iScience

Article
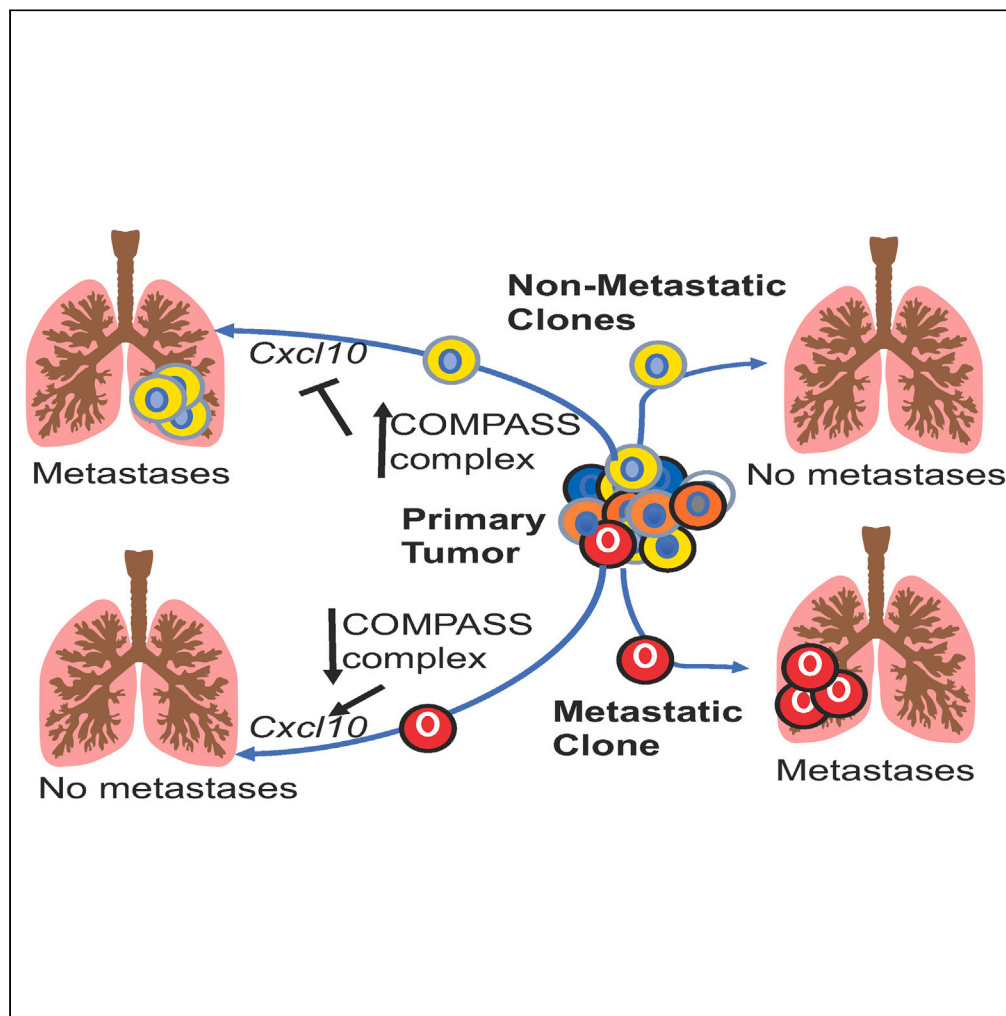
The COMPASS complex maintains the metastatic capacity imparted by a subpopulation of cells in UPS

Ga I. Ban, Vijitha Puviindran, Yu Xiang, ..., Yarui Diao, David G. Kirsch, Benjamin A. Alman

ben.alman@duke.edu

Highlights

Setd1a maintains the metastatic phenotype of a subpopulation of sarcoma cells

Setd1a regulates the expression of genes including *Cxcl10* in metastatic cells

Cxcl10 regulates the metastatic capacity in sarcomas

COMPASS complex inhibition suppresses metastasis in human sarcoma xenografts

Article

The COMPASS complex maintains the metastatic capacity imparted by a subpopulation of cells in UPS

Ga I. Ban,[1] Vijitha Puviindran,[1] Yu Xiang,[2] Puvi Nadesan,[1] Jackie Tang,[1] Jianhong Ou,[2] Nicholas Guardino,[1] Makoto Nakagawa,[1] MaKenna Browne,[2] Asjah Wallace,[1] Koji Ishikawa,[1] Eijiro Shimada,[1] John T. Martin,[1] Yarui Diao,[2] David G. Kirsch,[3,4] and Benjamin A. Alman[1,5,*]

SUMMARY

Intratumoral heterogeneity is common in cancer, particularly in sarcomas like undifferentiated pleomorphic sarcoma (UPS), where individual cells demonstrate a high degree of cytogenic diversity. Previous studies showed that a small subset of cells within UPS, known as the metastatic clone (MC), as responsible for metastasis. Using a CRISPR-based genomic screen _in-vivo_, we identified the COMPASS complex member _Setd1a_ as a key regulator maintaining the metastatic phenotype of the MC in murine UPS. Depletion of _Setd1a_ inhibited metastasis development in the MC. Transcriptome and chromatin sequencing revealed COMPASS complex target genes in UPS, such as _Cxcl10_, downregulated in the MC. Deleting _Cxcl10_ in non-MC cells increased their metastatic potential. Treating mice with human UPS xenografts with a COMPASS complex inhibitor suppressed metastasis without affecting tumor growth in the primary tumor. Our data identified an epigenetic program in a subpopulation of sarcoma cells that maintains metastatic potential.

INTRODUCTION

Cells within tumors exhibit morphologic, cytologic, and genetic heterogeneity.[1] During tumor progression, individual tumor clones may develop genetic or epigenetic alterations responsible for heterogeneity that might endow them with unique phenotypes.[2–4] Such clones could compete over the course of disease causing tumor progression.[5] Despite the characterization of cellular heterogeneity at the morphologic level, the cause of cellular heterogeneity and the biological significance of diverse tumor sub-clones are not well understood.[6–9]

Undifferentiated pleomorphic sarcoma (UPS, previously known as malignant fibrous histiocytoma, or MFH) is the most common soft tissue sarcoma diagnosed in adults. This tumor type is composed of a heterogeneous population of cells with distinct cytologic features.[10–12] UPS displays a complex karyotype, not associated with a known single oncogenic mutation or translocation.[13–17] Although recent investigations interrogated the subtypes and developmental etiology of this tumor type,[18,19] this does not explain intratumoral heterogeneity and has not been translated to improvements in clinical management. Current treatment typically involves wide surgical resection, adjuvant radiation therapy and sometimes chemotherapy. Unfortunately, the 5-year survival is only 50–60%.[14,15,20,21]

The underlying cause of cellular heterogeneity in sarcoma is unclear. However, there is evidence for both genetic and epigenetic regulation of heterogeneity in other tumor types. Most studies of tumor clonal dynamics focus on next generation sequencing of tumor genomes from spatially distinct tumor biopsies.[6,7,22–26] Based on shared somatic mutations, the phylogenetic relationships of tumor clones and their life history can be retrospectively constructed.[27] While genomic analyses of clonal dynamics reveal important insights into tumor evolution, such as the order in which specific somatic mutations can arise, their ability to prospectively track tumor evolution in an unperturbed manner is limited.[28–30] This is especially the case for many solid cancers, where longitudinal sampling is difficult. Interestingly, most of the mutations accrued during genome diversification are likely to be passenger mutations that do not impact phenotype. Indeed, mutations that cf. metastatic capacity or therapy resistance are usually present early during tumorigenesis.[9,31]

Studies using single-cell RNA sequencing demonstrated diversity between cancer cells independent of genetic changes.[25,32,33] This supports the notion that epigenomic changes play a crucial role in driving intratumoral heterogeneity. For instance, histone modification is important in driving functional heterogeneities such as stem-like cell states and variable therapy responses. Additionally, conditional deletion of DNA methyltransferase _Dmnt1_ blocks the development of leukemia by impairing self-renewal of the leukemia initiating cell population,[34]

[1]Department of Orthopedic Surgery, Duke University School of Medicine, Durham, NC, USA
[2]Department of Cell Biology and Duke Regeneration Center, Duke University School of Medicine, Durham, NC, USA
[3]Department of Radiation Oncology, Duke University School of Medicine, Durham, NC, USA
[4]The Princes Margaret Cancer Centre, Department of Radiation Oncology, University Health Network and the University of Toronto, Toronto, ON, Canada
[5]Lead contact
*Correspondence: ben.alman@duke.edu
https://doi.org/10.1016/j.isci.2024.110187

and in genetically homogeneous non-small cell lung cancer cell lines, a small subpopulation of cells that are chemotherapy resistant showed an altered chromatin landscape.[35]

The progression from a primary tumor to metastatic colonization follows an invasion-metastatic cascade.[36] Cancer cells invade into the surrounding tissue and intravasate into the circulation. Cells must suppress anoikis and evade immune detection to survive.[37] After reaching the site of colonization, the circulating cells extravasate into the organ, and may undergo a period of latency.[38] Emergence from latency as cancer cells initiate outgrowth and overtake the native tissue is the final step of metastasis. Whether all or a subset of cancer cells are capable of metastasis is an unanswered question. Emerging evidence is beginning to shed light on this question, and the results appear to be dependent on the type of cancer. Studies of prostate cancer using copy number and whole genome sequencing comparison of matched human metastases and primary tumor implicate a monoclonal origin.[26,39] In contrast, studies in breast and pancreatic cancers, using similar genetic analysis as well as studies of human cancer cells in mice using lentiviral barcoding of stable human cancer cell lines and multi-lineage tracing techniques revealed that multiple clones are able to initiate overt metastases.[9,40–42]

We previously studied clonal dynamics in UPS using a spatially restricted mouse model. Mouse sarcomas mimicking UPS were induced by injection of an adenovirus expressing Cre recombinase (Adeno-Cre) into the gastrocnemius muscle of mice expressing conditional gene mutations to activate a $Kras^{G12D}$ oncogene and to delete the $Trp53$ tumor suppressor gene.[43] Although no single oncogene is universally mutated in human UPS,[43] p53 mutations are frequent. While $KRAS$ mutations are not frequent in human UPS, they can be present, and activation of MAP Kinase signaling downstream of RAS frequently occurs.[44] These murine tumors phenocopy human UPS and exhibit a similar gene expression profile.[45–48]

To label individual tumor clones, we used the ROSA26-Confetti lineage tracing system, and endogenous barcoding with CRISPR-Cas9. This results in a tumor in which a multicolor fluorescent reporter or a barcode labeled individual clones respectively.[48] Study of tumor initiation, local recurrence after therapy, and distant metastasis showed that different clones arising from distinct tumor-initiating cells have different behaviors. Each independent UPS tumor studied contained a clone, termed the metastatic clone (MC), which had a high propensity for developing metastasis. The MC induces lung metastasis more effectively than non-MCs and maintains metastatic capacity over serial transplantations. Furthermore, when examining RNA sequencing data there are distinct differences in gene expression are observed between MC and non-MC cells, and intriguingly, the genetic profile of MCs resembles that of that seen in lung metastases.[48]

Here, we use a CRISPR-based functional in-vivo genomic screen to examine mechanisms responsible for maintaining the metastatic capacity of the MC clone.

