

# Positively selected enhancer elements endow osteosarcoma cells with metastatic competence

James J Morrow<sup>1,2</sup>, Ian Bayles<sup>2</sup>, Alister P W Funnell<sup>3</sup>, Tyler E Miller<sup>1</sup>, Alina Saiakhova<sup>2</sup>, Michael M Lizardo<sup>4</sup>, Cynthia F Bartels<sup>2</sup>, Maaike Y Kapteijn<sup>5</sup>, Stephen Hung<sup>2</sup>, Arnulfo Mendoza<sup>4</sup>, Gursimran Dhillon<sup>2</sup>, Daniel R Chee<sup>6</sup>, Jay T Myers<sup>7</sup>, Frederick Allen<sup>1</sup>, Marco Gambarotti<sup>8</sup>, Alberto Righi<sup>8</sup>, Analisa DiFeo<sup>9</sup>, Brian P Rubin<sup>10</sup>, Alex Y Huang<sup>1,7</sup>, Paul S Meltzer<sup>11</sup>, Lee J Helman<sup>4</sup>, Piero Picci<sup>8</sup>, Henri Versteeg<sup>5</sup> , John Stamatoyannopolus<sup>3</sup>, Chand Khanna<sup>4,12</sup> & Peter C Scacheri<sup>2,8</sup>

**Metastasis results from a complex set of traits acquired by tumor cells, distinct from those necessary for tumorigenesis. Here, we investigate the contribution of enhancer elements to the metastatic phenotype of osteosarcoma. Through epigenomic profiling, we identify substantial differences in enhancer activity between primary and metastatic human tumors and between near isogenic pairs of highly lung metastatic and nonmetastatic osteosarcoma cell lines. We term these regions metastatic variant enhancer loci (Met-VELs). Met-VELs drive coordinated waves of gene expression during metastatic colonization of the lung. Met-VELs cluster nonrandomly in the genome, indicating that activity of these enhancers and expression of their associated gene targets are positively selected. As evidence of this causal association, osteosarcoma lung metastasis is inhibited by global interruptions of Met-VEL-associated gene expression via pharmacologic BET inhibition, by knockdown of AP-1 transcription factors that occupy Met-VELs, and by knockdown or functional inhibition of individual genes activated by Met-VELs, such as that encoding coagulation factor III/tissue factor (*F3*). We further show that genetic deletion of a single Met-VEL at the *F3* locus blocks metastatic cell outgrowth in the lung. These findings indicate that Met-VELs and the genes they regulate play a functional role in metastasis and may be suitable targets for antimetastatic therapies.**

More than 90% of all cancer-related deaths are the result of tumor metastases<sup>1</sup>. The physical process of tumor cell dissemination and metastatic colonization of distant secondary sites has been well described<sup>2</sup>. Whole-genome sequencing studies have elucidated the evolutionary phylogeny of metastatic dissemination<sup>3,4</sup>, and gene expression studies have revealed many of the genes that mediate the progressive steps of metastasis and drive organ-specific colonization<sup>5–7</sup>. These studies suggest that adaptation of metastatic tumor cells to the microenvironment of their destination organ is accompanied by a shift in cell state attained through widespread changes in the transcriptional output of metastatic cell genomes. Whether the shift is driven by genetic factors, epigenetic factors, or a combination of both of these mechanisms is not yet clear.

During normal development, gene expression changes that accompany cell state transitions are driven by altered activity of gene enhancer elements<sup>8–10</sup>. Enhancers govern cell-type-specific expression programs and are defined by signature chromatin features,

including the histone modifications H3K4me1 and H3K27ac and hypersensitivity to DNase<sup>11</sup>. Enhancers appear to be important in tumorigenesis as well. Previous studies have demonstrated that malignant transformation is accompanied by locus-specific gains and losses in enhancer activity across the epigenome; these are termed variant enhancer loci (VELs)<sup>12,13</sup>. Others have shown that in many types of cancers, clusters of active enhancers called super-enhancers (SEs) mediate dysregulated expression of oncogenes<sup>14,15</sup>. Collectively these studies suggest that aberrant enhancer activity is a key driver of tumor formation and maintenance.

Altered transcriptional programs play a role in metastatic tumor progression. These transcriptional programs have been associated with metastatic colonization of specific secondary organs<sup>5–7,16</sup>. Recently, epigenetic changes have been associated with transcriptional changes during metastasis<sup>17</sup>. However, the contribution of gene enhancers to metastatic transcription is not well understood. On the basis of the knowledge that enhancers drive cell state transitions during normal

<sup>1</sup>Department of Pathology, Case Western Reserve University, Cleveland, Ohio, USA. <sup>2</sup>Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, Ohio, USA. <sup>3</sup>Altius Institute for Biomedical Sciences, Seattle, Washington, USA. <sup>4</sup>Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA. <sup>5</sup>Thrombosis and Hemostasis Division, Department of Internal Medicine, Leiden University Medical Center, Leiden, the Netherlands. <sup>6</sup>Department of Genome Sciences, University of Washington, Seattle, Washington, USA. <sup>7</sup>Department of Pediatrics, Case Western Reserve University, Cleveland, Ohio, USA. <sup>8</sup>Research Laboratory, Istituto Ortopedico Rizzoli, Bologna, Italy. <sup>9</sup>Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, Ohio, USA. <sup>10</sup>Departments of Anatomic Pathology and Molecular Genetics, Cleveland Clinic, Lerner Research Institute and Taussig Cancer Center, Cleveland, Ohio, USA. <sup>11</sup>Genetics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA. <sup>12</sup>Present address: Ethos Discovery, Washington, DC, USA and Ethos Veterinary Health, Woburn, Massachusetts, USA. Correspondence should be addressed to P.S. (peter.scacheri@case.edu).

