papadimitriou, *PLU-1/JARID1B/KDM5B is required for embryonic survival and contributes to cell proliferation in the mammary gland and in ER+ breast cancer cells*. Int J Oncol, 2011. **38**(5): p. 1267-77.
217. Defeo-Jones, D., P.S. Huang, R.E. Jones, K.M. Haskell, G.A. Vuocolo, M.G. Hanobik, H.E. Huber, and A. Oliff, *Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product*. Nature, 1991. **352**(6332): p. 251-4.
218. Beshiri, M.L., K.B. Holmes, W.F. Richter, S. Hess, A.B. Islam, Q. Yan, L. Plante, L. Litovchick, N. Gevry, N. Lopez-Bigas, et al., *Coordinated repression of cell cycle genes by KDM5A and E2F4 during differentiation*. Proc Natl Acad Sci U S A, 2012. **109**(45): p. 18499-504.
219. Rasmussen, P.B. and P. Staller, *The KDM5 family of histone demethylases as targets in oncology drug discovery*. Epigenomics, 2014. **6**(3): p. 277-86.
220. Dahl, J.A., I. Jung, H. Aanes, G.D. Greggains, A. Manaf, M. Lerdrup, G. Li, S. Kuan, B. Li, A.Y. Lee, et al., *Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition*. Nature, 2016. **537**(7621): p. 548-552.
221. Madsen, B., B. Spencer-Dene, R. Poulsom, D. Hall, P.J. Lu, K. Scott, A.T. Shaw, J.M. Burchell, P. Freemont, and J. Taylor-Papadimitriou, *Characterisation and developmental expression of mouse Plu-1, a homologue of a human nuclear protein (PLU-1) which is specifically up-regulated in breast cancer*. Mech Dev, 2002. **119 Suppl 1**: p. S239-46.

222. Zou, M.R., J. Cao, Z. Liu, S.J. Huh, K. Polyak, and Q. Yan, *Histone demethylase jumonji AT-rich interactive domain 1B (JARID1B) controls mammary gland development by regulating key developmental and lineage specification genes*. J Biol Chem, 2014. **289**(25): p. 17620-33.
223. Kidder, B.L., G. Hu, Z.X. Yu, C. Liu, and K. Zhao, *Extended self-renewal and accelerated reprogramming in the absence of Kdm5b*. Mol Cell Biol, 2013. **33**(24): p. 4793-810.
224. Schmitz, S.U., M. Albert, M. Malatesta, L. Morey, J.V. Johansen, M. Bak, N. Tommerup, I. Abarrategui, and K. Helin, *Jarid1b targets genes regulating development and is involved in neural differentiation*. Embo j, 2011. **30**(22): p. 4586-600.
225. Albert, M., S.U. Schmitz, S.M. Kooistra, M. Malatesta, C. Morales Torres, J.C. Rekling, J.V. Johansen, I. Abarrategui, and K. Helin, *The histone demethylase Jarid1b ensures faithful mouse development by protecting developmental genes from aberrant H3K4me3*. PLoS Genet, 2013. **9**(4): p. e1003461.
226. Shen, H., W. Xu, and F. Lan, *Histone lysine demethylases in mammalian embryonic development*. Exp Mol Med, 2017. **49**(4): p. e325.
227. Cox, B.J., M. Vollmer, O. Tamplin, M. Lu, S. Biechele, M. Gertsenstein, C. van Campenhout, T. Floss, R. Kuhn, W. Wurst, et al., *Phenotypic annotation of the mouse X chromosome*. Genome Res, 2010. **20**(8): p. 1154-64.
228. Iwase, S., F. Lan, P. Bayliss, L. de la Torre-Ubieta, M. Huarte, H.H. Qi, J.R. Whetstine, A. Bonni, T.M. Roberts, and Y. Shi, *The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases*. Cell, 2007. **128**(6): p. 1077-88.
229. Santos-Reboucas, C.B., N. Fintelman-Rodrigues, L.R. Jensen, A.W. Kuss, M.G. Ribeiro, M. Campos, Jr., J.M. Santos, and M.M. Pimentel, *A novel nonsense mutation in KDM5C/JARID1C gene causing intellectual disability, short stature and speech delay*. Neurosci Lett, 2011. **498**(1): p. 67-71.

230. Tzschach, A., S. Lenzner, B. Moser, R. Reinhardt, J. Chelly, J.P. Fryns, T. Kleefstra, M. Raynaud, G. Turner, H.H. Ropers, et al., *Novel JARID1C/SMCX mutations in patients with X-linked mental retardation*. Hum Mutat, 2006. **27**(4): p. 389.
231. Jensen, L.R., M. Amende, U. Gurok, B. Moser, V. Gimmel, A. Tzschach, A.R. Janecke, G. Tariverdian, J. Chelly, J.P. Fryns, et al., *Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation*. Am J Hum Genet, 2005. **76**(2): p. 227-36.
232. Al-Mubarak, B., M. Abouelhoda, A. Omar, H. Aldhalaan, M. Aldosari, M. Nester, H.A. Alshamrani, M. El-Kalioby, E. Goljan, R. Albar, et al., *Whole exome sequencing reveals inherited and de novo variants in autism spectrum disorder: a trio study from Saudi families*. Sci Rep, 2017. **7**(1): p. 5679.
233. Athanasakis, E., D. Licastro, F. Faletra, A. Fabretto, S. Dipresa, D. Vozzi, A. Morgan, A.P. d'Adamo, V. Pecile, X. Biarnes, et al., *Next generation sequencing in nonsyndromic intellectual disability: from a negative molecular karyotype to a possible causative mutation detection*. Am J Med Genet A, 2014. **164a**(1): p. 170-6.
234. Vallianatos, C.N. and S. Iwase, *Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders*. Epigenomics, 2015. **7**(3): p. 503-19.
235. Akimoto, C., H. Kitagawa, T. Matsumoto, and S. Kato, *Spermatogenesis-specific association of SMCY and MSH5*. Genes Cells, 2008. **13**(6): p. 623-33.
236. Banelli, B., E. Carra, F. Barbieri, R. Wurth, F. Parodi, A. Pattarozzi, R. Carosio, A. Forlani, G. Allemanni, D. Marubbi, et al., *The histone demethylase KDM5A is a key factor for the resistance to temozolomide in glioblastoma*. Cell Cycle, 2015. **14**(21): p. 3418-29.