RESULTS

Epigenetic modification maintains the metastatic phenotype of the MC

The metastatic phenotype in the MC could be defined genetically by de novo somatic mutations relative to the parent subclone.[5] To determine if this might be the case, whole genome sequencing was utilized to compare MC and non-MC clones from an established mouse UPS sarcoma model.[45,48–50] Sequencing at >50× coverage was performed on nine sorted tumor clones from three primary tumors in which the MC was confirmed in-vivo. We previously showed that the $Kras^{G12D}$ mutation and homozygous $Tp53$ deletions were present in all the clones in this study, consistent with sharing the same founder mutations.[48] Duplex as well as standard detection techniques were utilized. To identify MC-specific copy number variants (CNVs), we generated segmentations for merged results of MC samples using non-MCs as controls, calculated by the Bioconductor package cn.mops (v 1.40.0). p values were determined for the read count numbers of each segmentation using the Exact Fisher-Pitman test from the R package coin (v 1.4.3), and then adjusted using the Benjamini & Hochberg method. The results indicated that copy number alterations are highly similar among different tumor clones. We analyzed MC-specific single nucleotide variants (SNVs) and compared the mutation counts. The findings revealed that recurrent genomic variants across all samples are located in non-coding regions, and the mutation count numbers in autosomes are not different between MCs and non-MCs (Figure S1; Tables S1 and S2). MC specific exon variants were 0.035%. This low percentage of genomic variants is well under the error rate of 0.1% for next generation sequencing platforms.[51] Thus, there was a high degree of similarity among different tumor clones, regardless of metastatic status (data deposited in Gene Expression Omnibus [GEO] repository: GSE208366). The high similarity between the genomes of MCs and non-MCs suggest that genomic variations are not the driving mechanism for enhanced metastatic potential.[52] Therefore, we investigated if epigenetic alteration causes genes expression level changes that induce lung metastasis in UPS.

To identify epigenetic modifiers affecting the capacity to form lung metastasis, we performed an in-vivo CRISPR-Cas9 loss-of-function screen using a lentiviral single guide RNA (sgRNA) library.[53] We used a pooled mouse sgRNA lentiviral library containing 1,816 sgRNAs targeting 454 coding genes and 150 non-targeting control sgRNAs (total 1,966 guide RNAs) The list of genes targeted by the sgRNAs and resultant data are in Tables S3 and S4. Infection of the sgRNA lentiviral library and puromycin selection were performed in MCs, with a minimum coverage of 400 times per sgRNA and an average multiplicity of infection of 0.4. To investigate the sgRNA representation for lung metastasis, we transplanted the cells into mice by tail-vein injection to observe UPS lung metastasis and by intra-muscular injection to observe primary UPS formation into congenic animals. An aliquot of the initial cells was also used for sequencing to determine the sgRNA distribution prior to injection into mice. Cells from intramuscular tumors or pulmonary metastases were collected three weeks after intra-muscular injection or tail vein injection. Genomic DNA was isolated from the tumor and processed using high-throughput sequencing. sgRNA distribution and abundance was analyzed with the mGeCKOa Robust Rank Aggregation algorithm. Specific sgRNAs selected through in vivo pressure were identified. Genes important for lung metastasis were identified by comparing sgRNAs identified in lung metastasis compared to sgRNAs in the primary UPS tumor or in the cells prior to injection (Figure 1A, data deposited in GEO repository GSE208367). The sgRNAs
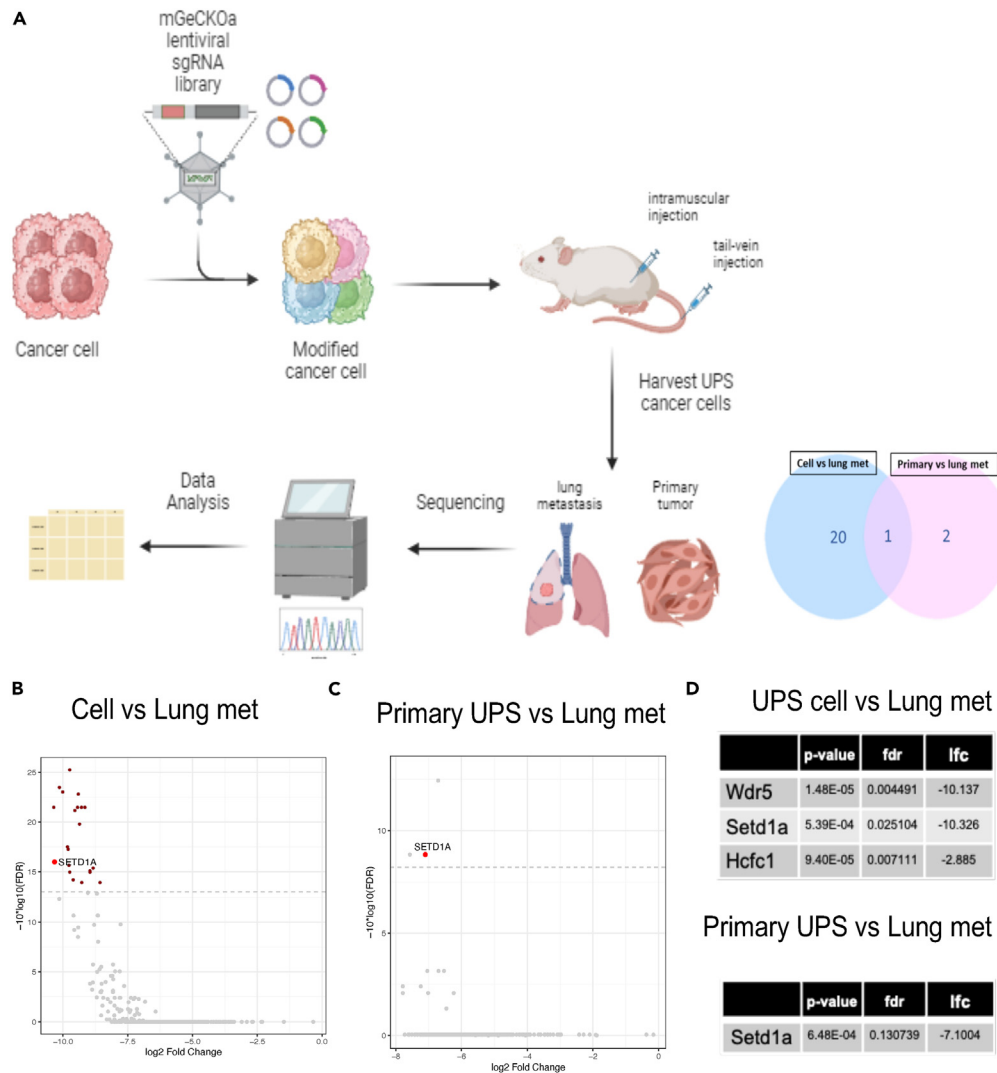
Figure 1. CRISPR-Cas9 screening *in-vivo*

(A) Scheme of CRISPR-Cas9 knock out screening *in-vivo*.

(B) Log2 fold changes of top depleted genes in lung metastasis samples compared with cell sample (FDR < 0.05). *Setd1a* is indicated.

(C) Log2 fold changes of top depleted genes in lung metastasis samples compared with primary tumor sample (FDR < 0.20). *Setd1a* is indicated.

(D) Identified COMPASS complex genes from CRISPR-Cas9 screening. False discovery rate (fdr), Log2 fold change (lfc).

cause loss-of-function and as such sgRNAs that are over represented in the primary tumor compared to the metastasis, or that are over represented in the input cells, target genes whose function potentially enhances metastasis.

We identified 21 targeting sgRNAs when comparing metastasis to input cells (Table S3). The 21 genes included key components of the COMPASS complex, Wdr5, Setd1a, and Hcfc1, that regulate gene transcription and histone H3K4 methylation (Figure 1B). We identified three genes whose function potentially enhances metastasis when comparing metastasis to the primary tumor (FDR < 0.15, p < 0.05, Figure 1C; Table S4). None of non-targeting control sgRNA was differentially identified in the comparisons, showing that this was not a random selection event. The two comparisons (lung metastases vs. primary and lung metastases vs. input cells) found one gene in common, Setd1a (Figure 1D).

The COMPASS complex regulates the metastatic potential of the metastasis initiating cell

Because both screens identified *Setd1a*, we focused on this gene. We next compared the protein levels of Setd1a in MC and non-MC cells and found that Setd1a is expressed at a higher level in the MC (Figure 2A). To determine the role of *Setd1a* in the MC in the capacity to form lung metastasis, we deleted the *Setd1a* gene in MC cells using CRISPR (MC-Setd1aKO, Figure 2B). A commercially available *Setd1a* mouse gene knockout kit which included two gRNAs targeting *Setd1a* was used to confirm that results were not be related to a nuance of the guide RNAs targeting *Setd1a* in the screen. Both of the gRNAs equally downregulated *Setd1a* compared to controls. Clones from both gRNAs showed the
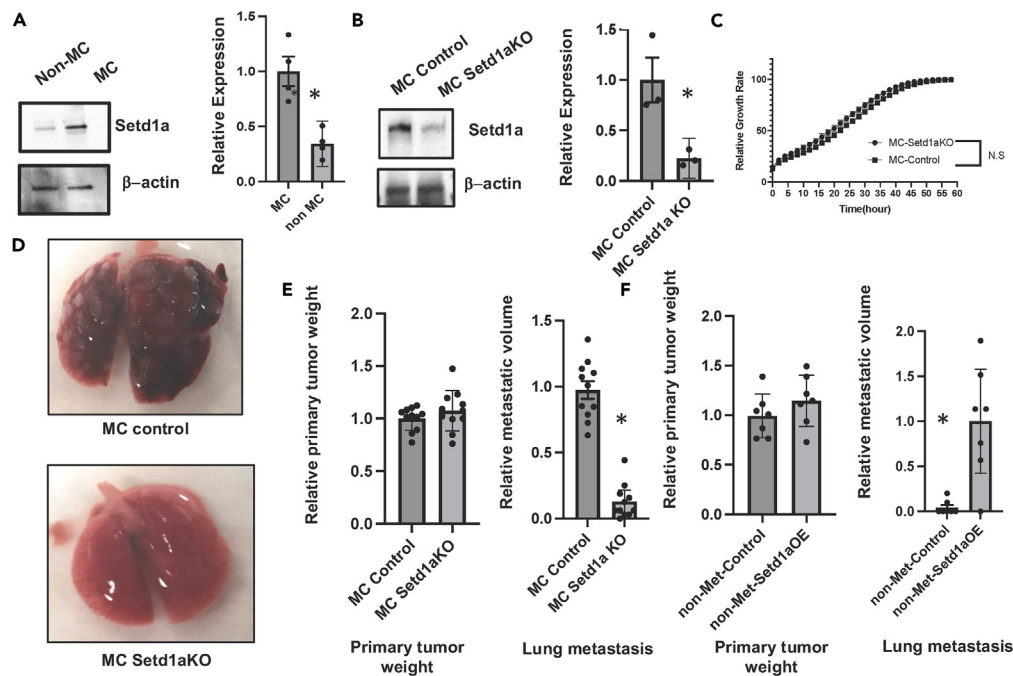
Figure 2. Depletion of Setd1a in the MC inhibits lung metastasis

(A) Higher expression of Setd1a in the MC as assessed by western blot analysis compared with expression in the non-MC. Relative expression with the MC set as "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p = 0.0045$.

(B) Effective depletion of Setd1a by CRISPR was assessed at the protein level by immune blotting Relative expression with the control set as "1." An asterisk indicates statistical significance. $p = 0.0075$.

(C) Growth rates of MC-Setd1aKO and MC-control. The growth rate was measured by cell confluence analysis software using the incuCyte system. No significant differences were identified.

(D) Representative whole lung images from mice implanted with MC-Setd1aKO or MC-control cells. Mets are gray tissue in the lungs in the top image.

(E) Quantification of primary tumor weight and lung metastasis volume from mice injected with MC-Setd1aKO and MC-control. Control cells set as "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p < 0.0001$.