Received 30 June 2017; accepted 18 December 2017; published online 15 January 2018; doi:10.1038/nm.4475

development and tumorigenesis, we hypothesized that enhancers play a similar role in the transition of cancer cells from one developmentally distinct tissue to another during metastatic progression.

Osteosarcoma is the most common primary malignancy of bone, with peak incidence in children and adolescents. Clinical outcomes for patients have not improved over the past 30 years, and there are currently no approved targeted antimetastatic therapies for osteosarcoma in wide clinical use<sup>18</sup>. More than 75% of osteosarcoma metastases occur in the lung, and these are the cause of the overwhelming majority of osteosarcoma-related deaths<sup>19</sup>. In this study, we leverage the knowledge that gene enhancer activity is a cornerstone of cellular phenotypes and cell-type-specific gene expression<sup>9,20</sup> to gain new insight into the regulatory mechanisms that allow metastatic osteosarcoma cells to overcome the barriers to colonization encountered as these cells engage the lung microenvironment. Our study establishes that enhancer elements endow tumor cells with the capacity to metastasize and that targeted inhibition of genes associated with high enhancer activity in metastatic cells, or genetic deletion of metastatic-specific enhancers themselves, can block metastasis.

## RESULTS

### The metastatic phenotype of human osteosarcoma is associated with variant enhancer loci

We mapped the locations of putative enhancer elements across the genome through chromatin immunoprecipitation and sequencing (ChIP-seq) analysis of the canonical enhancer histone marks H3K4me1 and H3K27ac in matched primary tumors and lung metastases from five subjects with osteosarcoma. We also performed ChIP-seq for H3K4me1 and H3K27ac and DNase-seq in a panel of five well-characterized<sup>21</sup> metastatic and nonmetastatic (metastatic/nonmetastatic) human osteosarcoma cell line pairs representing three distinct mechanisms of metastatic derivation, including *in vivo* selection (MG63.3/MG63, LM7/SaOS-2, M112/Hu09), treatment with a mutagenic compound (MNNG/HOS), and introduction of an oncogenic driver (143B/HOS) (Fig. 1a). We used H3K4me1 for our initial comparisons based on the previous finding that this histone mark broadly correlates with both poised and active enhancers<sup>22,23</sup>.

We found thousands of regions where H3K4me1 signals showed at least a threefold difference in enrichment between conditions (Fig. 1b,c). The metastasis-associated gains and losses of the H3K4me1 signal were similar to those of the VELs that we had previously identified in the setting of primary tumor development through comparisons of primary colon tumors and healthy colon tissue<sup>12,13</sup>. We herein term the regions that show differential enrichment of H3K4me1 between metastatic samples and nonmetastatic controls Met-VELs for distinction. Enhancers, defined by differential enrichment of H3K4me1, generally showed concordant changes in H3K27ac ChIP-seq and DNase-seq signals (Fig. 1d and Supplementary Fig. 1), indicating robust commissioning and decommissioning of active enhancer elements at these loci. We found that, on average across all samples, 9.3% of all enhancers in a given metastatic cell line or tumor were gained relative to controls, whereas 16.4% of enhancers present in nonmetastatic cell lines or primary tumors were lost (Fig. 1e).