237. Zeng, J., Z. Ge, L. Wang, Q. Li, N. Wang, M. Bjorkholm, J. Jia, and D. Xu, *The histone demethylase RBP2 Is overexpressed in gastric cancer and its inhibition triggers senescence of cancer cells*. Gastroenterology, 2010. **138**(3): p. 981-92.
238. Liang, X., J. Zeng, L. Wang, M. Fang, Q. Wang, M. Zhao, X. Xu, Z. Liu, W. Li, S. Liu, et al., *Histone demethylase retinoblastoma binding protein 2 is overexpressed in hepatocellular carcinoma and negatively regulated by hsa-miR-212*. PLoS One, 2013. **8**(7): p. e69784.
239. Li, L., L. Wang, P. Song, X. Geng, X. Liang, M. Zhou, Y. Wang, C. Chen, J. Jia, and J. Zeng, *Critical role of histone demethylase RBP2 in human gastric cancer angiogenesis*. Mol Cancer, 2014. **13**: p. 81.
240. Teng, Y.C., C.F. Lee, Y.S. Li, Y.R. Chen, P.W. Hsiao, M.Y. Chan, F.M. Lin, H.D. Huang, Y.T. Chen, Y.M. Jeng, et al., *Histone demethylase RBP2 promotes lung tumorigenesis and cancer metastasis*. Cancer Res, 2013. **73**(15): p. 4711-21.
241. Cao, J., Z. Liu, W.K. Cheung, M. Zhao, S.Y. Chen, S.W. Chan, C.J. Booth, D.X. Nguyen, and Q. Yan, *Histone demethylase RBP2 is critical for breast cancer progression and metastasis*. Cell Rep, 2014. **6**(5): p. 868-77.
242. Zhou, D., V. Kannappan, X. Chen, J. Li, X. Leng, J. Zhang, and S. Xuan, *RBP2 induces stem-like cancer cells by promoting EMT and is a prognostic marker for renal cell carcinoma*. Exp Mol Med, 2016. **48**: p. e238.
243. Rondinelli, B., D. Rosano, E. Antonini, M. Frenquelli, L. Montanini, D. Huang, S. Segalla, K. Yoshihara, S.B. Amin, D. Lazarevic, et al., *Histone demethylase JARID1C inactivation triggers genomic instability in sporadic renal cancer*. J Clin Invest, 2015. **125**(12): p. 4625-37.
244. Guo, G., Y. Gui, S. Gao, A. Tang, X. Hu, Y. Huang, W. Jia, Z. Li, M. He, L. Sun, et al., *Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma*. Nature Genetics, 2011. **44**: p. 17.

245. Ricketts, C.J. and W.M. Linehan, *Gender Specific Mutation Incidence and Survival Associations in Clear Cell Renal Cell Carcinoma (CCRCC)*. PLoS One, 2015. **10**(10): p. e0140257.
246. Dunford, A., D.M. Weinstock, V. Savova, S.E. Schumacher, J.P. Cleary, A. Yoda, T.J. Sullivan, J.M. Hess, A.A. Gimelbrant, R. Beroukhim, et al., *Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias*. Nat Genet, 2017. **49**(1): p. 10-16.
247. Ricketts, C.J., A.A. De Cubas, H. Fan, C.C. Smith, M. Lang, E. Reznik, R. Bowlby, E.A. Gibb, R. Akbani, R. Beroukhim, et al., *The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma*. Cell Rep, 2018. **23**(1): p. 313-326.e5.
248. Stein, J., M. Majores, M. Rohde, S. Lim, S. Schneider, E. Krappe, J. Ellinger, M. Dietel, C. Stephan, K. Jung, et al., *KDM5C is overexpressed in prostate cancer and is a prognostic marker for prostate-specific antigen-relapse following radical prostatectomy*. Am J Pathol, 2014. **184**(9): p. 2430-7.
249. Wang, Q., J. Wei, P. Su, and P. Gao, *Histone demethylase JARID1C promotes breast cancer metastasis cells via down regulating BRMS1 expression*. Biochem Biophys Res Commun, 2015. **464**(2): p. 659-66.
250. Ji, X., S. Jin, X. Qu, K. Li, H. Wang, H. He, F. Guo, and L. Dong, *Lysine-specific demethylase 5C promotes hepatocellular carcinoma cell invasion through inhibition BMP7 expression*. BMC Cancer, 2015. **15**: p. 801.
251. Li, N., S.S. Dhar, T.Y. Chen, P.Y. Kan, Y. Wei, J.H. Kim, C.H. Chan, H.K. Lin, M.C. Hung, and M.G. Lee, *JARID1D Is a Suppressor and Prognostic Marker of Prostate Cancer Invasion and Metastasis*. Cancer Res, 2016. **76**(4): p. 831-43.
252. Arseneault, M., J. Monlong, N.S. Vasudev, R.S. Laskar, M. Safisamghabadi, P. Harnden, L. Egevad, N. Nourbehesht, P. Panichnantakul, I. Holcatova, et al., *Loss of chromosome Y leads to down regulation of KDM5D and KDM6C epigenetic modifiers in clear cell renal cell carcinoma*. Sci Rep, 2017. **7**: p. 44876.