(F) Quantification of primary tumor weight and lung metastasis volume from mice injected with injected with non-MC-Setd1aKO and non-MC-control. Control cells set as "1." An asterisk indicates statistical significance. $p = 0.0017$).

same *in-vivo* results, and we thus combined the results from both clones. Controls were cells treated with the control gRNAs. We measured the growth rate of MC-control and MC-Setd1aKO cells *in-vitro* and found no difference with deletion of *Setd1a* (Figure 2C). Primary tumors were then reconstituted using the MC-Setd1aKO and MC-control cells and the non-MC cells from the primary tumor. Reconstituted tumor cells using the MC modified by CRISPR with a non-targeting control gRNA was used as a control. The reconstituted tumor cells were transplanted into the gastrocnemius muscle of congenic mice. After eight weeks the tumor bearing leg was amputated and the volume of the tumor measured and the mice were analyzed four weeks later for the development of lung metastases similar to our prior work.[48] Deletion of *Setd1a* in the MC population substantially decreased the capacity to form metastases (Figures 2D and 2E). We then examined non-MC cells alone in which we overexpressed *Setd1a* or a control construct in the identical manner. We found no difference in the volume of the primary tumor that formed but a substantial increase in metastatic capacity from non-MC cells overexpressing *Setd1a* (Figure 2F).

The COMPASS complex regulates gene expression in MC cells

The COMPASS complex can regulate histone H3 lysine-4 methylation (H3K4me), which in turn controls gene expression. SETD1A is a core protein in the complex, the absence of which interferes with histone methylation status.[54–57] Mutations in COMPASS complex members are identified in several cancers and enhancer reprogramming related to the complex promotes tumorigenesis in some cancer types.[56,58–62] To determine gene expression changes that occur after loss of *Setd1a*, we undertook RNA sequencing comparing MC-Setd1aKO cells with control cells that retained *Setd1a*. This analysis identified 82 upregulated and 39 downregulated genes in the three independent MC replicates we analyzed (Figures 3A and 3B, data deposited in GEO repository GSE208363). Gene Ontology (GO) analysis showed that chemokine mediated and inflammatory responses were the most differentially regulated processes. Since the COMPASS complex acts through H3K4me to control activity of gene promoters, we used an antibody for H3K4 trimethylation for chromatin immunoprecipitation followed by sequencing (ChIP-seq) to identify the genes whose promoter sequences are regulated by Setd1a.[63,64] ChIP-seq analysis revealed changes
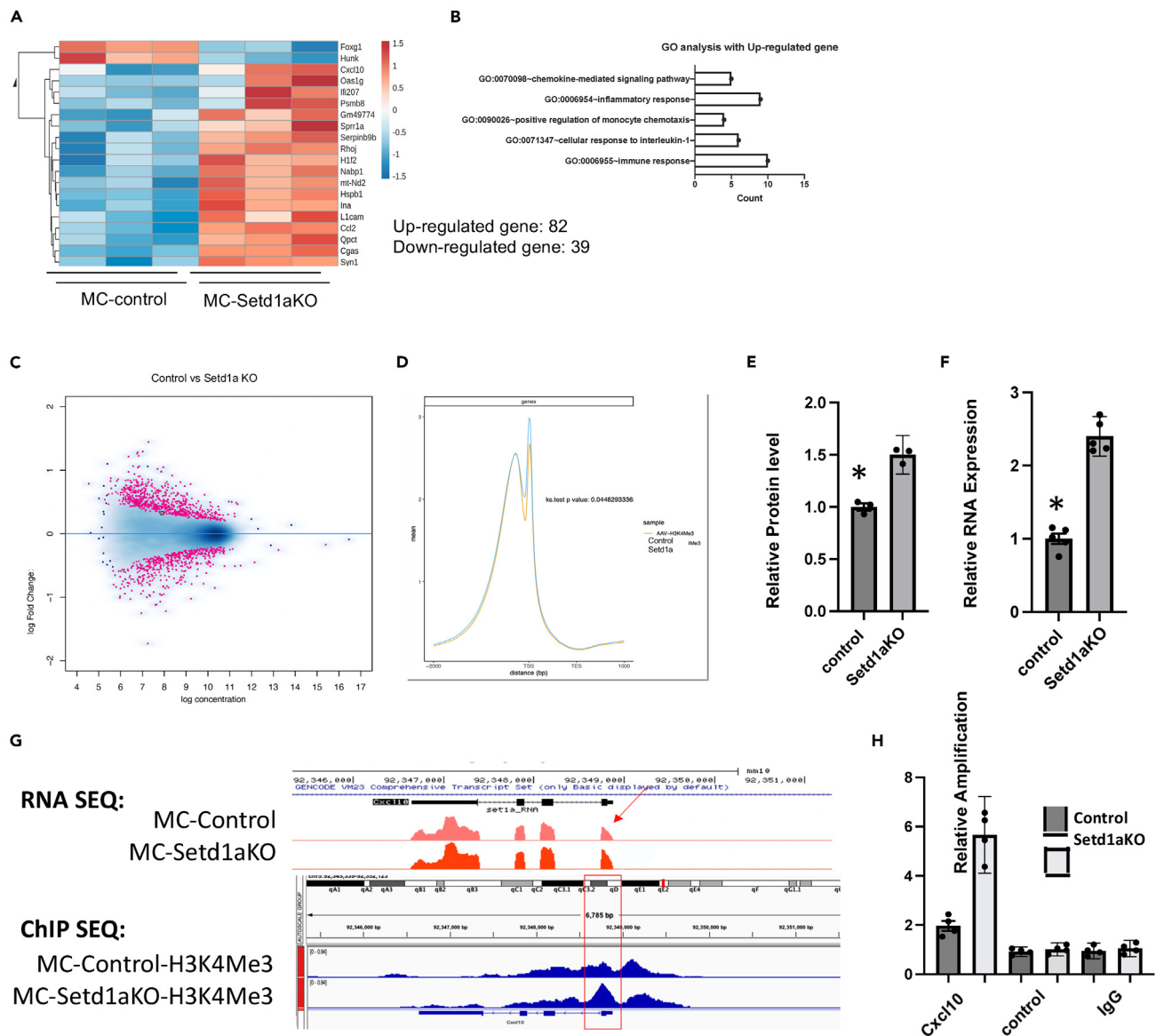
Figure 3. Setd1a changes Cxcl10 expression through H3K4

(A) Differential genes expression analysis of RNA sequencing from MC-Setd1aKO and MC-control displayed as a heatmap. The top genes that were differentially expressed in MC-Setd1aKO compared with MC-control are shown.

(B) GO analysis of down-regulated genes.

(C) Dot plot demonstrating significant differential binding of H3K4me3, plotted using log fold change and log concentration. This illustrates the change in the differential H3K4me3 binding profile. The circled point is Cxcl10.

(D) Average ChIP-seq signal of H3K4me3 at transcription start site from MC-Setd1aKO and MC-control.

(E) Secretion of Cxcl10 protein determined by ELISA. Graph show relative value with control normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p = 0.0008$.

(F) *Cxcl10* expression was measured by RT-PCR. Graph show relative value with control normalized to "1." An asterisk indicates statistical significance. $p < 0.0001$.

(G) The RNA-seq and H3K4me3 profiles at the Cxcl10 gene locus, which shows an increased signal in Cxcl10 transcriptional start sires as indicated by a red arrow or red box.

(H) ChIP-PCR was performed to confirm H3K4me3 occupancy on Cxcl10 gene from ChIP sequencing data. Graph show relative value with IgG control normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p < 0.0001$.

in numbers of copies of the transcription start sites in the *Setd1a* knockout MC cells compared to the control MC and a strong correlation between the gene expression data and ChIP-seq data at genome-wide scales (Figure 3C). The gene most differentially regulated at the RNA level, in which there is changes in ChIP-seq enrichment, and that encodes a gene involved in chemokine mediated and inflammatory

responses was *Cxcl10*[65,66] (Figure 3C). Protein and RNA levels correlated with loss of *Setd1a* (Figures 3E and 3F), and promoter binding in the sequencing data (Figure 3G), was confirmed with ChIP-PCR studies (Figure 3H).

As an additional method to compare the MC and non-MC populations and expression of *Setd1a* and *Cxcl10* we examined previously generated single cell RNA sequencing data on a mouse UPS tumor generated by Trp53 deletion and Kras[G12D] oncogene expression.[67] Sub-populations of cells that exhibited the expression of feature genes of sarcoma cells from an untreated murine UPS sample were analyzed for *Setd1a* and *Cxcl10* expression. This analysis showed that a small subpopulation of cells that expressed *Setd1a* at a high level, and these might be the MC population (Figure S2). There was a negative correlation between cells that cells that expressed *Setd1a* and *Cxcl10* (R = −0.38). We then undertook immunofluorescence to visualize cells producing SETD1A and CXCL10 protein. We found cells expressing SETD1A did not also produce CXCL10. (Figure S3).

Cxcl10 regulates the metastatic potential of the metastasis initiating cell

Since *Cxcl10* expression increased when *Setd1a* was depleted, it would be predicted to joy suppress metastasis. The non-MC cells have a higher level of expression of *Cxcl10* than the MC, consistent with this notion as non-MC cells do not efficiently produce metastases (Figures 4A and 4B). To test for the role of *Cxcl10* in metastatic capacity, we used CRISPR to generate genetically modified non-MC cells in which *Cxcl10* is depleted. A commercially available *Cxcl10* mouse gene knockout kit was utilized. Two gRNAs targeting *Cxcl10* were used, both of which equally downregulated *Setd1a* compared to controls, and thus we combined the data from these two clones. Controls were non-MC cells expressing a non-targeting control sgRNA (Figure 4C). Cells were injected into the limb of mice; the limb was amputated after two months and mice sacrificed 4 weeks later. Animals were analyzed for primary tumor growth and metastatic potential in the same manner as in our previous analyses. There was no difference in the volume of the primary tumor that formed, but the non-MC cells readily formed metastases with *Cxcl10* depletion (Figures 4D and 4F). To determine the role of *Cxcl10* in the MC, we overexpressed *Cxcl10* in MC cells (Figure 4G) in a similar manner as in our *Setd1a* studies. Overexpressing *Cxcl10* in the MC suppressed their metastatic capacity (Figures 4H and 4J).

Taken together, these data are consistent with the notion that the metastasis initiating cell maintains its metastatic potential through an epigenetic mechanism regulated by the COMPASS complex and that one epigenetic regulated gene is *Cxcl10.* Interestingly, the function of CXCL10 and other genes differentially regulated raises the possibility that differences between metastatic and non-metastatic clones is related to expression of inflammatory chemokines.

COMPASS complex inhibition suppresses human UPS xenograft metastasis

Primary human undifferentiated pleomorphic tumors from dissociated cells were established in NOD-SCID gamma null mice as in our previous work.[68] We analyzed tumor growth of a primary tumor from cells injected into the hindlimb and then lung metastases that developed two months after amputation of the primary tumor. 10,000 cells mixed with Matrigel were injected into the hindlimb of the NOD/SCID gamma null mouse. Mice were then treated with a COMPASS complex inhibitor that disrupts the function of member interactions,[69–71] OICR-9429 (S7833; Selleck, Houston, TX), at a dose of 1.5 mg/kg per day[72] administered via intraperitoneal injection starting at four weeks. The limb was amputated at 8 weeks and the primary tumor size determined. Animals were sacrificed for analysis of lung metastasis 8 weeks later. COMPASS complex pharmacologic inhibition did not alter the size of the primary tumor, but significantly reduced metastatic potential (Figure 5A). Interestingly, survival analysis in publicly available human data (The Cancer Genome Atlas [TCGA]) showed that human UPS tumors with *SETD1A* expression above the median expression level had significantly worse survival outcomes ($p = 0.037$, Figure 5B), and a tend toward better survival for tumors expressing *CXCL10* that did not reach statistical significance (Figure 5C). Similar correlation with survival data was not observed for leiomyosarcoma or chondrosarcoma (Figure S4) showing the specificity of this finding to UPS. Since SETD1A would be expected to be expressed at high levels by only a small subpopulation of cells, its expression should be detected in bulk tumor sequencing as the remainder of cells would express the gene at only a low level. In contrast, since *CXCL10* is expected to be expressed by the majority of cells with only a small number having a lower expression level, which would be harder to detect. Finding a trend in data for *CXCL10* is consistent with this notion.