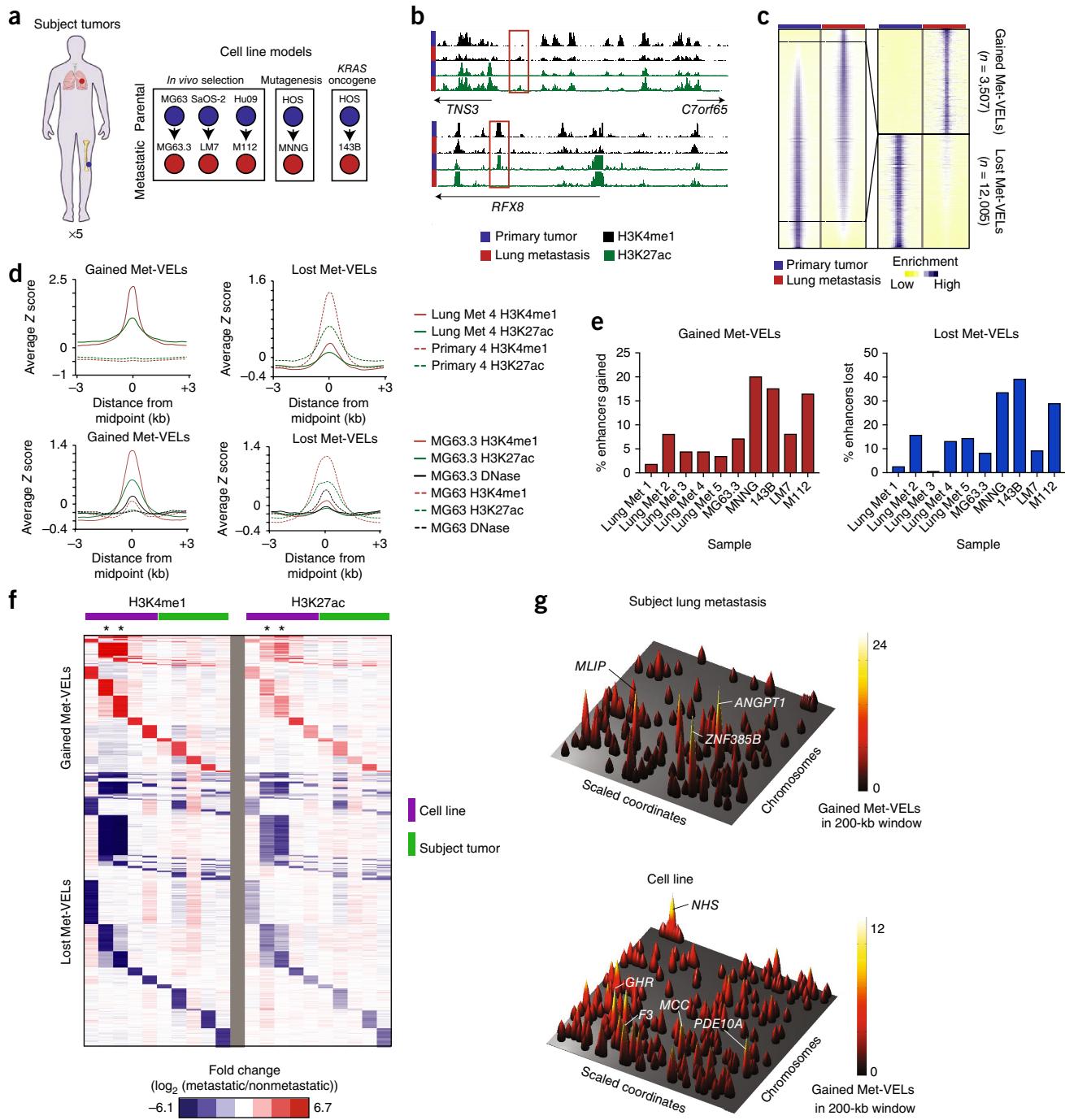
We next assessed the degree of Met-VEL heterogeneity across the cohort (Fig. 1f). Met-VELs were more concordant between two metastatic cell lines (MNNG and 143B) derived from a single parental cell line (HOS) than between those derived from distinct parental cell populations (40.8% versus 0.2–9.3%,  $P < 0.001$ ;  $\chi^2$  test with Monte Carlo correction), suggesting that the specific enhancer elements that undergo activation and silencing are nonrandom and may be

driven in part by the genetic and/or epigenetic makeup of the parental cell line. Met-VELs were heterogeneous among the remaining samples, with 23–69.1% overlapping with at least one other sample and 4.2–18.4% overlapping with two or more samples. No Met-VELs were shared among all cell lines or tissue samples. The heterogeneity observed could be a result of different selective pressures between the experimental approaches used to derive the metastatic cell lines and also heterogeneity in the selective process of metastasis in subjects. Additionally, the genetic heterogeneity among primary osteosarcoma tumors may contribute to the epigenetic heterogeneity we observed in cell lines and tumors.

An initial survey of Met-VEL distribution revealed dense clusters at distinct regions across the epigenome that were often in the vicinity of individual genes, similar to the pattern seen with SEs (Supplementary Fig. 2a). This finding led us to hypothesize that enhancer activity in these regions was nonrandomly acquired owing to selective pressures incurred during metastatic progression. We systematically tested this hypothesis and found numerous loci with Met-VEL counts significantly greater than expected by chance in all samples (clusters in exemplar pairs are displayed in Fig. 1g and Supplementary Fig. 2b). Several of the genes associated with Met-VEL clusters in both primary human samples and in cell lines have been previously implicated in tumor biology and/or progression. Angiopoietin 1 (ANGPT1) is a TIE2 receptor tyrosine kinase agonist that plays a crucial role in angiogenesis and is currently being studied as a therapeutic target in multiple tumor types<sup>24</sup>. Growth hormone receptor (GHR)<sup>25,26</sup>, phosphodiesterase 10A (PDE10A)<sup>27</sup>, and tissue factor (F3) have all been previously implicated in tumor biology and/or progression. F3 is a well-described activator of normal blood coagulation. F3 also plays an endogenous role within tumor cells to promote tumor growth and metastasis in multiple cancers, but the mechanism underlying its activation is not fully defined<sup>28</sup>. All metastatic/nonmetastatic cell lines and primary and metastatic tumor pairs studied showed evidence of nonrandom acquisition and loss of enhancer clusters (Supplementary Fig. 2c). On average across the cohort, 22% of all Met-VELs were found to reside in Met-VEL clusters (Supplementary Fig. 2d).