253. Jangravi, Z., M.S. Tabar, M. Mirzaei, P. Parsamatin, H. Vakilian, M. Alikhani, M. Shabani, P.A. Haynes, A.K. Goodchild, H. Gourabi, et al., *Two Splice Variants of Y Chromosome-Located Lysine-Specific Demethylase 5D Have Distinct Function in Prostate Cancer Cell Line (DU-145)*. J Proteome Res, 2015. **14**(9): p. 3492-502.
254. Lu, P.J., K. Sundquist, D. Baeckstrom, R. Poulsom, A. Hanby, S. Meier-Ewert, T. Jones, M. Mitchell, P. Pitha-Rowe, P. Freemont, et al., *A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer*. J Biol Chem, 1999. **274**(22): p. 15633-45.
255. Roesch, A., A.M. Mueller, T. Stempf, C. Moehle, M. Landthaler, and T. Vogt, *RBP2-H1/JARID1B is a transcriptional regulator with a tumor suppressive potential in melanoma cells*. Int J Cancer, 2008. **122**(5): p. 1047-57.
256. Dai, B., Z. Hu, H. Huang, G. Zhu, Z. Xiao, W. Wan, P. Zhang, W. Jia, and L. Zhang, *Overexpressed KDM5B is associated with the progression of glioma and promotes glioma cell growth via downregulating p21*. Biochem Biophys Res Commun, 2014. **454**(1): p. 221-7.
257. Wang, D., S. Han, R. Peng, C. Jiao, X. Wang, X. Yang, R. Yang, and X. Li, *Depletion of histone demethylase KDM5B inhibits cell proliferation of hepatocellular carcinoma by regulation of cell cycle checkpoint proteins p15 and p27*. J Exp Clin Cancer Res, 2016. **35**: p. 37.
258. Ohta, K., N. Haraguchi, Y. Kano, Y. Kagawa, M. Konno, S. Nishikawa, A. Hamabe, S. Hasegawa, H. Ogawa, T. Fukusumi, et al., *Depletion of JARID1B induces cellular senescence in human colorectal cancer*. Int J Oncol, 2013. **42**(4): p. 1212-8.
259. Wang, Z., F. Tang, G. Qi, S. Yuan, G. Zhang, B. Tang, and S. He, *KDM5B is overexpressed in gastric cancer and is required for gastric cancer cell proliferation and metastasis*. Am J Cancer Res, 2015. **5**(1): p. 87-100.

260. Xiang, Y., Z. Zhu, G. Han, X. Ye, B. Xu, Z. Peng, Y. Ma, Y. Yu, H. Lin, A.P. Chen, et al., *JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer*. Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19226-31.
261. Wang, L., Y. Mao, G. Du, C. He, and S. Han, *Overexpression of JARID1B is associated with poor prognosis and chemotherapy resistance in epithelial ovarian cancer*. Tumour Biol, 2015. **36**(4): p. 2465-72.
262. Wang, W., K. Zheng, Y. Pei, and X. Zhang, *Histone Demethylase JARID1B Is Overexpressed in Osteosarcoma and Upregulates Cyclin D1 Expression via Demethylation of H3K27me3*. Oncol Res, 2018. **26**(3): p. 373-384.
263. Huang, D., Y. Qiu, G. Li, C. Liu, L. She, D. Zhang, X. Chen, G. Zhu, X. Zhang, Y. Tian, et al., *KDM5B overexpression predicts a poor prognosis in patients with squamous cell carcinoma of the head and neck*. J Cancer, 2018. **9**(1): p. 198-204.
264. Scibetta, A.G., S. Santangelo, J. Coleman, D. Hall, T. Chaplin, J. Copier, S. Catchpole, J. Burchell, and J. Taylor-Papadimitriou, *Functional analysis of the transcription repressor PLU-1/JARID1B*. Mol Cell Biol, 2007. **27**(20): p. 7220-35.
265. Fork, C., L. Gu, J. Hitzel, I. Josipovic, J. Hu, M. SzeKa Wong, Y. Ponomareva, M. Albert, S.U. Schmitz, S. Uchida, et al., *Epigenetic Regulation of Angiogenesis by JARID1B-Induced Repression of HOXA5*. Arterioscler Thromb Vasc Biol, 2015. **35**(7): p. 1645-52.
266. Hayami, S., M. Yoshimatsu, A. Veerakumarasivam, M. Unoki, Y. Iwai, T. Tsunoda, H.I. Field, J.D. Kelly, D.E. Neal, H. Yamaue, et al., *Overexpression of the JmjC histone demethylase KDM5B in human carcinogenesis: involvement in the proliferation of cancer cells through the E2F/RB pathway*. Mol Cancer, 2010. **9**: p. 59.
267. Muller, H. and K. Helin, *The E2F transcription factors: key regulators of cell proliferation*. Biochim Biophys Acta, 2000. **1470**(1): p. M1-12.

268. Bamodu, O.A., W.C. Huang, W.H. Lee, A. Wu, L.S. Wang, M. Hsiao, C.T. Yeh, and T.Y. Chao, *Aberrant KDM5B expression promotes aggressive breast cancer through MALAT1 overexpression and downregulation of hsa-miR-448*, in *BMC Cancer*. 2016.
269. Shen, X., Z. Zhuang, Y. Zhang, Z. Chen, L. Shen, W. Pu, L. Chen, and Z. Xu, *JARID1B modulates lung cancer cell proliferation and invasion by regulating p53 expression*. *Tumour Biol*, 2015. **36**(9): p. 7133-42.
270. Han, L., X.H. Liang, L.X. Chen, S.M. Bao, and Z.Q. Yan, *SIRT1 is highly expressed in brain metastasis tissues of non-small cell lung cancer (NSCLC) and in positive regulation of NSCLC cell migration*. *Int J Clin Exp Pathol*, 2013. **6**(11): p. 2357-65.
271. Tang, B., G. Qi, F. Tang, S. Yuan, Z. Wang, X. Liang, B. Li, S. Yu, J. Liu, Q. Huang, et al., *JARID1B promotes metastasis and epithelial-mesenchymal transition via PTEN/AKT signaling in hepatocellular carcinoma cells*. *Oncotarget*, 2015. **6**(14): p. 12723-39.
272. Yoshida, M., A. Ishimura, M. Terashima, Z. Enkhbaatar, N. Nozaki, K. Satou, and T. Suzuki, *PLU1 histone demethylase decreases the expression of KAT5 and enhances the invasive activity of the cells*. *Biochem J*, 2011. **437**(3): p. 555-64.
273. Enkhbaatar, Z., M. Terashima, D. Oktyabri, S. Tange, A. Ishimura, S. Yano, and T. Suzuki, *KDM5B histone demethylase controls epithelial-mesenchymal transition of cancer cells by regulating the expression of the microRNA-200 family*. *Cell Cycle*, 2013. **12**(13): p. 2100-12.
274. Sharma, S.V., D.Y. Lee, B. Li, M.P. Quinlan, F. Takahashi, S. Maheswaran, U. McDermott, N. Azizian, L. Zou, M.A. Fischbach, et al., *A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations*. *Cell*, 2010. **141**(1): p. 69-80.
275. Hou, J., J. Wu, A. Dombkowski, K. Zhang, A. Holowatyj, J.L. Boerner, and Z.Q. Yang, *Genomic amplification and a role in drug-resistance for the KDM5A histone demethylase in breast cancer*. *Am J Transl Res*, 2012. **4**(3): p. 247-56.