DISCUSSION

Here we identified an epigenetic regulator, the COMPASS complex, that maintains the metastatic phenotype of a subpopulation of UPS cells. Inhibiting the COMPASS complex suppresses metastasis in pre-clinical mouse models. This supports the notion that individual cells in sarcomas have unique properties, and that only some cells can form metastases. Pharmacologic targeting of this epigenetic regulation suppressed metastases in human UPS xenografts, raising the possibility that such an approach could improve clinical outcomes.

The COMPASS complex is a transcriptional regulator required for normal cell function. Dysregulation of the complex function can promote tumorigenesis and mutations in complex members are identified in several cancer types.[56,58–62] For instance, knockdown of *SETD1A* suppress breast cancer formation[73] and aberrant expression of *SETD1A* promotes survival of breast cancer cells.[74] Other members of the COMPASS complex regulate its overall function. WDR5 and HCFC1 act as co-factors to facilitate the function of Setd1a, thereby regulating gene expression.[75] As such, these factors can have a similar function as SETD1A, or could cooperate together as epigenetic regulators. The identification of additional COMPASS complex components in our functional screen strengthens the concept that mechanism regulates metastatic potential.
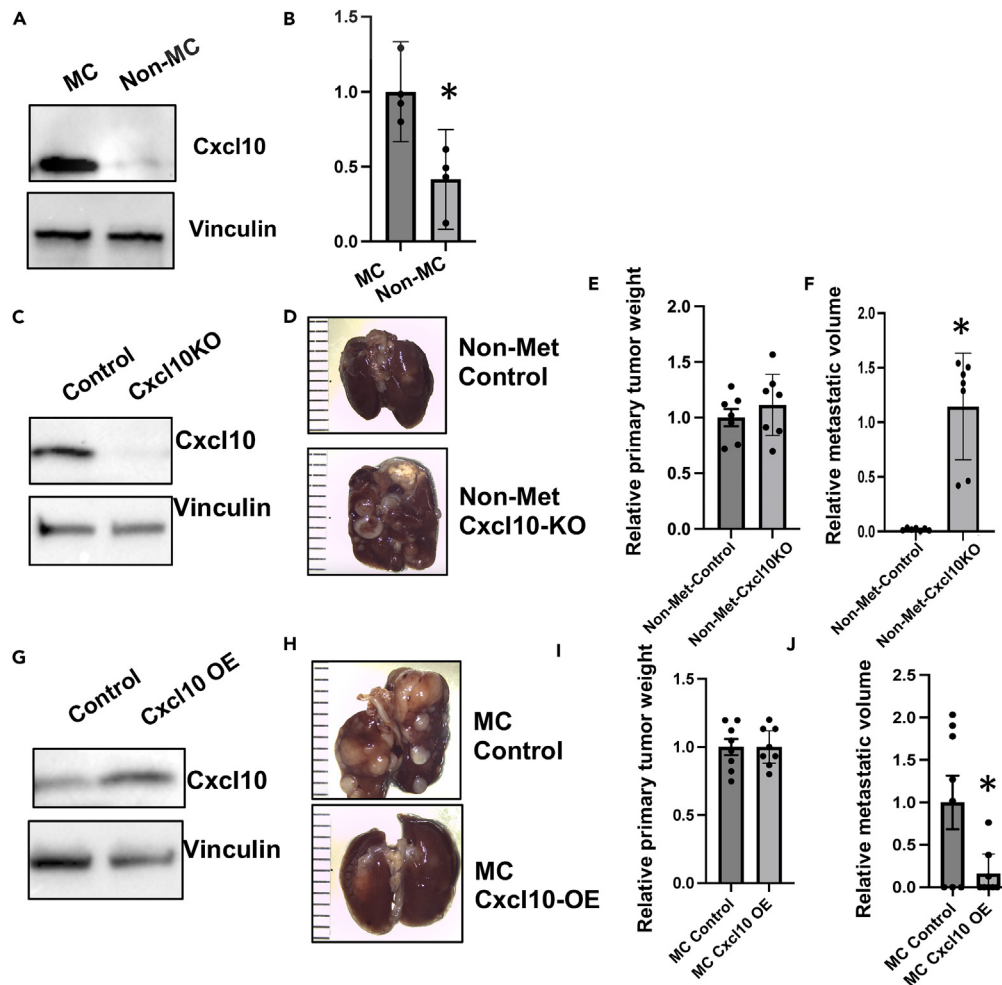
Figure 4. *Cxcl10* suppresses lung metastasis

(A) Cxcl10 expression in the MC and non-MC cells measured using western blot analysis.

(B) Graph show relative value with MC samples normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p = 0.01$.

(C) Cxcl10 expression in *Cxcl10* ko-non-met cells was measured with western blot analysis.

(D) Representative images of lungs showing an increased lung metastasis volume from mice implanted with Cxcl10ko-non-met cells and control cells. The gray colored tissues are the metastases. Scale bar shows millimeters.

(E) Quantification of primary tumor weight and lung metastasis volume (F), with control cells normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p < 0.0001$.

(G) Overexpression in MC cells expressing a *Cxcl10* expression construct.

(H) Representative images of mouse lungs.

(I and J) Quantification of primary tumor weight (I) and lung metastasis volume (J), with control cells normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p < 0.02$.

The COMPASS complex's function is not fully elucidated. While the complex regulates methylation of H3K4, recent data identified additional mechanisms by which it can regulate transcription, such as through pausing RNA polymerase II and regulating transcriptional elongation.[76] Additionally, COMPASS complex members can regulate individual gene transcription without altering global methylation.[75,77] Interestingly CXCL10 was recently shown to interact with COMPASS complex members and the glucocorticoid receptor in co-regulating transcription, suggesting a feedback mechanism.[78] Thus, there are many mechanisms by which the COMPASS complex regulates transcription and much remains to be learned about how it regulates transcription. Our data show that in UPS, the COMPASS complex downregulates expression of *Cxcl10* through regulation of a local H3K4 methylation site near the *Cxcl10* transcriptional start site. CXCL10 has been studied in cancer, and is reported to have diverse effects. It has a tumor inhibitory effect in ovarian and gastric cancer but a tumor-promoting effect in pancreatic and renal cancer.[79,80]

Insights into if all or only a subset of cells can form metastases are beginning to be generated. Studies of prostate cancer using copy number and whole genome sequencing comparison of matched human metastases and primary tumor implicate a monoclonal origin.[26,39] In
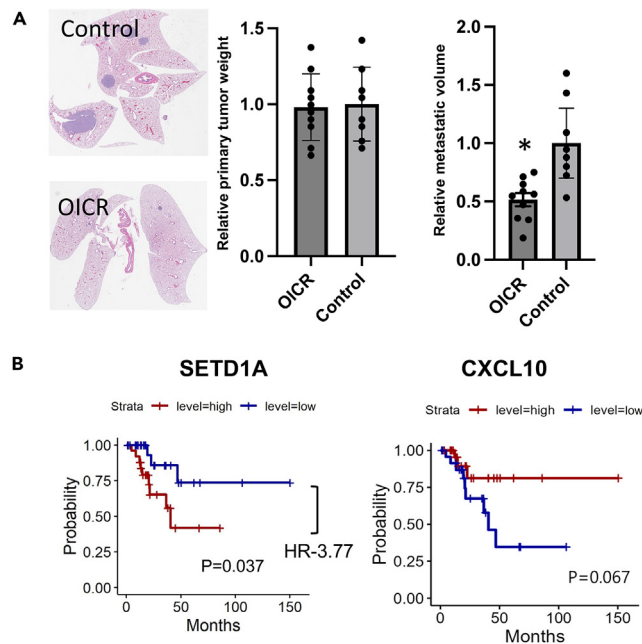
Figure 5. Inhibitor targeting the COMPASS complex suppresses lung metastasis of human UPS cells

(A) Representative H&E staining of lung tissue from mice treated with OICR-9429 and vehicle control. Quantification of lung metastasis volume with control cells normalized to "1." Each data point, means and 95% confidence intervals are shown. An asterisk indicates statistical significance. $p = 0.0018$.

(B) Kaplan-Meier survival curve and hazard ratio of SETD1A gene. The RNA sequencing data in human UPS tumors from The Cancer Genome Atlas was used.

contrast, studies in breast and pancreatic cancers, using similar genetic analysis as well as studies of human cancer cells in mice using lentiviral barcoding of stable human cancer cell lines and multi-lineage tracing techniques in mice revealed that multiple clones are able to initiate overt metastases.[9,40–42] In breast cancer, the activity of phosphoglycerate dehydrogenase in individual cells regulates metastasis potential.[81] Data in sarcoma shows that while there are differences in gene expression, genetic changes in the primary tumors are similar in metastases.[82,83] These later data are consistent with our findings and support the notion that a subpopulation (or clone) of cells is epigenetically conferred with the capacity to form metastasis. We found that the COMPASS complex is activated in only select cell subpopulations. While other cancers have mutations in COMPASS complex members or show COMPASS complex activation,[56,58–62,75,84] we are not are of another study in which the COMPASS complex is activated in onlys a subpopulation of tumor cells. While our genome sequencing did not identify significant differences between the clonal populations, and in particular no recurrent copy number variations of mutations in coding regions, it is still possible that subtle genetic changes could be responsible for the observed phenotype. Nevertheless, our functional screens did identify a functional role for an epigenetic regulator in the MC cell population phenotype, and regardless of other factors which may influence metastasis, this epigenetic mechanism does play a role maintaining the MC.

Studies identifying the MC population were undertaken using murine UPS tumors. It is difficult to investigate human tumors in the same manner as we examined the mouse tumors by labeling clones at tumor initiation. This is a limitation of our study. Furthermore, given the genetic heterogeneity in human UPS tumors, mouse models cannot exactly copy the genetic situation in human tumors. Data from TCGA revealing that the expression of SETD1A and CXCL10 correlates with survival in human UPS is consistent with a role for this mechanism in human tumors. This is further supported but our xenograft data. However, the effects we observed may be not be generalizable to all sarcoma types. In support of this, survival analysis using human chondrosarcoma and leiomyosarcoma data from TCGA did not show a correlation with survival.

In the MC, the COMPASS complex activation regulates *Cxcl10* expression. *CXCL10* expression levels are associated with better prognosis in some tumor types, which is thought to be related to the ability to activate immune responses.[65,66,85–87] In UPS, the COMPASS complex inhibition of *Cxcl10* and other related genes, might allow this cell subpopulation to evade immune surveillance in the lung to become established as a metastasis. Our prior data showing that the MC population and non-MC populations from murine tumors can both enter the circulation,[48] raises the notion that the immune modulation in the metastatic site is important in sarcoma metastatic ability. Indeed, the microenvironment of the lung[88] in concert with lower chemokine mediated inflammatory response in the MC regulated by the COMPASS complex, may allow UPS cells to preferentially metastasize to the lung.

Limitations of the study

A limitation of our study is that experiments on the MC population were undertaken using murine UPS tumors. It is difficult to investigate human tumors in the same manner as we examined the mouse tumors by labeling clones at tumor initiation. Thus it is unclear how our results pertain to human tumors.

There is not a characterized mutation that universally causes UPS. As such, while the mouse model we utilized is commonly used to study UPS, it does not completely copy the human disease.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
- METHOD DETAILS
  - Protein analysis
  - Lentiviral transduction and CRISPR modification
  - Tumor analysis
  - ChIP sequencing
  - RNA-sequencing
  - Whole genome variant analysis
  - Survival analysis from cancer genome atlas data
  - CRISPR screen
  - Single cell RNA sequencing analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110187.