### Met-VELs dynamically modulate gene expression as tumor cells engage the lung microenvironment

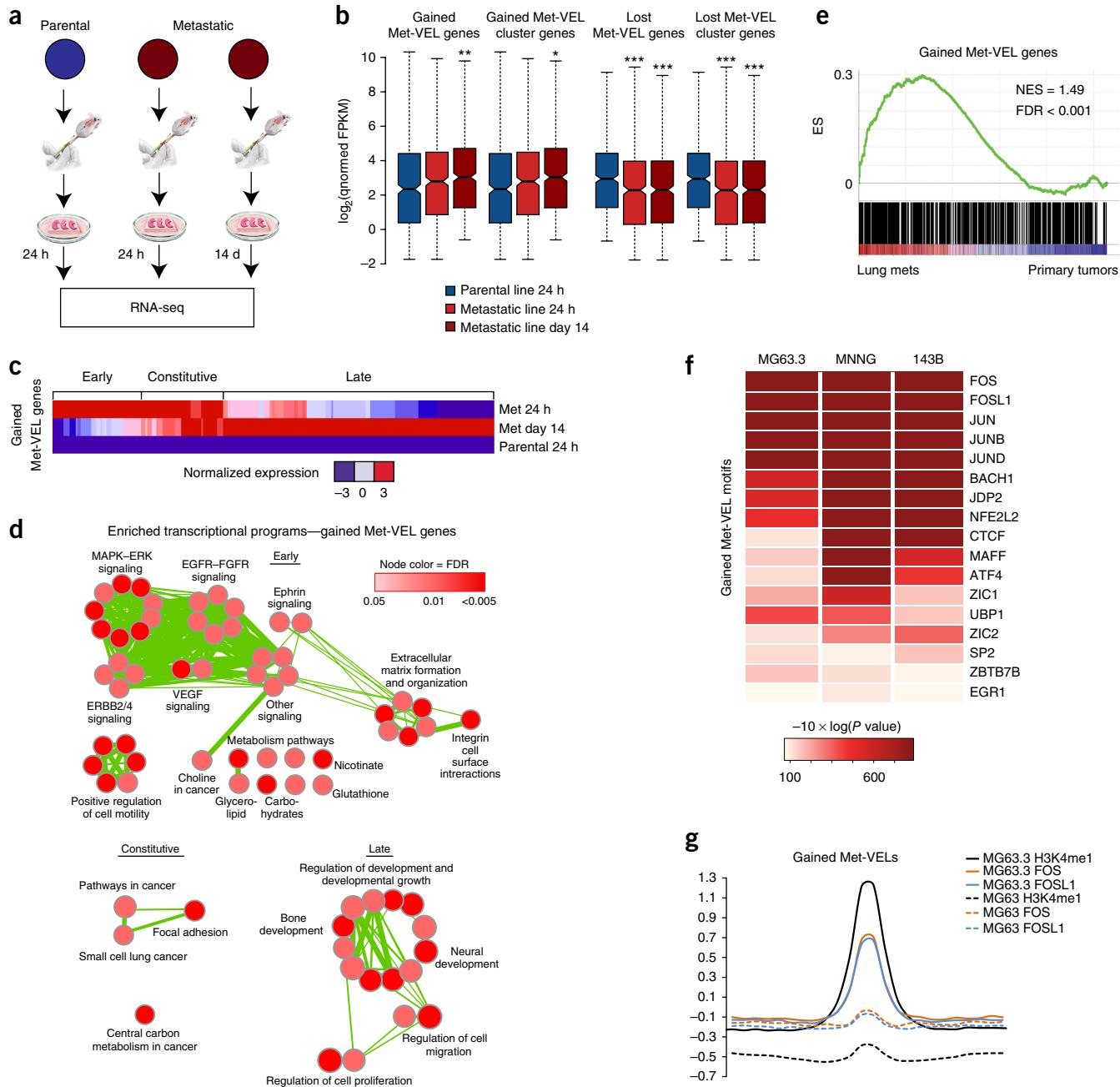
We used an *ex vivo* mouse model of osteosarcoma lung metastasis<sup>29</sup> to investigate the role of Met-VELs in modulating gene expression during metastasis. In this model, we seeded GFP-expressing tumor cells to the lungs of mice through intravenous injection. We then removed the lungs and grew lung sections in *ex vivo* culture. This approach allowed us to track metastatic outgrowth of GFP-labeled tumor cells in real time and to assess dynamic changes in gene expression (Fig. 2a). We performed RNA-seq at both early (24 h) and late (day 14) time points following injection in three metastatic/nonmetastatic cell line pairs. We determined the association between Met-VELs and their predicted target genes using PreSTIGE<sup>30</sup> and validated that >94.5% of predicted Met-VEL gene targets were located within the same topologically associating domains as their corresponding enhancer. In all cases, genes associated with gained Met-VELs and Met-VEL clusters were generally expressed at higher levels in metastatic cells within the lung microenvironment than in the corresponding nonmetastatic parental cell lines, whereas genes associated with lost Met-VELs and Met-VEL clusters were expressed at lower levels (Fig. 2b and Supplementary Fig. 3a,b). We compared the expression of gene sets associated with Met-VELs across conditions to investigate whether these represent a transcriptional program that is specifically modulated in metastasis. The degree