276. Roesch, A., M. Fukunaga-Kalabis, E.C. Schmidt, S.E. Zabierowski, P.A. Brafford, A. Vultur, D. Basu, P. Gimotty, T. Vogt, and M. Herlyn, *A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth*. *Cell*, 2010. **141**(4): p. 583-94.
277. Roesch, A., A. Vultur, I. Bogeski, H. Wang, K.M. Zimmermann, D. Speicher, C. Korbel, M.W. Laschke, P.A. Gimotty, S.E. Philipp, et al., *Overcoming intrinsic multidrug resistance in melanoma by blocking the mitochondrial respiratory chain of slow-cycling JARID1B(high) cells*. *Cancer Cell*, 2013. **23**(6): p. 811-25.
278. Kuo, Y.T., Y.L. Liu, B.O. Adebayo, P.H. Shih, W.H. Lee, L.S. Wang, Y.F. Liao, W.M. Hsu, C.T. Yeh, and C.M. Lin, *JARID1B Expression Plays a Critical Role in Chemoresistance and Stem Cell-Like Phenotype of Neuroblastoma Cells*. *PLoS One*, 2015. **10**(5): p. e0125343.
279. Hinohara, K., H.J. Wu, S. Vigneau, T.O. McDonald, K.J. Igarashi, K.N. Yamamoto, T. Madsen, A. Fassl, S.B. Egri, M. Papanastasiou, et al., *KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance*. *Cancer Cell*, 2018. **34**(6): p. 939-953.e9.
280. Lin, C.S., Y.C. Lin, B.O. Adebayo, A. Wu, J.H. Chen, Y.J. Peng, M.F. Cheng, W.H. Lee, M. Hsiao, T.Y. Chao, et al., *Silencing JARID1B suppresses oncogenicity, stemness and increases radiation sensitivity in human oral carcinoma*. *Cancer Lett*, 2015. **368**(1): p. 36-45.
281. Bayo, J., T.A. Tran, L. Wang, S. Pena-Llopis, A.K. Das, and E.D. Martinez, *Jumonji Inhibitors Overcome Radioresistance in Cancer through Changes in H3K4 Methylation at Double-Strand Breaks*. *Cell Reports*, 2018. **25**(4): p. 1040-1050.e5.
282. Sayegh, J., J. Cao, M.R. Zou, A. Morales, L.P. Blair, M. Norcia, D. Hoyer, A.J. Tackett, J.S. Merkel, and Q. Yan, *Identification of small molecule inhibitors of Jumonji AT-rich interactive domain 1B (JARID1B) histone demethylase by a sensitive high throughput screen*. *J Biol Chem*, 2013. **288**(13): p. 9408-17.

283. Vinogradova, M., V.S. Gehling, A. Gustafson, S. Arora, C.A. Tindell, C. Wilson, K.E. Williamson, G.D. Guler, P. Gangurde, W. Manieri, et al., *An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells*. Nat Chem Biol, 2016. **12**(7): p. 531-8.
284. Tumber, A., A. Nuzzi, E.S. Hookway, S.B. Hatch, S. Velupillai, C. Johansson, A. Kawamura, P. Savitsky, C. Yapp, A. Szykowska, et al., *Potent and Selective KDM5 Inhibitor Stops Cellular Demethylation of H3K4me3 at Transcription Start Sites and Proliferation of MM1S Myeloma Cells*. Cell Chem Biol, 2017. **24**(3): p. 371-380.
285. Bavetsias, V., R.M. Lanigan, G.F. Ruda, B. Atrash, M.G. McLaughlin, A. Tumber, N.Y. Mok, Y.V. Le Bihan, S. Dempster, K.J. Boxall, et al., *8-Substituted Pyrido[3,4-d]pyrimidin-4(3H)-one Derivatives As Potent, Cell Permeable, KDM4 (JMJD2) and KDM5 (JARID1) Histone Lysine Demethylase Inhibitors*. J Med Chem, 2016. **59**(4): p. 1388-409.
286. Liang, J., B. Zhang, S. Labadie, D.F. Ortwine, M. Vinogradova, J.R. Kiefer, V.S. Gehling, J.C. Harmange, R. Cummings, T. Lai, et al., *Lead optimization of a pyrazolo[1,5-a]pyrimidin-7(4H)-one scaffold to identify potent, selective and orally bioavailable KDM5 inhibitors suitable for in vivo biological studies*. Bioorg Med Chem Lett, 2016. **26**(16): p. 4036-41.
287. Thinnes, C.C., K.S. England, A. Kawamura, R. Chowdhury, C.J. Schofield, and R.J. Hopkinson, *Targeting histone lysine demethylases - progress, challenges, and the future*. Biochim Biophys Acta, 2014. **1839**(12): p. 1416-32.
288. Konen, J., E. Summerbell, B. Dwivedi, K. Galior, Y. Hou, L. Rusnak, A. Chen, J. Saltz, W. Zhou, L. Boise, et al., *Image-guided genomics of phenotypically heterogeneous populations reveals vascular signaling during symbiotic collective cancer invasion*. Nature Communications, 2017. **8**: p. 15078.
289. Ladoux, B. and R.M. Mege, *Mechanobiology of collective cell behaviours*. Nat Rev Mol Cell Biol, 2017. **18**(12): p. 743-757.