AUTHOR CONTRIBUTIONS

G.I.B., D.G.K., and B.A.A. designed the study. G.I.B., V.P., Y.X., P.N., J.T., N.G., M.N., M.B., A.W., K.I., and E.S. performed the experiments and collected the data. G.I.B., B.A.A., Y.D., and J.T.M. interpreted the results. G.I.B., Y.X., J.T.M., and J.O. performed statistical analysis. G.I.B. and B.A.A. drafted the initial manuscript which was read and modified by all of the authors. All the authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Almendro, V., Marusyk, A., and Polyak, K. (2013). Cellular heterogeneity and molecular evolution in cancer. Annu. Rev. Pathol. 8, 277–302. https://doi.org/10.1146/annurev-pathol-020712-163923.

2. Swanton, C. (2012). Intratumor heterogeneity: evolution through space and time. Cancer Res. 72, 4875–4882. https://doi.org/10.1158/0008-5472.CAN-12-2217.

3. Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer 12, 323–334. https://doi.org/10.1038/nrc3261.

4. McGranahan, N., and Swanton, C. (2017). Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 168, 613–628. https://doi.org/10.1016/j.cell.2017.01.018.

5. Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature 481, 306–313. https://doi.org/10.1038/nature10762.

6. Gerlinger, M., Rowan, A.J., Horswell, S., Math, M., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med. 366, 883–892. https://doi.org/10.1056/NEJMoa1113205.

7. de Bruin, E.C., McGranahan, N., Mitter, R., Salm, M., Wedge, D.C., Yates, L., Jamal-Hanjani, M., Shafi, S., Murugaesu, N., Rowan, A.J., et al. (2014). Spatial and temporal diversity in genomic instability processes defines lung cancer evolution.

Science *346*, 251–256. https://doi.org/10.1126/science.1253462.

8. Gulati, S., Martinez, P., Joshi, T., Birkbak, N.J., Santos, C.R., Rowan, A.J., Pickering, L., Gore, M., Larkin, J., Szallasi, Z., et al. (2014). Systematic evaluation of the prognostic impact and intratumour heterogeneity of clear cell renal cell carcinoma biomarkers. Eur. Urol. *66*, 936–948. https://doi.org/10.1016/j.eururo.2014.06.053.

9. Makohon-Moore, A.P., Zhang, M., Reiter, J.G., Bozic, I., Allen, B., Kundu, D., Chatterjee, K., Wong, F., Jiao, Y., Kohutek, Z.A., et al. (2017). Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. Nat. Genet. *49*, 358–366. https://doi.org/10.1038/ng.3764.

10. Orndal, C., Rydholm, A., Willén, H., Mitelman, F., and Mandahl, N. (1994). Cytogenetic intratumor heterogeneity in soft tissue tumors. Cancer Genet. Cytogenet. *78*, 127–137.

11. Francis, P., Fernebro, J., Edén, P., Laurell, A., Rydholm, A., Domanski, H.A., Breslin, T., Hegardt, C., Borg, A., and Nilbert, M. (2005). Intratumor versus intertumor heterogeneity in gene expression profiles of soft-tissue sarcomas. Genes Chromosomes Cancer *43*, 302–308. https://doi.org/10.1002/gcc.20191.

12. Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105–111.

13. Matushansky, I., Charytonowicz, E., Mills, J., Siddiqi, S., Hricik, T., and Cordon-Cardo, C. (2009). MFH classification: differentiating undifferentiated pleomorphic sarcoma in the 21st Century. Expert Rev. Anticancer Ther. *9*, 1135–1144. https://doi.org/10.1586/era.09.76.

14. Lehnhardt, M., Daigeler, A., Homann, H.H., Schwaiberger, V., Goertz, O., Kuhnen, C., and Steinau, H.U. (2009). MFH revisited: outcome after surgical treatment of undifferentiated pleomorphic or not otherwise specified (NOS) sarcomas of the extremities – an analysis of 140 patients. Langenbeck's Arch. Surg. *394*, 313–320. https://doi.org/10.1007/s00423-008-0368-5.

15. Nascimento, A.F., and Raut, C.P. (2008). Diagnosis and management of pleomorphic sarcomas (so-called "MFH") in adults. J. Surg. Oncol. *97*, 330–339. https://doi.org/10.1002/jso.20972.

16. Matushansky, I., Hernando, E., Socci, N.D., Mills, J.E., Matos, T.A., Edgar, M.A., Singer, S., Maki, R.G., and Cordon-Cardo, C. (2007). Derivation of sarcomas from mesenchymal stem cells via inactivation of the Wnt pathway. J. Clin. Invest. *117*, 3248–3257. https://doi.org/10.1172/JCI31377.

17. Kresse, S.H., Ohnstad, H.O., Bjerkehagen, B., Myklebost, O., and Meza-Zepeda, L.A. (2010). DNA copy number changes in human malignant fibrous histiocytomas by array comparative genomic hybridisation. PLoS One *5*, e15378. https://doi.org/10.1371/journal.pone.0015378.

18. Steele, C.D., Tarabichi, M., Oukrif, D., Webster, A.P., Ye, H., Fittall, M., Lombard, P., Martincorena, I., Tarpey, P.S., Collord, G., et al. (2019). Undifferentiated Sarcomas Develop through Distinct Evolutionary Pathways. Cancer Cell *35*, 441–456.e8. https://doi.org/10.1016/j.ccell.2019.02.002.

19. Blum, J.M., Añó, L., Li, Z., Van Mater, D., Bennett, B.D., Sachdeva, M., Lagutina, I., Zhang, M., Mito, J.K., Dodd, L.G., et al. (2013). Distinct and overlapping sarcoma subtypes initiated from muscle stem and progenitor cells. Cell Rep. *5*, 933–940. https://doi.org/10.1016/j.celrep.2013.10.020.

20. Fletcher, C.D. (1992). Pleomorphic malignant fibrous histiocytoma: fact or fiction? A critical reappraisal based on 159 tumors diagnosed as pleomorphic sarcoma. Am. J. Surg. Pathol. *16*, 213–228.

21. Hollowood, K., and Fletcher, C.D. (1995). Malignant fibrous histiocytoma: morphologic pattern or pathologic entity? Semin. Diagn. Pathol. *12*, 210–220.

22. Yates, L.R., Gerstung, M., Knappskog, S., Desmedt, C., Gundem, G., Van Loo, P., Aas, T., Alexandrov, L.B., Larsimont, D., Davies, H., et al. (2015). Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat. Med. *21*, 751–759. https://doi.org/10.1038/nm.3886.

23. Boutros, P.C., Fraser, M., Harding, N.J., de Borja, R., Trudel, D., Lalonde, E., Meng, A., Hennings-Yeomans, P.H., McPherson, A., Sabelnykova, V.Y., et al. (2015). Spatial genomic heterogeneity within localized, multifocal prostate cancer. Nat. Genet. *47*, 736–745. https://doi.org/10.1038/ng.3315.

24. Hao, J.J., Lin, D.C., Dinh, H.Q., Mayakonda, A., Jiang, Y.Y., Chang, C., Jiang, Y., Lu, C.C., Shi, Z.Z., Xu, X., et al. (2016). Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. Nat. Genet. *48*, 1500–1507. https://doi.org/10.1038/ng.3683.

25. Wang, Y., Waters, J., Leung, M.L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., et al. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature *512*, 155–160. https://doi.org/10.1038/nature13600.

26. Gundem, G., Van Loo, P., Kremeyer, B., Alexandrov, L.B., Tubio, J.M.C., Papaemmanuil, E., Brewer, D.S., Kallio, H.M.L., Högnäs, G., Annala, M., et al. (2015). The evolutionary history of lethal metastatic prostate cancer. Nature *520*, 353–357. https://doi.org/10.1038/nature14347.

27. Schwartz, R., and Schäffer, A.A. (2017). The evolution of tumour phylogenetics: principles and practice. Nat. Rev. Genet. *18*, 213–229. https://doi.org/10.1038/nrg.2016.170.

28. Landau, D.A., Carter, S.L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M.S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., et al. (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell *152*, 714–726. https://doi.org/10.1016/j.cell.2013.01.019.

29. Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L.A., et al. (2012). Mutational processes molding the genomes of 21 breast cancers. Cell *149*, 979–993. https://doi.org/10.1016/j.cell.2012.04.024.

30. Welch, J.S. (2014). Mutation position within evolutionary subclonal architecture in AML. Semin. Hematol. *51*, 273–281. https://doi.org/10.1053/j.seminhematol.2014.08.004.

31. Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.e.C., Krishnamurthy Radhakrishna, V., et al. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition.

Nat. Med. *22*, 262–269. https://doi.org/10.1038/nm.4040.

32. Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94. https://doi.org/10.1038/nature09807.

33. Zack, T.I., Schumacher, S.E., Carter, S.L., Cherniack, A.D., Saksena, G., Tabak, B., Lawrence, M.S., Zhsng, C.Z., Wala, J., Mermel, C.H., et al. (2013). Pan-cancer patterns of somatic copy number alteration. Nat. Genet. *45*, 1134–1140. https://doi.org/10.1038/ng.2760.

34. Trowbridge, J.J., Sinha, A.U., Zhu, N., Li, M., Armstrong, S.A., and Orkin, S.H. (2012). Haploinsufficiency of Dnmt1 impairs leukemia stem cell function through derepression of bivalent chromatin domains. Genes Dev. *26*, 344–349. https://doi.org/10.1101/gad.184341.111.

35. Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell *141*, 69–80. https://doi.org/10.1016/j.cell.2010.02.027.

36. Lambert, A.W., Pattabiraman, D.R., and Weinberg, R.A. (2017). Emerging Biological Principles of Metastasis. Cell *168*, 670–691. https://doi.org/10.1016/j.cell.2016.11.037.

37. Douma, S., Van Laar, T., Zevenhoven, J., Meuwissen, R., Van Garderen, E., and Peeper, D.S. (2004). Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB. Nature *430*, 1034–1039. https://doi.org/10.1038/nature02765.

38. Malladi, S., Macalinao, D.G., Jin, X., He, L., Basnet, H., Zou, Y., de Stanchina, E., and Massagué, J. (2016). Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. Cell *165*, 45–60. https://doi.org/10.1016/j.cell.2016.02.025.

39. Liu, W., Laitinen, S., Khan, S., Vihinen, M., Kowalski, J., Yu, G., Chen, L., Ewing, C.M., Eisenberger, M.A., Carducci, M.A., et al. (2009). Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. Nat. Med. *15*, 559–565. https://doi.org/10.1038/nm.1944.

40. Wagenblast, E., Soto, M., Gutiérrez-Ángel, S., Hartl, C.A., Gable, A.L., Maceli, A.R., Erard, N., Williams, A.M., Kim, S.Y., Dickopf, S., et al. (2015). A model of breast cancer heterogeneity reveals vascular mimicry as a driver of metastasis. Nature *520*, 358–362. https://doi.org/10.1038/nature14403.

41. Cheung, K.J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J.D., Fairchild, A.N., Gorin, M.A., Verdone, J.E., Pienta, K.J., Bader, J.S., and Ewald, A.J. (2016). Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. Proc. Natl. Acad. Sci. USA *113*, E854–E863. https://doi.org/10.1073/pnas.1508541113.

42. Maddipati, R., and Stanger, B.Z. (2015). Pancreatic Cancer Metastases Harbor Evidence of Polyclonality. Cancer Discov. *5*, 1086–1097. https://doi.org/10.1158/2159-8290.CD-15-0120.

43. Cancer Genome Atlas Research Network. Electronic address: elizabeth.demicco@sinaihealthsystem.ca; Cancer Genome Atlas Research Network (2017). Comprehensive and Integrated

Genomic Characterization of Adult Soft Tissue Sarcomas. Cell *171*, 950–965.e28. https://doi.org/10.1016/j.cell.2017.10.014.