**Figure 1** Identification of Met-VELs and Met-VEL clusters through H3K4me1 ChIP-seq. **(a)** Schematic of the human tumor and the metastatic human osteosarcoma cell line cohorts. **(b)** UCSC browser views of H3K4me1 profiles from MG63.3 (metastatic) and MG63 (parental) cell lines illustrating an example of gained (top) and lost (bottom) Met-VEL(s). The Met-VELs are boxed in red. **(c)** Left, a heatmap showing the H3K4me1 ChIP-seq signal  $\pm 5$  kb from the H3K4me1 peak midpoints for all putative enhancers in the MG63.3/MG63 pair sorted by differences in signal. Right, a heatmap showing this signal for gained and lost Met-VELs alone. **(d)** Aggregate plots showing the H3K4me1 and H3K27ac ChIP-seq signals  $\pm 3$  kb from the midpoints of gained (left) and lost (right) Met-VELs for a representative matched primary/lung metastatic human tumor pair (top) and the MG63.3/MG63 cell line pair (bottom). The DNase-seq signal  $\pm 3$  kb from Met-VEL midpoints is also shown for the MG63.3 and MG63 cell lines. **(e)** Percentage of enhancers gained and lost in metastatic samples relative to primary tumors or nonmetastatic cell lines. **(f)** A heatmap showing the fold change of normalized reads per kilobase of transcript per million mapped reads (RPKM) in metastatic samples versus primary tumor or nonmetastatic cell lines for the aggregated list of all gained and lost Met-VELs across all samples. The H3K4me1 signal is shown at left and the H3K27ac signal at right. The samples in each heatmap from left to right are as follows: M112, 143B, MNNG, MG63.3, LM7, lung Met 1, Lung Met 2, Lung Met 3, Lung Met 4, and Lung Met 5. Asterisks mark 143B and MNNG samples. **(g)** The genome-wide gained Met-VEL landscape for human osteosarcoma metastatic tumor (Lung Met 4) and the MG63.3 cell line. Rows represent scaled chromosomal coordinates. Peaks represent maximum gained Met-VEL counts in 200-kb sliding windows. The predicted target genes for selected peaks are labeled.

of differential expression of Met-VEL-associated genes in the parental versus metastatic cells was greater in *ex vivo* lung culture than in standard *in vitro* culture conditions (Supplementary Fig. 4a,b), indicating that modulation of these transcriptional programs represents a

cellular response to external cues from the lung microenvironment. Met-VEL-associated gene sets showed little overlap (<27%) with the most differentially expressed genes in each metastatic/nonmetastatic cell line pair, indicating that the Met-VEL-associated genes represent



**Figure 2** Met-VELs modulate gene expression during metastatic colonization of the lung. (a) Schematic of the experimental design for assessment of Met-VEL-associated gene expression in parental and metastatic cell lines in the *ex vivo* lung metastasis model. Adapted by permission from ref. 29. (b) The log<sub>2</sub> quantile-normalized fragments per kilobase of transcript per million mapped reads (qnormed FPKM) values for gained (left) and lost (right) Met-VEL- and Met-VEL cluster-associated genes in the MG63/MG63.3 cell line pair. Asterisks indicate significant differences in FPKM distributions between parental and metastatic cell lines (\* $P < 0.05$ ; \*\*  $P < 10 \times 10^{-3}$ ; \*\*\*  $P < 10 \times 10^{-4}$ ). P values were calculated using a Mann–Whitney test. (c) A heatmap of upregulated genes associated with gained Met-VELs in the MG63/MG63.3 cell line pair illustrating a phasic expression pattern. (d) Enriched map representation of all Gene Ontology (GO) terms for three classes of gained Met-VEL genes determined through aggregating gene lists from all three cell line pairs. Green lines represent similarity (overlap) of related nodes. (e) Gene set enrichment analysis plot of the upregulated gained-Met-VEL-associated gene set compiled from three metastatic cell lines derived from human lung metastases (MG63.3, MNNG, 143B) versus that derived from primary tumors. ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate. (f) Expressed transcription factors with enriched binding-site motifs in gained Met-VELs in three metastatic/parental cell line pairs and corresponding P values for this motif enrichment. (g) Aggregate plot of H3K4me1, FOS, and FOSL1 ChIP-seq signals at all gained Met-VELs in the MG63.3 cell line.

a set distinct from those likely to be identified using expression data alone (**Supplementary Fig. 4c,d**).

Subsets of genes associated with gained Met-VELs became highly expressed within 24 h of the arrival of metastatic cells in lung; other gene subsets were only activated later during metastatic outgrowth, and a third subset was constitutively upregulated (**Fig. 2c** and **Supplementary Fig. 3c**). We assessed these gene sets for functional enrichment and found that the phasic waves of gene expression coordinated by Met-VELs during lung colonization were associated with distinct cellular functions (**Fig. 2d** and **Supplementary Fig. 3e**). We verified that genes associated with gained Met-VELs and upregulated in *ex vivo* lung culture frequently had elevated expression in lung metastases from individuals with osteosarcoma relative to primary tumors (**Fig. 2e**). We also found that genes associated with gained Met-VELs in patient metastases showed similar functional enrichment to the cell line gene sets (**Supplementary Fig. 5**). These results verify that the genes identified using the cell line models and the *ex vivo* approach are representative of those that are dysregulated in human tumors.

We next performed motif enrichment analysis and identified a number of commonly expressed transcription factors (TFs) with enriched binding-site motifs in gained and lost Met-VELs across all three pairs analyzed. The most highly enriched motifs include those for many members of the activator protein 1 (AP-1) complex (JUN, JUNB, JUND, FOS, and FOSL1; **Fig. 2f** and **Supplementary Fig. 3d**), which has previously been shown to play a key role in osteosarcoma metastasis<sup>31</sup>. Intriguingly, the binding-site motifs for AP-1 complex TFs were enriched in both gained and lost Met-VELs. This finding suggests that Met-VELs likely alter the transcriptional programs mediated by AP-1 during osteosarcoma metastasis. We verified through ChIP-seq that FOS and FOSL1 are bound at gained Met-VELs (**Fig. 2g**).