290. Chen, T., T.B. Saw, R.M. Mege, and B. Ladoux, *Mechanical forces in cell monolayers*. J Cell Sci, 2018. **131**(24).
291. American Cancer, S., *Cancer Facts & Figures*. 2016.
292. Efferth, T. and M. Leber, *Molecular principles of cancer invasion and metastasis (Review)*. International Journal of Oncology, 2009. **34**(4).
293. Westcott, J.M., A.M. Prechtel, E.A. Maine, T.T. Dang, M.A. Esparza, H. Sun, Y. Zhou, Y. Xie, and G.W. Pearson, *An epigenetically distinct breast cancer cell subpopulation promotes collective invasion*. J Clin Invest, 2015. **125**(5): p. 1927-43.
294. Cheung, K.J., E. Gabrielson, Z. Werb, and A.J. Ewald, *Collective invasion in breast cancer requires a conserved basal epithelial program*. Cell, 2013. **155**(7): p. 1639-1651.
295. Kabla, A.J., *Collective cell migration: leadership, invasion and segregation*. J R Soc Interface, 2012. **9**(77): p. 3268-78.
296. Friedl, P., J. Locker, E. Sahai, and J.E. Segall, *Classifying collective cancer cell invasion*. Nat Cell Biol, 2012. **14**(8): p. 777-783.
297. Ewald, A.J., R.J. Huebner, H. Palsdottir, J.K. Lee, M.J. Perez, D.M. Jorgens, A.N. Tauscher, K.J. Cheung, Z. Werb, and M. Auer, *Mammary collective cell migration involves transient loss of epithelial features and individual cell migration within the epithelium*. J Cell Sci, 2012. **125**(Pt 11): p. 2638-54.
298. Gilbert-Ross, M., J. Konen, J. Koo, J. Shupe, B.S. Robinson, W.G.t. Wiles, C. Huang, W.D. Martin, M. Behera, G.H. Smith, et al., *Targeting adhesion signaling in KRAS, LKB1 mutant lung adenocarcinoma*. JCI Insight, 2017. **2**(5): p. e90487.
299. Richardson, A.M., L.S. Havel, A.E. Koyen, J.M. Konen, J. Shupe, W.G. Wiles, W.D. Martin, H.E. Grossniklaus, G. Sica, M. Gilbert-Ross, et al., *Vimentin Is Required for Lung*

- Adenocarcinoma Metastasis via Heterotypic Tumor Cell–Cancer-Associated Fibroblast Interactions during Collective Invasion*. *Clinical Cancer Research*, 2018. **24**(2): p. 420-432.
300. Friedl, P., K.S. Zanker, P. Friedl, P.B. Noble, K.S. Zanker, P.A. Walton, D.W. Laird, P.J. Chauvin, R.J. Tabah, and M. Black, *Migration of Coordinated Cell Clusters in Mesenchymal and Epithelial Cancer Explants in Vitro*. *Cancer Research*, 1995. **55**(20): p. 4557-4560.
301. Hegerfeldt, Y., M. Tusch, E.-b. Bro, and P. Friedl, *Collective Cell Movement in Primary Melanoma Explants*. *Cancer Res*, 2002. **62**: p. 2125-2130.
302. Nabeshima, K., T. Inoue, Y. Shimao, H. Kataoka, and M. Koono, *Cohort migration of carcinoma cells: differentiated colorectal carcinoma cells move as coherent cell clusters or sheets*. *Histol Histopathol.*, 1999. **14**(4): p. 1183-1197.
303. Leighton, J., R. Kalla, J. Turner, and R. Fennell, *Pathogenesis of Tumor Invasion II. Aggregate Replication*. *Cancer Research*, 1960. **20**(575): p. 586.
304. Friedl, P., *Prespecification and plasticity: shifting mechanisms of cell migration*. *Curr Opin Cell Biol*, 2004. **16**(1): p. 14-23.
305. Alexander, S., B. Weigelin, F. Winkler, and P. Friedl, *Preclinical intravital microscopy of the tumour-stroma interface: invasion, metastasis, and therapy response*. *Curr Opin Cell Biol*, 2013. **25**(5): p. 659-71.
306. Friedl, P. and K. Wolf, *Plasticity of cell migration: a multiscale tuning model*. *J Cell Biol*, 2010. **188**(1): p. 11-9.
307. Haeger, A., M. Krause, K. Wolf, and P. Friedl, *Cell jamming: collective invasion of mesenchymal tumor cells imposed by tissue confinement*. *Biochim Biophys Acta*, 2014. **1840**(8): p. 2386-95.
308. van Zijl, F., G. Krupitza, and W. Mikulits, *Initial steps of metastasis: cell invasion and endothelial transmigration*. *Mutat Res*, 2011. **728**(1-2): p. 23-34.

309. Pandya, P., J.L. Orgaz, and V. Sanz-Moreno, *Actomyosin contractility and collective migration: may the force be with you*. *Curr Opin Cell Biol*, 2017. **48**: p. 87-96.
310. Waclaw, B., I. Bozic, M.E. Pittman, R.H. Hruban, B. Vogelstein, and M.A. Nowak, *A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity*. *Nature*, 2015. **525**(7568): p. 261-4.
311. Lipinski, K.A., L.J. Barber, M.N. Davies, M. Ashenden, A. Sottoriva, and M. Gerlinger, *Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine*. *Trends Cancer*, 2016. **2**(1): p. 49-63.
312. McPherson, A., A. Roth, E. Laks, T. Masud, A. Bashashati, A.W. Zhang, G. Ha, J. Biele, D. Yap, A. Wan, et al., *Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer*. *Nat Genet*, 2016. **48**(7): p. 758-67.
313. Rinner, B., B. Galle, S. Trajanoski, C. Fischer, M. Hatz, T. Maierhofer, G. Michelitsch, F. Moinfar, I. Stelzer, R. Pfragner, et al., *Molecular evidence for the bi-clonal origin of neuroendocrine tumor derived metastases*. *BMC Genomics*, 2012. **13**: p. 594.
314. Gerlinger, M., A.J. Rowan, S. Horswell, M. Math, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing*. *N Engl J Med*, 2012. **366**(10): p. 883-892.
315. Goley, E.D. and M.D. Welch, *The ARP2/3 complex: an actin nucleator comes of age*. *Nature reviews. Molecular cell biology*, 2006. **7**(10): p. 713-26.
316. Zheng, H.C., Y.S. Zheng, X.H. Li, H. Takahashi, T. Hara, S. Masuda, X.H. Yang, Y.F. Guan, and Y. Takano, *Arp2/3 overexpression contributed to pathogenesis, growth and invasion of gastric carcinoma*. *Anticancer Research*, 2008. **28**(4 B): p. 2225-2232.