44. Serrano, C., Romagosa, C., Hernández-Losa, J., Simonetti, S., Valverde, C., Moliné, T., Somoza, R., Pérez, M., Vélez, R., Vergés, R., et al. (2016). RAS/MAPK pathway hyperactivation determines poor prognosis in undifferentiated pleomorphic sarcomas. Cancer *122*, 99–107. https://doi.org/10.1002/cncr.29733.

45. Mito, J.K., Riedel, R.F., Dodd, L., Lahat, G., Lazar, A.J., Dodd, R.D., Stangenberg, L., Eward, W.C., Hornicek, F.J., Yoon, S.S., et al. (2009). Cross species genomic analysis identifies a mouse model as undifferentiated pleomorphic sarcoma/malignant fibrous histiocytoma. PLoS One *4*, e8075. https://doi.org/10.1371/journal.pone.0008075.

46. Sachdeva, M., Mito, J.K., Lee, C.L., Zhang, M., Li, Z., Dodd, R.D., Cason, D., Luo, L., Ma, Y., Van Mater, D., et al. (2014). MicroRNA-182 drives metastasis of primary sarcomas by targeting multiple genes. J. Clin. Invest. *124*, 4305–4319. https://doi.org/10.1172/JCI77116.

47. Eisinger-Mathason, T.S.K., Zhang, M., Qiu, Q., Skuli, N., Nakazawa, M.S., Karakasheva, T., Mucaj, V., Shay, J.E.S., Stangenberg, L., Sadri, N., et al. (2013). Hypoxia-dependent modification of collagen networks promotes sarcoma metastasis. Cancer Discov. *3*, 1190–1205. https://doi.org/10.1158/2159-8290.CD-13-0118.

48. Tang, Y.J., Huang, J., Tsushima, H., Ban, G.I., Zhang, H., Oristian, K.M., Puviindran, V., Williams, N., Ding, X., Ou, J., et al. (2019). Tracing Tumor Evolution in Sarcoma Reveals Clonal Origin of Advanced Metastasis. Cell Rep. *28*, 2837–2850.e5. https://doi.org/10.1016/j.celrep.2019.08.029.

49. Huang, J., Chen, M., Whitley, M.J., Kuo, H.C., Xu, E.S., Walens, A., Mowery, Y.M., Van Mater, D., Eward, W.C., Cardona, D.M., et al. (2017). Generation and comparison of CRISPR-Cas9 and Cre-mediated genetically engineered mouse models of sarcoma. Nat. Commun. *8*, 15999. https://doi.org/10.1038/ncomms15999.

50. Kirsch, D.G., Dinulescu, D.M., Miller, J.B., Grimm, J., Santiago, P.M., Young, N.P., Nielsen, G.P., Quade, B.J., Chaber, C.J., Schultz, C.P., et al. (2007). A spatially and temporally restricted mouse model of soft tissue sarcoma. Nat. Med. *13*, 992–997. https://doi.org/10.1038/nm1602.

51. Schmitt, M.W., Kennedy, S.R., Salk, J.J., Fox, E.J., Hiatt, J.B., and Loeb, L.A. (2012). Detection of ultra-rare mutations by next-generation sequencing. Proc. Natl. Acad. Sci. USA *109*, 14508–14513. https://doi.org/10.1073/pnas.1208715109.

52. Fox, E.J., Reid-Bayliss, K.S., Emond, M.J., and Loeb, L.A. (2014). Accuracy of Next Generation Sequencing Platforms. Next Gener. Seq. Appl. *1*, 1000106. https://doi.org/10.4172/jngsa.1000106.

53. DeLuca, S., and Bursac, N. (2022). CRISPR Library Screening in Cultured Cardiomyocytes. Methods Mol. Biol. *2485*, 1–13. https://doi.org/10.1007/978-1-0716-2261-2_1.

54. Yao, J., Chen, J., Li, L.Y., and Wu, M. (2020). Epigenetic plasticity of enhancers in cancer. Transcription *11*, 26–36. https://doi.org/10.1080/21541264.2020.1713682.

55. Qu, Q., Takahashi, Y.H., Yang, Y., Hu, H., Zhang, Y., Brunzelle, J.S., Couture, J.F., Shilatifard, A., and Skiniotis, G. (2018). Structure and Conformational Dynamics of a COMPASS Histone H3K4 Methyltransferase Complex. Cell *174*, 1117–1126.e12. https://doi.org/10.1016/j.cell.2018.07.020.

56. Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu. Rev. Biochem. *81*, 65–95. https://doi.org/10.1146/annurev-biochem-051710-134100.

57. Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl. Acad. Sci. USA *98*, 12902–12907. https://doi.org/10.1073/pnas.231473398.

58. Neilsen, B.K., Chakraborty, B., McCall, J.L., Frodyma, D.E., Sleightholm, R.L., Fisher, K.W., and Lewis, R.E. (2018). WDR5 supports colon cancer cells by promoting methylation of H3K4 and suppressing DNA damage. BMC Cancer *18*, 673. https://doi.org/10.1186/s12885-018-4580-6.

59. Dhar, S.S., Zhao, D., Lin, T., Gu, B., Pal, K., Wu, S.J., Alam, H., Lv, J., Yun, K., Gopalakrishnan, V., et al. (2018). MLL4 Is Required to Maintain Broad H3K4me3 Peaks and Super-Enhancers at Tumor Suppressor Genes. Mol. Cell *70*, 825–841.e6. https://doi.org/10.1016/j.molcel.2018.04.028.

60. Malek, R., Gajula, R.P., Williams, R.D., Nghiem, B., Simons, B.W., Nugent, K., Wang, H., Taparra, K., Lemtiri-Chlieh, G., Yoon, A.R., et al. (2017). TWIST1-WDR5-Hottip Regulates Hoxa9 Chromatin to Facilitate Prostate Cancer Metastasis. Cancer Res. *77*, 3181–3193. https://doi.org/10.1158/0008-5472.CAN-16-2797.

61. Sze, C.C., and Shilatifard, A. (2016). MLL3/MLL4/COMPASS Family on Epigenetic Regulation of Enhancer Function and Cancer. Cold Spring Harbor Perspect. Med. *6*, a026427. https://doi.org/10.1101/cshperspect.a026427.

62. Salz, T., Deng, C., Pampo, C., Siemann, D., Qiu, Y., Brown, K., and Huang, S. (2015). Histone Methyltransferase hSETD1A Is a Novel Regulator of Metastasis in Breast Cancer. Mol. Cancer Res. *13*, 461–469. https://doi.org/10.1158/1541-7786.MCR-14-0389.

63. Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature *537*, 558–562. https://doi.org/10.1038/nature19362.

64. Kim, D.H., Tang, Z., Shimada, M., Fierz, B., Houck-Loomis, B., Bar-Dagen, M., Lee, S., Lee, S.K., Muir, T.W., Roeder, R.G., and Lee, J.W. (2013). Histone H3K27 trimethylation inhibits H3 binding and function of SET1-like H3K4 methyltransferase complexes. Mol. Cell Biol. *33*, 4936–4946. https://doi.org/10.1128/MCB.00601-13.

65. Karin, N., and Razon, H. (2018). Chemokines beyond chemo-attraction: CXCL10 and its significant role in cancer and autoimmunity. Cytokine *109*, 24–28. https://doi.org/10.1016/j.cyto.2018.02.012.

66. Tokunaga, R., Zhang, W., Naseem, M., Puccini, A., Berger, M.D., Soni, S., McSkane, M., Baba, H., and Lenz, H.J. (2018). CXCL9,

CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. Cancer Treat Rev. *63*, 40–47. https://doi.org/10.1016/j.ctrv.2017.11.007.

67. Wisdom, A.J., Hong, C.S., Lin, A.J., Xiang, Y., Cooper, D.E., Zhang, J., Xu, E.S., Kuo, H.C., Mowery, Y.M., Carpenter, D.J., et al. (2019). Neutrophils promote tumor resistance to radiation therapy. Proc. Natl. Acad. Sci. USA *116*, 18584–18589. https://doi.org/10.1073/pnas.1901562116.

68. Wang, C.Y.Y., Wei, Q., Han, I., Sato, S., Ghanbari-Azarnier, R., Whetstone, H., Poon, R., Hu, J., Zheng, F., Zhang, P., et al. (2012). Hedgehog and Notch signaling regulate self-renewal of undifferentiated pleomorphic sarcomas. Cancer Res. *72*, 1013–1022. https://doi.org/10.1158/0008-5472.CAN-11-2531.

69. Lu, K., Tao, H., Si, X., and Chen, Q. (2018). The Histone H3 Lysine 4 Presenter WDR5 as an Oncogenic Protein and Novel Epigenetic Target in Cancer. Front. Oncol. *8*, 502. https://doi.org/10.3389/fonc.2018.00502.

70. Cao, F., Townsend, E.C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillette, P., Zhu, J., et al. (2014). Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. Mol. Cell *53*, 247–261. https://doi.org/10.1016/j.molcel.2013.12.001.

71. Grebien, F., Vedadi, M., Getlik, M., Giambruno, R., Grover, A., Avellino, R., Skucha, A., Vittori, S., Kuznetsova, E., Smil, D., et al. (2015). Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia. Nat. Chem. Biol. *11*, 571–578. https://doi.org/10.1038/nchembio.1859.

72. Shimoda, H., Doi, S., Nakashima, A., Sasaki, K., Doi, T., and Masaki, T. (2019). Inhibition of the H3K4 methyltransferase MLL1/WDR5 complex attenuates renal senescence in ischemia reperfusion mice by reduction of p16$^{(INK4a)}$. Kidney Int. *96*, 1162–1175. https://doi.org/10.1016/j.kint.2019.06.021.

73. Tajima, K., Matsuda, S., Yae, T., Drapkin, B.J., Morris, R., Boukhali, M., Niederhoffer, K., Comaills, V., Dubash, T., Nieman, L., et al. (2019). SETD1A protects from senescence through regulation of the mitotic gene expression program. Nat. Commun. *10*, 2854. https://doi.org/10.1038/s41467-019-10786-w.

74. Jin, M.L., Kim, Y.W., Jin, H.L., Kang, H., Lee, E.K., Stallcup, M.R., and Jeong, K.W. (2018). Aberrant expression of SETD1A promotes survival and migration of estrogen receptor alpha-positive breast cancer cells. Int. J. Cancer *143*, 2871–2883. https://doi.org/10.1002/ijc.31853.

75. Cenik, B.K., and Shilatifard, A. (2021). COMPASS and SWI/SNF complexes in development and disease. Nat. Rev. Genet. *22*, 38–58. https://doi.org/10.1038/s41576-020-0278-0.

76. Wang, H., Fan, Z., Shliaha, P.V., Miele, M., Hendrickson, R.C., Jiang, X., and Helin, K. (2023). H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. Nature *615*, 339–348. https://doi.org/10.1038/s41586-023-05780-8.

77. Sze, C.C., Cao, K., Collings, C.K., Marshall, S.A., Rendleman, E.J., Ozark, P.A., Chen, F.X., Morgan, M.A., Wang, L., and Shilatifard, A. (2017). Histone H3K4 methylation-dependent and -independent functions of Set1A/COMPASS in embryonic stem cell self-renewal and differentiation.

Genes Dev. *31*, 1732–1737. https://doi.org/10.1101/gad.303768.117.

78. Greulich, F., Wierer, M., Mechtidou, A., Gonzalez-Garcia, O., and Uhlenhaut, N.H. (2021). The glucocorticoid receptor recruits the COMPASS complex to regulate inflammatory transcription at macrophage enhancers. Cell Rep. *34*, 108742. https://doi.org/10.1016/j.celrep.2021.108742.