### Met-VEL-associated gene expression is required for metastatic colonization

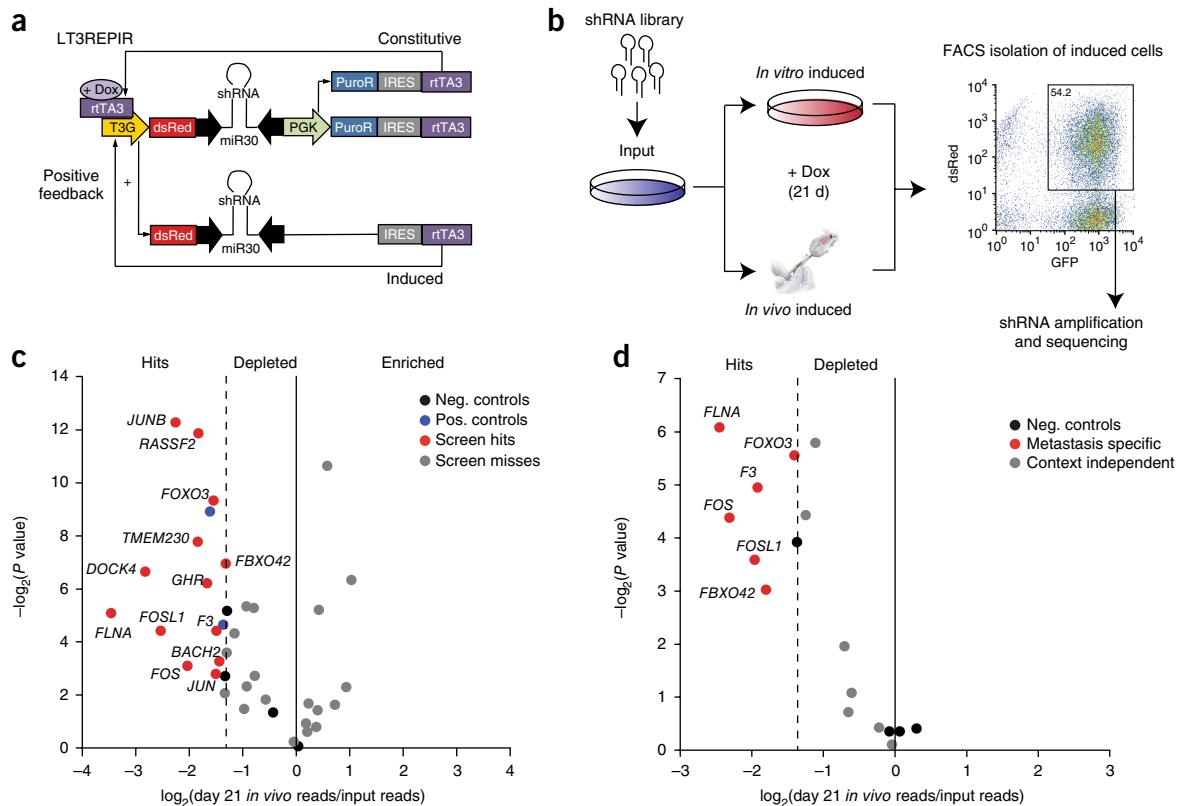
We reasoned that BRD4 inhibition may interrupt Met-VEL-associated gene expression and metastasis on the basis of the finding that gained Met-VELs have high levels of the activating histone mark H3K27ac and the knowledge that the bromodomain and extraterminal (BET)-family protein BRD4 is critical for transcriptional activation through H3K27ac-marked enhancers. To test this hypothesis, we used the BRD4 inhibitor JQ1, which displaces BRD4 from H3K27ac-marked enhancers<sup>15,32</sup>. JQ1 has been shown to inhibit primary tumor formation in osteosarcoma through its effects on both tumor cells and bone cells within the tumor microenvironment<sup>33</sup>. We found that JQ1 had potent antiproliferative effects on metastatic tumor cells growing in the lung microenvironment without affecting the surrounding healthy lung tissue (**Supplementary Fig. 6a–c**). These antiproliferative effects were associated with selective suppression of target genes of gained Met-VELs that are normally upregulated in metastatic cells in response to cues from the lung microenvironment (**Supplementary Fig. 6d,e**). Intriguingly, JQ1 more potently suppressed genes associated with gained Met-VELs than SE-associated genes (**Supplementary Fig. 6f**) that have previously been linked to JQ1's antitumor properties in other tumor models<sup>15,34,35</sup>.