317. Iwaya, K., K. Oikawa, S. Semba, B. Tsuchiya, Y. Mukai, T. Otsubo, T. Nagao, M. Izumi, M. Kuroda, H. Domoto, et al., *Correlation between liver metastasis of the colocalization of actin-related protein 2 and 3 complex and WAVE2 in colorectal carcinoma*. *Cancer Sci*, 2007. **98**(7): p. 992-9.
318. Lv, J., J. Liu, M. Xiao, H. Xu, C. Xu, X. Zhang, L. Tang, F. Jiang, Y. Zhou, Z. Zhang, et al., *ARP3 promotes tumor metastasis and predicts a poor prognosis in hepatocellular carcinoma*. *Pathol Res Pract*, 2018. **214**(9): p. 1356-1361.
319. Yang, Z.-L., X. Miao, L. Xiong, Q. Zou, Y. Yuan, J. Li, L. Liang, M. Chen, and S. Chen, *CFL1 and Arp3 are Biomarkers for Metastasis and Poor Prognosis of Squamous Cell/ Adenosquamous Carcinomas and Adenocarcinomas of Gallbladder*. *Cancer Investigation*, 2013. **31**(2): p. 132-139.
320. Mertins, P., J.W. Qiao, J. Patel, N.D. Udeshi, K.R. Clauser, D.R. Mani, M.W. Burgess, M.A. Gillette, J.D. Jaffe, and S.A. Carr, *Integrated proteomic analysis of post-translational modifications by serial enrichment*. *Nat Methods*, 2013. **10**(7): p. 634-7.
321. Wagner, S.a., P. Beli, B.T. Weinert, M.L. Nielsen, J. Cox, M. Mann, and C. Choudhary, *A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles*. *Molecular & Cellular Proteomics*, 2011. **10**(10): p. M111.013284-M111.013284.
322. Wagner, S.a., P. Beli, B.T. Weinert, C. Scholz, C.D. Kelstrup, C. Young, M.L. Nielsen, J.V. Olsen, C. Brakebusch, and C. Choudhary, *Proteomic analyses reveal divergent ubiquitylation site patterns in murine tissues*. *Molecular & Cellular Proteomics*, 2012: p. 1578-1585.
323. Torres, M.P., H. Dewhurst, and N. Sundararaman, *Proteome-wide Structural Analysis of PTM Hotspots Reveals Regulatory Elements Predicted to Impact Biological Function and Disease*. *Molecular & Cellular Proteomics*, 2016. **15**(11): p. 3513.
324. Qin, J.Y., L. Zhang, K.L. Clift, I. Hular, A.P. Xiang, B.Z. Ren, and B.T. Lahn, *Systematic comparison of constitutive promoters and the doxycycline-inducible promoter*. *PLoS One*, 2010. **5**(5): p. e10611.

325. Han, M., W. Xu, P. Cheng, H. Jin, and X. Wang, *Histone demethylase lysine demethylase 5B in development and cancer*, in *Oncotarget*. 2017. p. 8980-91.
326. Zheng, Y.C., J. Chang, L.C. Wang, H.M. Ren, J.R. Pang, and H.M. Liu, *Lysine demethylase 5B (KDM5B): A potential anti-cancer drug target*. *Eur J Med Chem*, 2019. **161**: p. 131-140.
327. Viré, E., C. Curtis, V. Davalos, A. Git, S. Robson, A. Villanueva, A. Vidal, I. Barbieri, S. Aparicio, M. Esteller, et al., *The Breast Cancer Oncogene EMSY Represses Transcription of Antimetastatic microRNA miR-31*, in *Mol Cell*. 2014. p. 806-18.
328. Kuo, K.T., W.C. Huang, O.A. Bamodu, W.H. Lee, C.H. Wang, M. Hsiao, L.S. Wang, and C.T. Yeh, *Histone demethylase JARID1B/KDM5B promotes aggressiveness of non-small cell lung cancer and serves as a good prognostic predictor*, in *Clin Epigenetics*. 2018.
329. Lu, W., S. Liu, B. Li, Y. Xie, C. Adhiambo, Q. Yang, B.R. Ballard, K.I. Nakayama, R.J. Matusik, and Z. Chen, *SKP2 inactivation suppresses prostate tumorigenesis by mediating JARID1B ubiquitination*. *Oncotarget*, 2015. **6**(2): p. 771-88.
330. Bronsert, P., K. Enderle-Ammour, M. Bader, S. Timme, M. Kuehs, A. Csanadi, G. Kayser, I. Kohler, D. Bausch, J. Hoepfner, et al., *Cancer cell invasion and EMT marker expression: a three-dimensional study of the human cancer-host interface*. *J Pathol*, 2014. **234**(3): p. 410-22.
331. Liu, Z., X. Yang, C. Chen, B. Liu, B. Ren, L. Wang, K. Zhao, S. Yu, and H. Ming, *Expression of the Arp2/3 complex in human gliomas and its role in the migration and invasion of glioma cells*. *Oncol Rep*, 2013. **30**(9): p. 2127-2136.
332. Semba, S., K. Iwaya, J. Matsubayashi, H. Serizawa, H. Kataba, T. Hirano, H. Kato, T. Matsuoka, and K. Mukai, *Coexpression of actin-related protein 2 and Wiskott-Aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung*. *Clin Cancer Res*, 2006. **12**(8): p. 2449-54.

333. Kinoshita, T., N. Nohata, H. Watanabe-Takano, H. Yoshino, H. Hidaka, L. Fujimura, M. Fuse, T. Yamasaki, H. Enokida, M. Nakagawa, et al., *Actin-related protein 2/3 complex subunit 5 (ARPC5) contributes to cell migration and invasion and is directly regulated by tumor-suppressive microRNA-133a in head and neck squamous cell carcinoma*. *Int J Oncol*, 2012. **40**(6): p. 1770-8.
334. Calbo, J., E. van Montfort, N. Proost, E. van Drunen, H.B. Beverloo, R. Meuwissen, and A. Berns, *A Functional Role for Tumor Cell Heterogeneity in a Mouse Model of Small Cell Lung Cancer*. *Cancer Cell*, 2011. **19**(2): p. 244-256.
335. Mateo, F., O. Meca-Cortes, T. Celia-Terrassa, Y. Fernandez, I. Abasolo, L. Sanchez-Cid, R. Bermudo, A. Sagasta, L. Rodriguez-Carunchio, M. Pons, et al., *SPARC mediates metastatic cooperation between CSC and non-CSC prostate cancer cell subpopulations*. *Mol Cancer*, 2014. **13**: p. 237.
336. Tsuji, T., S. Ibaragi, K. Shima, M.G. Hu, M. Katsurano, A. Sasaki, and G.F. Hu, *Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth*. *Cancer Res*, 2008. **68**(24): p. 10377-86.
337. Liu, J.S., J.T. Farlow, A.K. Paulson, M.A. Labarge, and Z.J. Gartner, *Programmed cell-to-cell variability in Ras activity triggers emergent behaviors during mammary epithelial morphogenesis*. *Cell Rep*, 2012. **2**(5): p. 1461-70.
338. Jolly, M.K., J.A. Somarelli, M. Sheth, A. Biddle, S.C. Tripathi, A.J. Armstrong, S.M. Hanash, S.A. Bapat, A. Rangarajan, and H. Levine, *Hybrid epithelial/mesenchymal phenotypes promote metastasis and therapy resistance across carcinomas*. *Pharmacol Ther*, 2018.
339. Abercrombie, M., *Contact inhibition in tissue culture*. *In Vitro*, 1970. **6**(2): p. 128-42.
340. Abercrombie, M. and J.E. Heaysman, *Observations on the social behaviour of cells in tissue culture. I. Speed of movement of chick heart fibroblasts in relation to their mutual contacts*. *Exp Cell Res*, 1953. **5**(1): p. 111-31.