79. Dangaj, D., Bruand, M., Grimm, A.J., Ronet, C., Barras, D., Duttagupta, P.A., Lanitis, E., Duraiswamy, J., Tanyi, J.L., Benencia, F., et al. (2019). Cooperation between Constitutive and Inducible Chemokines Enables T Cell Engraftment and Immune Attack in Solid Tumors. Cancer Cell *35*, 885–900.e10. https://doi.org/10.1016/j.ccell.2019.05.004.

80. Romero, J.M., Grünwald, B., Jang, G.H., Bavi, P.P., Jhaveri, A., Masoomian, M., Fischer, S.E., Zhang, A., Denroche, R.E., Lungu, I.M., et al. (2020). A Four-Chemokine Signature Is Associated with a T-cell-Inflamed Phenotype in Primary and Metastatic Pancreatic Cancer. Clin. Cancer Res. *26*, 1997–2010. https://doi.org/10.1158/1078-0432.CCR-19-2803.

81. Rossi, M., Altea-Manzano, P., Demicco, M., Doglioni, G., Bornes, L., Fukano, M., Vandekeere, A., Cuadros, A.M., Fernández-García, J., Riera-Domingo, C., et al. (2022). Author Correction: PHGDH heterogeneity potentiates cancer cell dissemination and metastasis. Nature *609*, E8. https://doi.org/10.1038/s41586-022-05226-7.

82. Rajan, S., Franz, E.M., McAloney, C.A., Vetter, T.A., Cam, M., Gross, A.C., Taslim, C., Wang, M., Cannon, M.V., Oles, A., and Roberts, R.D. (2023). Osteosarcoma tumors maintain intra-tumoral transcriptional heterogeneity during bone and lung colonization. BMC Biol. *21*, 98. https://doi.org/10.1186/s12915-023-01593-3.

83. Wang, D., Niu, X., Wang, Z., Song, C.L., Huang, Z., Chen, K.N., Duan, J., Bai, H., Xu, J., Zhao, J., et al. (2019). Multiregion Sequencing Reveals the Genetic Heterogeneity and Evolutionary History of Osteosarcoma and Matched Pulmonary Metastases. Cancer Res. *79*, 7–20. https://doi.org/10.1158/0008-5472.CAN-18-1086.

84. Wang, L.H., Aberin, M.A.E., Wu, S., and Wang, S.P. (2021). The MLL3/4 H3K4 methyltransferase complex in establishing an active enhancer landscape. Biochem.

Soc. Trans. *49*, 1041–1054. https://doi.org/10.1042/BST20191164.

85. Li, X., Lu, M., Yuan, M., Ye, J., Zhang, W., Xu, L., Wu, X., Hui, B., Yang, Y., Wei, B., et al. (2022). CXCL10-armed oncolytic adenovirus promotes tumor-infiltrating T-cell chemotaxis to enhance anti-PD-1 therapy. OncoImmunology *11*, 2118210. https://doi.org/10.1080/2162402X.2022.2118210.

86. Carew, J.S., Espitia, C.M., Zhao, W., Mita, M.M., Mita, A.C., and Nawrocki, S.T. (2017). Oncolytic reovirus inhibits angiogenesis through induction of CXCL10/IP-10 and abrogation of HIF activity in soft tissue sarcomas. Oncotarget *8*, 86769–86783. https://doi.org/10.18632/oncotarget.21423.

87. Reschke, R., Yu, J., Flood, B., Higgs, E.F., Hatogai, K., and Gajewski, T.F. (2021). Immune cell and tumor cell-derived CXCL10 is indicative of immunotherapy response in metastatic melanoma. J. Immunother. Cancer *9*, e003521. https://doi.org/10.1136/jitc-2021-003521.

88. Altorki, N.K., Markowitz, G.J., Gao, D., Port, J.L., Saxena, A., Stiles, B., McGraw, T., and Mittal, V. (2019). The lung microenvironment: an important regulator of tumour growth and metastasis. Nat. Rev. Cancer *19*, 9–31. https://doi.org/10.1038/s41568-018-0081-9.

89. Burns, J., Brown, J.M., Jones, K.B., and Huang, P.H. (2022). The Cancer Genome Atlas: Impact and Future Directions in Sarcoma. Surg. Oncol. Clin. N. Am. *31*, 559–568. https://doi.org/10.1016/j.soc.2022.03.013.

90. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

91. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

92. Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res. *44*, e45. https://doi.org/10.1093/nar/gkv1191.

93. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner.

Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

94. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. *15*, 554. https://doi.org/10.1186/s13059-014-0554-4.

95. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

96. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

97. Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat. Protoc. *12*, 828–863. https://doi.org/10.1038/nprot.2017.016.

98. Meier, J.A., Zhang, F., and Sanjana, N.E. (2017). GUIDES: sgRNA design for loss-of-function screens. Nat. Methods *14*, 831–832. https://doi.org/10.1038/nmeth.4423.

99. Banito, A., Li, X., Laporte, A.N., Roe, J.S., Sanchez-Vega, F., Huang, C.H., Dancsok, A.R., Hatzi, K., Chen, C.C., Tschaharganeh, D.F., et al. (2018). The SS18-SSX Oncoprotein Hijacks KDM2B-PRC1.1 to Drive Synovial Sarcoma. Cancer Cell *34*, 346–348. https://doi.org/10.1016/j.ccell.2018.07.006.

100. Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015). Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell *160*, 1246–1260. https://doi.org/10.1016/j.cell.2015.02.038.

101. Dai, M., Yan, G., Wang, N., Daliah, G., Edick, A.M., Poulet, S., Boudreault, J., Ali, S., Burgos, S.A., and Lebrun, J.J. (2021). In vivo genome-wide CRISPR screen reveals breast cancer vulnerabilities and synergistic mTOR/Hippo targeted combination therapy. Nat. Commun. *12*, 3055. https://doi.org/10.1038/s41467-021-23316-4.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Setd1a	Abcam	Cat# ab70378; RRID:AB_1951955
b-actin	Thermo Fisher Scientific	Cat# MA5-15739; RRID:AB_10979409
Cxcl10	R and D Systems	Cat# AF-466-NA; RRID:AB_2292487
Vinculin	Millipore	Cat# MAB3574; RRID:AB_2304338
H3K4Me3 C42D8 Rabbit mAb	Abcam,	Cat# ab8580; RRID:AB_306649
Bacterial and virus strains		
Ad5-CMV-Cre	Viral Vector Core Facility University of Iowa Health Care	N/A
pLenti-control	Duke core facility	N/A
pLenti-Setd1aKO	Duke core facility	N/A
pLenti-Cxcl10KO	Duke core facility	N/A
Chemicals, peptides, and recombinant proteins		
OICR-9429	Selleck	Cat#S7833
Deposited data		
RNA seq data	This study	NCBI GEO Database Accession number: GSE208363
ChIP seq data	This study	NCBI GEO Database Accession number: GSE208365
Whole Genome Sequencing data	This study	NCBI GEO Database Accession number: GSE208366
CRISPR modified clones	This study	NCBI GEO Database Accession number: GSE208367
Re analysis of single cell RNA sequencing data from a mouse tumor.	Wisdom et al.[67]	https://www.synapse.org (Synapse ID: syn18918968).
Re analysis of data for survival analysis in human tumors.	National Cancer Institute Genomics Data Commons.[89]	https://gdc.cancer.gov/
Experimental models: Organisms/strains		
LSL-KrasG12D; p53flox/flox; R26R-Confetti/ Confetti (KPCC)	This Study	N/A
Nude mice	Duke University Rodent Breeding Core	N/A
NOD scid gamma mice	Duke University Rodent Breeding Core	N/A
Recombinant DNA		
LentiCRISPR.v2	Addgene	Cat#52961
Lentiviral CRISPR Setd1a Mouse Gene Knockout Kit	OriGene	Cat#KN515623
Lentiviral CRISPR Cxcl10 Mouse Gene Knockout Kit	OriGene	Cat#KN504037
Mouse Setd1a lentiviral expression vector and control vector	ABM	Catalog #. 435200640295, LV 587
Mouse Cxcl10 lentiviral expression vector and control vector	ABM	Catalogue# 17166064, LV7

(Continued on next page)

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
ImageJ	ImageJ Software	https://imagej.net/ij/download.html
Bowtie2	Langmead et al.[90]	N/A
MACS2	Zhang et al.[91]	N/A
DESeq2	Lun et al.[92]	N/A
STAR	Dobin et al.[93]	N/A
MAGECK-RRA	Li et al.[94]	N/A

RESOURCE AVAILABILITY

Lead contact

Benjamin Alman (ben.alman@duke.edu).

Materials availability

This study did not generate new unique reagents and all materials in this study are commercially available. Plasmids and associated vector maps generated is this study are available upon request. Any additional analysis information for this work is available by request to the lead contact.

Data and code availability

- RNA sequencing, ChIP sequencing, Whole Genome sequencing and CRISPR modified clones sequencing data have been deposited in Gene Expression Omnibus Database Accession Display and available through the National Center for Biotechnology Information (NCBI). Accession codes are listed in the key resources table.
- There were no codes generated in this manuscript.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

We used a genetically engineered mouse model of sarcoma that mimics human Undifferentiated Pleomorphic Sarcoma (UPS)[49,50] that was adapted to define the MC population.[48] In this animal, an adenovirus expressing Cre recombinase (Adeno-Cre) is injected into the gastrocnemius muscle of mice with conditional gene mutations to activate a Kras^{G12D} oncogene and delete both alleles of the Trp53 tumor suppressor gene. p53 is the most commonly mutated tumor suppressor gene across many human cancers including human UPS.[43] Although no single oncogene is frequently mutated in UPS,[43] mutations of KRAS have been reported and activation of MAP Kinase signaling downstream of RAS frequently occurs in aggressive UPS.[44] A sarcoma develops at the site of injection within 3 months.[49,50] These sarcomas mimic human UPS histology and gene expression,[45] and by the predilection for lung metastasis in approximately 40% of mice.[46–48] We previously adapted this KP model with CRISPR technology by modifying the adenovirus to include Cas9 and a sgRNA to p53 to initiate KP sarcomas by creating insertions and deletions (indels) in p53 or by adding a cre-recombinase activated Confetti allele[49] to allow tracing of individual clones. Because each tumor-initiating cell has the identical founder mutations (Kras^{G12D} and loss of both copies of p53), this KP sarcoma model mimics human cancer heterogeneity where different cancer cells share the same founder mutations but display different biological behaviors. Indeed, within the KP sarcomas, different clones arising from distinct tumor-initiating cells have different metastatic potential as only the clone that we term MCs has a high propensity for developing metastasis and a unique gene expression profile.[48] To identify the subpopulation of MC in the mouse UPSs we used KP mice expressing the Confetti allele, as cells from these murine tumors maintain their phenotype in long term culture. By transplanting the primary tumors expressing multiple fluorescent markers in multiple mice and observing the fluorescent markers of metastases that develop, we found that any given primary tumor only developed metastases from cells expressing the same fluorescent reporter.[48] We generated primary cultures from independent clones[48] which were used in this study. The Kras mutation and P53 deletion from mice were observed in the clones including MC and non-MCs. These data for the clones utilized in this investigation were previously reported.[49]

Cells were transplanted into the hindlimb of congenic animals. The limb was amputated after two months to determine the primary tumor weight. Four weeks later, the animals were sacrificed and lung tumor volume determined using a technique as described in a subsequent paragraph. Equal numbers of male and female mice were used, and animals were on the black 6 background.