The results of the JQ1 studies suggest that activation of gained Met-VEL-associated genes is necessary for metastatic outgrowth. To further test this hypothesis, we conducted a functional *in vivo* RNAi screen. We constructed a custom shRNA library targeting 33 genes. This gene list included 20 genes associated with gained Met-VELs or Met-VEL clusters, 11 TFs with binding-site motifs enriched in Met-VELs, and 2 genes of interest from other ongoing studies. We cloned the shRNA library

into a tetracycline-inducible lentiviral construct (LT3REPIR; **Fig. 3a**) modified from a similar construct previously published<sup>36</sup>. Using a *Discosoma* sp. red fluorescent protein (dsRed) fluorescent reporter of shRNA induction, we show that this construct was robustly induced upon exposure to doxycycline and not leaky in the absence of doxycycline (**Supplementary Fig. 7**). We conducted parallel screens *in vivo* and *in vitro* to allow us to distinguish genes that specifically inhibit metastatic outgrowth in the *in vivo* microenvironment, termed metastasis dependency genes, from genes whose inhibition reduces cellular growth independent of context (**Fig. 3b**). In the *in vivo* screen, transduced cells were delivered via injection in tail veins of mice pretreated with doxycycline. Mice were maintained on doxycycline throughout the 21-d course of the experiment. In the parallel *in vitro* screen, transduced cells were treated in culture with doxycycline for 21 d. At the conclusion of the experiment, induced cells actively expressing shRNA (GFP<sup>+</sup>dsRed<sup>+</sup>) were sorted from mouse lung or *in vitro* culture through fluorescence-activated cell sorting (FACS). DNA was isolated from these cells along with noninduced cells from the initial population (input), and the shRNAs were amplified, sequenced, and aligned to the reference shRNA sequences in the library to determine normalized representation for each shRNA.

We defined metastasis dependency genes as those whose knockdown inhibited *in vivo* metastasis at least 2.5-fold more than *in vitro* growth. First, shRNA representations in metastatic tumor cells were compared to input representations to identify shRNAs depleted from the population of cells during metastatic outgrowth in lung (**Fig. 3c**). Genes were defined as initial hits when they were a target of at least two unique shRNAs (the pool contained three or four shRNAs per gene) that inhibited metastatic outgrowth to a greater degree than all negative controls. Because we used a second filter for these hits in this screen and planned further functional validation experiments, we intentionally chose a relatively inclusive threshold for initial-hit calling. We found that 13 of 33 genes (39%) included in our screen met this criterion. To determine whether the depletion of cells expressing these shRNAs was specific to metastasis, we compared the relative representation of hits in the population of cells induced *in vivo* to the controls induced *in vitro* (**Fig. 3d**). Genes targeted by shRNAs that were significantly more depleted *in vivo* than *in vitro* were considered metastasis dependency genes; 6 of the 13 initial hits met this criterion (**Supplementary Table 1**). Metastasis dependency genes included four genes (*F3*, *FBXO42*, *FLNA*, and *FOXO3*) associated with gained Met-VEL clusters as well as two genes (*FOS* and *FOSL1*) encoding AP-1 complex TFs that had binding-site motifs enriched in Met-VELs and that were shown to bind at these enhancers. These results indicate metastatic colonization of the lung by osteosarcoma cells is dependent on expression of a subset of individual genes associated with gained Met-VEL clusters as well as AP-1 complex TFs that are likely to regulate Met-VEL transcriptional programs.

Among the Met-VEL-associated genes, *F3* emerged as a top candidate driver of metastasis in osteosarcoma. In the MG63.3 cell line, *F3* was associated with a gained Met-VEL cluster containing the second-highest gained Met-VEL count among the entire data set (**Fig. 1g**), suggesting that this locus was under particularly strong positive selection during metastatic derivation. Gained Met-VELs in the *F3* cluster also showed higher levels of H3K27ac and increased DNase sensitivity in MG63.3 cells as compared to the parental MG63 cell line, and chromatin conformation capture studies confirmed that these enhancers physically contact the transcription start site of *F3* (**Supplementary Fig. 8a**). *F3* was more highly expressed in MG63.3 cells during metastatic outgrowth than in the parental MG63 cells (**Supplementary Fig. 8b**). In addition, two other metastatic cell lines, MNNG and 143B,



**Figure 3** Results from an *in vivo* high-throughput RNAi functional assay of candidate metastasis dependency genes. **(a)** Schematic of the doxycycline (dox)-inducible LT3REPIR shRNA construct. Modified from ref. 36. T3G, tetracycline-responsive element; rTA3, reverse tetracycline transactivator; IRES, internal ribosomal entry site; PGK, phosphoglycerate kinase promoter. **(b)** Schematic of the experimental design for the *in vivo* high-throughput functional assay of candidate metastasis dependency genes. **(c)** Volcano plot showing the relative abundance of shRNAs targeting 33 genes in isolated GFP<sup>+</sup>dsRed<sup>+</sup> osteosarcoma cells from doxycycline-treated mice ( $n = 5$  mice per replicate, with 3 replicates) versus that of the input cell population. The second-most depleted shRNAs for each gene are plotted as well as negative and positive shRNA controls. Negative controls contained groups of 2–4 shRNAs. **(d)** Volcano plot showing the relative abundance of shRNAs targeting 13 genes meeting initial-hit criteria (Fig. 3c) in isolated GFP<sup>+</sup>dsRed<sup>+</sup> osteosarcoma cells from doxycycline-treated mice ( $n = 5$  mice per replicate, with 3 replicates) versus isolated GFP<sup>+</sup>dsRed<sup>+</sup> osteosarcoma cells treated with doxycycline *in vitro*. The second-most-depleted shRNAs for each gene are plotted as well as negative controls, which contained groups of 2–4 shRNAs. In c and d, each dot marks a shRNA, and labels indicate the gene that the shRNA targeted.

showed active enhancer signals at the *F3* locus, similar to MG63.3 (Supplementary Fig. 8d), and expressed *F3* at higher levels during metastatic outgrowth than their respective parental cell line (Supplementary Fig. 8b). To verify that elevated *F3* transcript levels were recapitulated at the protein level and also were not an artifact of *ex vivo* culture, we performed immunofluorescence analysis of lung metastases from a fully *in vivo* model of metastasis and confirmed that metastatic osteosarcoma cells expressed higher levels of *F3* protein than nonmetastatic cells (Fig. 4a,b). Quantification of *F3* expression levels directly in human osteosarcoma samples showed that *F3* expression was elevated in lung metastases relative to primary tumors (Supplementary Fig. 8c). Using a tissue microarray, we confirmed that *F3* protein was highly expressed in lung metastases from human osteosarcoma patients. *F3* was expressed in >50% of tumor cells in 18 of 18 lung metastases, and *F3* had strongly positive staining in 17 of 18 samples (Fig. 4c,d and Supplementary Fig. 9).