341. Valentin, G., P. Haas, and D. Gilmour, *The chemokine SDF1 α coordinates tissue migration through the spatially restricted activation of Cxcr7 and Cxcr4b*. *Curr Biol*, 2007. **17**(12): p. 1026-31.
342. Dona, E., J.D. Barry, G. Valentin, C. Quirin, A. Khmelinskii, A. Kunze, S. Durdu, L.R. Newton, A. Fernandez-Minan, W. Huber, et al., *Directional tissue migration through a self-generated chemokine gradient*. *Nature*, 2013. **503**(7475): p. 285-9.
343. Horton, J.R., X. Liu, M. Gale, L. Wu, J.R. Shanks, X. Zhang, P.J. Webber, J.S.K. Bell, S.C. Kales, B.T. Mott, et al., *Structural Basis for KDM5A Histone Lysine Demethylase Inhibition by Diverse Compounds*. *Cell Chem Biol*, 2016. **23**(7): p. 769-781.
344. Vazquez-Rodriguez, S., M. Wright, C.M. Rogers, A.P. Cribbs, S. Velupillai, M. Philpott, H. Lee, J.E. Dunford, K.V.M. Huber, M.B. Robers, et al., *Design, Synthesis and Characterization of Covalent KDM5 Inhibitors*. *Angew Chem Int Ed Engl*, 2018.
345. Nie, Z., L. Shi, C. Lai, S.M. O'Connell, J. Xu, R.K. Stansfield, D.J. Hosfield, J.M. Veal, and J.A. Stafford, *Structure-based design and discovery of potent and selective KDM5 inhibitors*. *Bioorg Med Chem Lett*, 2018. **28**(9): p. 1490-1494.
346. Nolen, B.J., N. Tomasevic, A. Russell, D.W. Pierce, Z. Jia, C.D. McCormick, J. Hartman, R. Sakowicz, and T.D. Pollard, *Characterization of two classes of small molecule inhibitors of Arp2/3 complex*. *Nature*, 2009. **460**(7258): p. 1031-4.
347. Ilatovskaya, D.V., V. Chubinskiy-Nadezhdin, T.S. Pavlov, L.S. Shuyskiy, V. Tomilin, O. Palygin, A. Staruschenko, and Y.A. Negulyaev, *Arp2/3 complex inhibitors adversely affect actin cytoskeleton remodeling in the cultured murine kidney collecting duct M-1 cells*. *Cell Tissue Res*, 2013. **354**(3): p. 783-92.
348. Kazazian, K., C. Go, H. Wu, O. Brashavitskaya, R. Xu, J.W. Dennis, A.-C. Gingras, and C.J. Swallow, *Plk4 Promotes Cancer Invasion and Metastasis through Arp2/3 Complex Regulation of the Actin Cytoskeleton*. *Cancer Res*, 2017. **77**(2): p. 1-14.

349. Aceto, N., A. Bardia, D.T. Miyamoto, M.C. Donaldson, B.S. Wittner, J.A. Spencer, M. Yu, A. Pely, A. Engstrom, H. Zhu, et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis*. Cell, 2014. **158**(5): p. 1110-22.
350. Chapman, A., L.F. del Ama, J. Ferguson, J. Kamarashev, C. Wellbrock, and A. Hurlstone, *Heterogeneous tumor subpopulations cooperate to drive invasion*. Cell Reports, 2014. **8**(3): p. 688-695.
351. Zhao, L.H. and H.G. Liu, *Immunohistochemical detection and clinicopathological significance of JARID1B/KDM5B and P16 expression in invasive ductal carcinoma of the breast*. Genet Mol Res, 2015. **14**(2): p. 5417-26.
352. Yamamoto, S., Z. Wu, H.G. Russnes, S. Takagi, G. Peluffo, C. Vaske, X. Zhao, H.K. Moen Vollan, R. Maruyama, M.B. Ekram, et al., *JARID1B is a luminal lineage-driving oncogene in breast cancer*. Cancer Cell, 2014. **25**(6): p. 762-77.
353. Mitra, D., P.M. Das, F.C. Huynh, and F.E. Jones, *Jumonji/ARID1 B (JARID1B) protein promotes breast tumor cell cycle progression through epigenetic repression of microRNA let-7e*. J Biol Chem, 2011. **286**(47): p. 40531-5.
354. Barrett, A., B. Madsen, J. Copier, P.J. Lu, L. Cooper, A.G. Scibetta, J. Burchell, and J. Taylor-Papadimitriou, *PLU-1 nuclear protein, which is upregulated in breast cancer, shows restricted expression in normal human adult tissues: a new cancer/testis antigen?* Int J Cancer, 2002. **101**(6): p. 581-8.
355. Kruidenier, L., C.W. Chung, Z. Cheng, J. Liddle, K. Che, G. Joberty, M. Bantscheff, C. Bountra, A. Bridges, H. Diallo, et al., *A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response*. Nature, 2012. **488**(7411): p. 404-8.
356. Heinemann, B., J.M. Nielsen, H.R. Hudlebusch, M.J. Lees, D.V. Larsen, T. Boesen, M. Labelle, L.O. Gerlach, P. Birk, and K. Helin, *Inhibition of demethylases by GSK-J1/J4*. Nature, 2014. **514**(7520): p. E1-2.