For human UPS tumors, NOD.SCID.gamma (NSG) and nude male mice 6-8 weeks old were utilized to allow effective tumor growth.[68] To test the COMPASS complex in human cells, NSG mice were treated with OICR-9429 (S7833, Selleck) at 10mg/kg dosage every day via

intraperitoneal injection for 8 weeks. For intramuscular injections, at 500,000 cells were injected and the hind limb was amputated after a .5cm diameter tumor formed. The animals were then observed afterwards for lung metastases as in our prior work.[48]

Male and female mice were used in equal numbers for all of the studies and mice were used at two months of age. Nude and NSG mice were purchased from Duke rodent breeding core. Animals were housed in a clean room providing high-quality feed and drinking water, with between two and five animals housed per cage.

All animal studies were approved by the Duke University and Duke University Medical Center Institutional Animal Care and Use Committee (A152-19-07) following guidelines set forth by the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

METHOD DETAILS

Protein analysis

To measure secretion of CXCL10 from UPS cells, CXCL10 Duoset ELISA kit (R&D, DY466) Enzyme-Linked Immunosorbent Assy (ELISA) was used and the optical density were measured using a microplate reader (spectramax) set to 450nm. For Western analysis the following antibodies were used: SETD1A (Abcam Cat# ab70378, RRID:AB_1951955), CXCL10 (R and D Systems Cat# AF-466-NA, RRID:AB_2292487), Vinculin (Millipore Cat# MAB3574, RRID:AB_2304338), Beta-actin: (Thermo Fisher Scientific Cat# MA5-15739, RRID:AB_10979409). Immuno-fluorescence was carried out using the Cxcl10 antibody at 1/100 dilution and the Setd1a, antibody at 1/100 dilution after blocking with Donkey Serum (D9663, SIGMA). Secondary antibodies were Donkey-Anti-Rabbit-647: A31573 and Donkey-Anti-Goat-594: A32758. NucBlue Stain (Thermo, P36981) was used to indentify nuclei. Slides were observed using a fluorescent microscope.

Lentiviral transduction and CRISPR modification

To generate stable knock-out cells, Lentivirus carrying CRISPR knock-out and CRISPR over-expression constructs, and empty vectors were transduced into the cells with 8 mg/mL of polybrene. After two days, the cells were selected with puromycin, and knock down was validated by Western Blot. To deplete *Setd1a* or *Cxcl10*, we used commercially available Gene Knockout Kits (OriGene KN515623 and KN504037 respectively), according to the manufacturer's instructions. Two gRNAs targeting *Setd1a* and two targeting *Cxcl10*, were used, both of which equally downregulated the gene of interest compared to controls. Clones from both gRNAs in for both genes showed the same *in-vivo* and we this combined the results from both clones. Controls were cells treated with a control gRNA in an identical manner.

Tumor analysis

Mice were euthanized. Primary tumors were dissected from surrounding tissues and weighed. Lungs were insufflated with 10% Neutral Buffered Formalin (NBF) and removed from the chest cavity. Fixation was allowed to continue for 24 hours, lungs were then washed thrice with PBS and placed in 70% Ethanol to dehydrate. An Automated Tissue Processor was used to further dehydrate the lungs in a gradient of alcohols and homogenize in Xylenes before being introduced to Paraplast. Lungs were then embedded into paraffin blocks for sectioning. Paraffin blocks were processed as 5 μm-thick sections. 30 slides were captured for each lung block, and stained for Hematoxylin & Eosin (H&E). Briefly, slides were deparaffinized and hydrated, followed by Mayer's Hematoxylin staining and Alcoholic Eosin counterstaining. Slides were mounted with Cytoseal, cover-slipped, and allowed to dry prior to imaging. Stained slides were imaged at 40x and stitched using the Leica Aperio GT 450 slide scanner. Images were opened in Qupath and exported as TIF files for analysis in Image J. Image analysis was completed as follows: In Image J, total area of lung tissue (including Tumor tissue within lungs, if applicable) was measured by individually measuring the area of each visible lung lobe and adding these areas together. Once total lung area was acquired, the same measurement was done for total tumor area. These numbers were used to calculate ratios of tumor burden as a function of each individual lung volume, to account for variation in lung size/insufflation between mice. Ten slices equal distanced across the entire lung fields were used to account for differences in tumor seeding locations, and data was combined to obtain the metastatic burden for each animal. Observers were blinded to the genotype or treatment.

ChIP sequencing

The ChIP-seq Kit from Cell Signaling was utilized. 2×10^7 MC Cells were fixed using 37% formaldehyde to crosslink proteins to DNA. Following two cold PBS washes, the cells were centrifuged at 1,000 x g for 5 minutes at 4°C with cold PBS containing a protease inhibitor cocktail. The resulting pellet was resuspended in ChIP Lysis Buffer and subjected to sonication for chromatin fragmentation. The proteins were then incubated with the Immunoprecipitating antibodies (H3K4Me3 C42D8, Cell Signaling, and IgG control) overnight at 4°C with rotation. Protein G Magnetic Beads were added to each immunoprecipitation reaction and incubated for 2 hours at 4°C with rotation. Chromatin was subsequently eluted using elution buffer, followed by reverse-crosslinking at 65°C overnight. DNA purification was carried out using the kit provided by Cell Signaling.

The ChIP-seq reads were aligned using Bowtie2[90] and the peak calling were performed using MACS2.[91] ChIP-seq reads were trimmed using TrimGalore (v 0.6.6) and then aligned to GRCm38/mm10 reference genome by bwa (v 0.7.17-r1188). Unmapped reads aligned to multiple locations were removed using samtools (v 1.11). Duplicated reads were marked by picard (v 2.26.0). The genomic coverage was saved in bigwig format and normalized by CPM method by bedTools (v 2.29.2) and University of California Santa Cruz (UCSC) Utilities bedGraphTo-BigWig (v 377). H3K4Me3 binding peaks were called by MACS2 (v 2.2.7.1) with parameters 'callpeak –keep-dup all –gsize 1.87e9 –format BAMPE' with IGG controls. The differential binding peaks were analyzed by Bioconductor package DiffBind (v 3.4.11) and csaw (v 1.28.0) .

Peaks with FDR < 0.05 and absolute log2 fold > 1 are considered as statistically significant and annotated by Bioconductor package ChIPpeakAnno (v 3.28.1. The coverage for gene body were performed by deepTools (v 3.1.3). The differential regions between conditions were determined by the CSAW[92] package. Regions with FDR < 0.05 and absolute log2 fold > 1 are considered as statistically significant.

RNA-sequencing

For RNA-seq, the reads were aligned using STAR,[93] and counted on Ensembl GTF via featureCounts.[95] The gene-level counts tables were then used for differential expression analysis via the DESeq2[96] package. Genes with FDR < 0.05 and absolute log2 fold > 1 are considered as statistically significant.

Whole genome variant analysis

The fastq files were aligned to the mouse reference GRCm38/mm10 sequence using bwa (v 0.7.15) mem with default parameters. The technique duplicates were merged, and library PCR duplicates were removed with picard (v 2.18.7). SNPs were called from the individual strain BAM files with Genome Analysis Toolkit (v4.0.5.1). The SNPs were filtered with expression "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0" and the Indels were filtered with expression "QD < 2.0 || FS > 200.0 || ReadPosRankSum < -20.0".

Survival analysis from cancer genome atlas data

Bulk RNA sequencing data in human UPS tumors (The Cancer Genome Atlas)[89] was used to determine if genes associated with the COMPASS complex predicted survival outcomes. Survival analysis was performed for each gene in R Studio (v 4.0.2). First, a median-split was performed to categorize UPS tumors (N=50) with "high" or "low" expression of *SETD1A*, and *CXCL10*. The observed distributions of gene expression of these genes for UPS, showed that the values are concentrated near the mean (tails are not long), thus the mean and median are almost identical for SETD1A (median = 10.061, mean = 10.057) and CXCL10 (median = 7.602, mean = 7.829), so the mean and median data is similar. Then, Kaplan-Meier survival curves (survfit, survival package [v3.2-11]) and hazard ratios (coxph, survival package) were generated for "high" and "low" expression groups. P<0.05 was considered statistically significant. We undertook a similar analysis for other sarcoma types and did not find such correlations.

CRISPR screen

A custom CRISPR sgRNA library targeting 454 mouse genes with 4 sgRNAs per gene was designed using GUIDES or CRISPick (https://portals.broadinstitute.org/gppx/crispick/public).[97,98] An additional 150 non-targeting sgRNAs were included to serve as controls, for a total of 1966 sgRNAs in the custom library.[99] Oligo pools were ordered from Twist Biosciences and PCR amplified and cloned into LentiCRISPR.v2 (Addgene #52961) by Gibson Assembly as described.[97] Assembled sgRNA vectors were transformed into Endura competent cells (Lucigen) with resulting library representation of 260-fold. Library DNA was isolated using ZymoPURE plasmid midiprep kits (Zymoresearch) and representation of the sgRNAs in the library was confirmed by high-throughput sequencing. Library lentivirus was prepared by cotransfecting HEK293T cells with lentivector, psPAX2 and pMD2.g and harvesting supernatant 48 hours post-transfection. Virus was filtered through 45 uM filters and stored at -80 until needed. The UPS tumor cells were infected with library lentivirus and selected with puromycin (0.4 M.O.I). The cells were injected to mice. After 2-3 weeks, the primary tumor and lung metastasis were collected for genomic DNA extraction (QIAGEN). The sgRNA sequences from each genomic DNA sample were PCR amplified using primers with Illumina adaptors and indexes for multiplexing. Samples were pooled and sequenced at the Duke Genome Sequencing Core Facility, to an average sequencing depth of 1900 reads per sgRNA. Screen analysis was performed using MAGECK-RRA[94] to identify genes depleted in the lung metastases compared to primary tumors.

Single cell RNA sequencing analysis

The data for analysis was downloaded from the Synapse database (Synapse ID: syn18918968). The unradiated and untreated mouse tumor sample (labeled as "iso_0") was used for further analysis. We performed quality control on the extracted data by removing the cells with low expression genes (minimum of 3 cells), nFeature_RNA<200 and percent.mt>5%. The data were normalized on a log scale. Principal components were calculated using the most variably expressed genes and the first 30 principal components were carried forward for clustering and visualization. Cells were embedded into a K-nearest neighbor graph using the FindNeighbors function and iteratively grouping cells together with the Louvain algorithm using the FindClusters function at resolution 0.5 Cluster biomarkers were identified using the FindAllMarkers function and visualized on the UMAP plot to show individual cell types (http://satijalab.org/seurat/pbmc3k_tutorial.html). Based on known feature genes and reference data, we annotated clusters 0, 2 & 6 as tumor cells. The identified tumor cell clusters were subset from the full dataset for further analysis. After subsetting, we used the Louvain algorithm for clustering with a resolution of 0.5. This was done using the FindClusters function from the Seurat package. We identified marker genes specific to each cluster using the FindAllMarkers function. This function performs differential expression analysis to determine genes that are significantly differentially expressed between clusters. Analysis for *Setd1a* and *Cxcl10* expression in individual cells undertaken.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Data was analyzed initially using the Bartlett test to determine if the comparison groups follow a normal distribution. The Student's t-test will be used for comparisons in which there are two groups; or an ANOVA test followed by a Turkey post hoc testing for three or more comparisons. When variances were not parametric, the Kruskal-Wallis and the Steel-Dwass tests was used. A power analysis was undertaken from prior data and 8 animals in each group were required to find a 25% difference in metastatic tumor burden.

Statistical cut offs in the CRISPR-Cas9 screening were chosen from those used in the literature. Because studies of metastasis result in widely spread distributions of sgRNAs leading to high variation, a higher FDR is generally selected. In the initial report of this approach, a false discovery rate (FDR)-based cutoff for enrichment (FDR < 0.2) was used as this resulted in the rigrous identification of significant genes that all had a p value of less than 0.05.[100] More recent studies applied an even higher FDR cutoff (FDR < 0.25) is studies of metastatic lesions.[101] Thus, we selected a more stringent FDR for the identification of genes involved in metastasis than other groups utilized.