357. Wang, L., J. Chang, D. Varghese, M. Dellinger, S. Kumar, A.M. Best, J. Ruiz, R. Bruick, S. Pena-Llopis, J. Xu, et al., *A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth*. Nat Commun, 2013. **4**: p. 2035.
358. Hatch, S.B., C. Yapp, R.C. Montenegro, P. Savitsky, V. Gamble, A. Tumber, G.F. Ruda, V. Bavetsias, O. Fedorov, B. Atrash, et al., *Assessing histone demethylase inhibitors in cells: lessons learned*. Epigenetics Chromatin, 2017. **10**: p. 9.
359. Johansson, C., A. Tumber, K. Che, P. Cain, R. Nowak, C. Gileadi, and U. Oppermann, *The roles of Jumonji-type oxygenases in human disease*. Epigenomics, 2014. **6**(1): p. 89-120.
360. Van Rechem, C., J.C. Black, M. Boukhali, M.J. Aryee, S. Graslund, W. Haas, C.H. Benes, and J.R. Whetstine, *Lysine demethylase KDM4A associates with translation machinery and regulates protein synthesis*. Cancer Discov, 2015. **5**(3): p. 255-63.
361. Liu, X., S.M. Zhang, M.K. McGeary, I. Krykbaeva, L. Lai, D.J. Jansen, S.C. Kales, A. Simeonov, M.D. Hall, D.P. Kelly, et al., *KDM5B Promotes Drug Resistance by Regulating Melanoma-Propagating Cell Subpopulations*. Mol Cancer Ther, 2019. **18**(3): p. 706-717.
362. Wu, L., J. Cao, W.L. Cai, S.M. Lang, J.R. Horton, D.J. Jansen, Z.Z. Liu, J.F. Chen, M. Zhang, B.T. Mott, et al., *KDM5 histone demethylases repress immune response via suppression of STING*. PLoS Biol, 2018. **16**(8): p. e2006134.
363. Liang, J., S. Labadie, B. Zhang, D.F. Ortwine, S. Patel, M. Vinogradova, J.R. Kiefer, T. Mauer, V.S. Gehling, J.C. Harmange, et al., *From a novel HTS hit to potent, selective, and orally bioavailable KDM5 inhibitors*. Bioorg Med Chem Lett, 2017. **27**(13): p. 2974-2981.
364. Horton, J.R., X. Liu, L. Wu, K. Zhang, J. Shanks, X. Zhang, G. Rai, B.T. Mott, D.J. Jansen, S.C. Kales, et al., *Insights into the Action of Inhibitor Enantiomers against Histone Lysine Demethylase 5A*. J Med Chem, 2018. **61**(7): p. 3193-3208.

365. Kwon, M.C., N. Proost, J.Y. Song, K.D. Sutherland, J. Zevenhoven, and A. Berns, *Paracrine signaling between tumor subclones of mouse SCLC: a critical role of ETS transcription factor Pea3 in facilitating metastasis*. *Genes Dev*, 2015. **29**(15): p. 1587-92.
366. Cuevas, I., H. Layman, L. Coussens, and N. Boudreau, *Sustained endothelial expression of HoxA5 in vivo impairs pathological angiogenesis and tumor progression*. *PLoS One*, 2015. **10**(3): p. e0121720.
367. Steeg, P.S., *Targeting metastasis*. *Nat Rev Cancer*, 2016. **16**(4): p. 201-18.
368. Skrypek, N., S. Goossens, E. De Smedt, N. Vandamme, and G. Berx, *Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity*. *Trends Genet*, 2017. **33**(12): p. 943-959.
369. Taylor, M.J., D. Perrais, and C.J. Merrifield, *A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis*. *PLoS Biol*, 2011. **9**(3): p. e1000604.
370. Andrews, S., *FastQC: A Quality Control tool for High Throughput Sequence Data*. 2010.
371. Bolger, A.M., B. Usadel, and M. Lohse, *Trimmomatic: a flexible trimmer for Illumina sequence data*. *Bioinformatics*, 2014. **30**(15): p. 2114-2120.
372. Dobin, A., C.A. Davis, C. Zaleski, F. Schlesinger, J. Drenkow, M. Chaisson, P. Batut, S. Jha, and T.R. Gingeras, *STAR: ultrafast universal RNA-seq aligner*. *Bioinformatics*, 2012. **29**(1): p. 15-21.
373. Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and S. Genome Project Data Processing, *The Sequence Alignment/Map format and SAMtools*. *Bioinformatics*, 2009. **25**(16): p. 2078-9.
374. Koboldt, D.C., Q. Zhang, D.E. Larson, D. Shen, M.D. McLellan, L. Lin, C.A. Miller, E.R. Mardis, L. Ding, and R.K. Wilson, *VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing*. *Genome Research*, 2012. **22**(3): p. 568-576.

375. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data*. *Nucleic Acids Res*, 2010. **38**(16): p. e164.
376. Konen, J., S. Wilkinson, B. Lee, H. Fu, W. Zhou, Y. Jiang, and A.I. Marcus, *LKB1 kinase-dependent and -independent defects disrupt polarity and adhesion signaling to drive collagen remodeling during invasion*. *Molecular biology of the cell*, 2016. **27**(7): p. mbc.E15-08-0569-1084.
377. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. *Nat Methods*, 2012. **9**(4): p. 357-9.
378. Lawrence, M., W. Huber, H. Pages, P. Aboyoun, M. Carlson, R. Gentleman, M.T. Morgan, and V.J. Carey, *Software for computing and annotating genomic ranges*. *PLoS Comput Biol*, 2013. **9**(8): p. e1003118.
379. Morgan, M., S. Anders, M. Lawrence, P. Aboyoun, H. Pages, and R. Gentleman, *ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data*. *Bioinformatics*, 2009. **25**(19): p. 2607-8.
380. Morgan, M., H. Pagès, V. Obenchain, and N. Hayden, *Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import*. 2017.
381. Team, R.C., *R: A language and environment for statistical computing*. *R Foundation for Statistical Computing*. 2017: Vienna, Austria.
382. Team, T.B.D., *BSgenome.Hsapiens.UCSC.hg19: Full genome sequences for Homo sapiens (UCSC version hg19)*. 2014.
383. Dalvi, M.P., L. Wang, R. Zhong, R.K. Kollipara, H. Park, J. Bayo, P. Yenerall, Y. Zhou, B.C. Timmons, J. Rodriguez-Canales, et al., *Taxane-Platin-Resistant Lung Cancers Co-develop Hypersensitivity to JumonjiC Demethylase Inhibitors*. *Cell Rep*, 2017. **19**(8): p. 1669-1684.