We analyzed the enhancer epigenomes of ten pairs of matched primary and metastatic tumors to determine whether *F3* upregulation in osteosarcoma lung metastases was a result of enhancer dysregulation. Although we did not observe *de novo* creation of a sufficient number of gained Met-VELs at the *F3* locus for it to meet the criteria to be called a gained cluster, this locus has an enhancer cluster that in 9

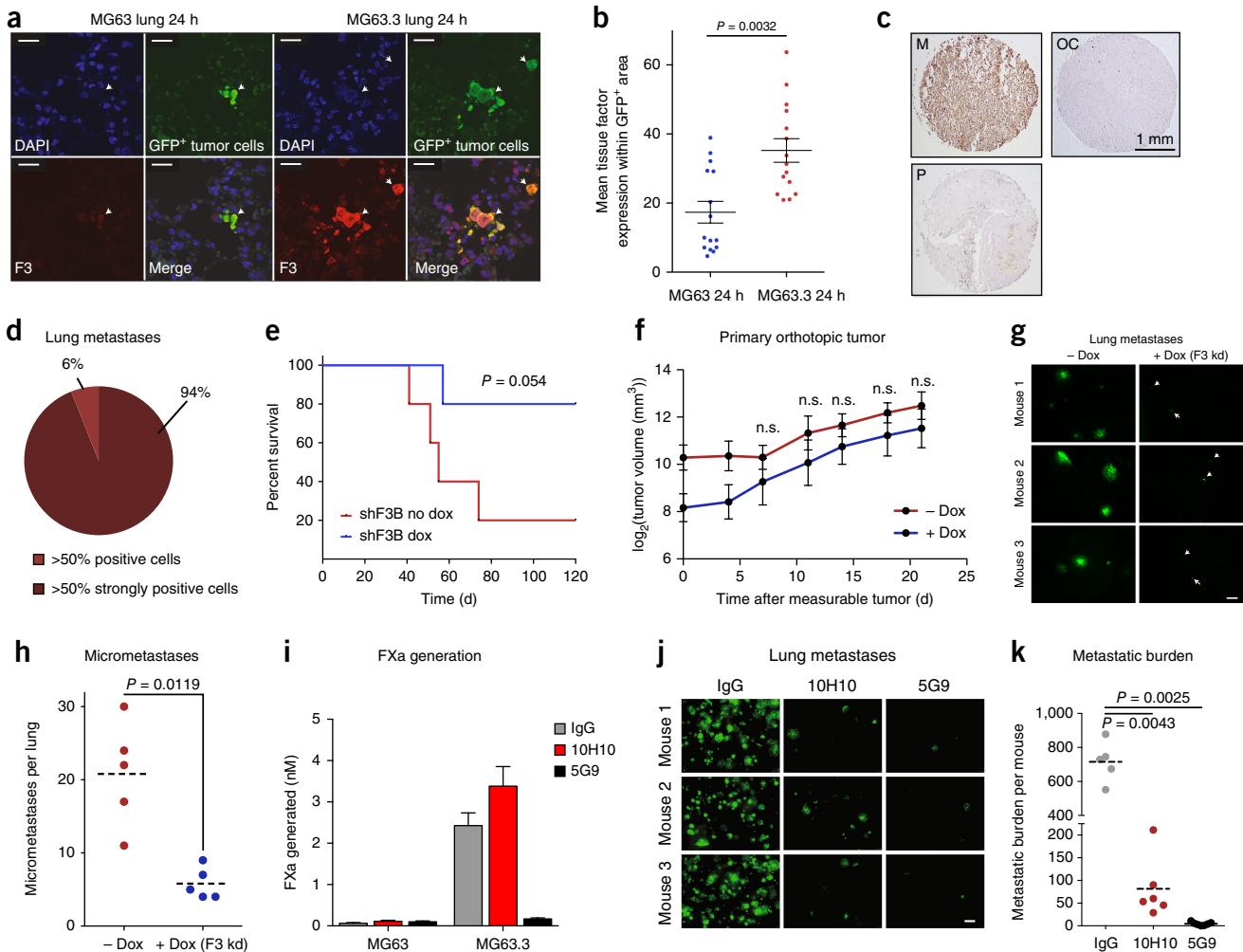
out of 10 metastatic samples was called as a SE and ranked among the sites most enriched for H3K27ac (Supplementary Fig. 10a). Comparison of the SE landscapes across all samples showed that the *F3* SE was among the top 1.3% of SEs enriched in lung metastases relative to primary tumors ( $P < 0.05$ , *t*-test). This was true when metastatic tumors were compared to primary tumors alone and when metastatic/nonmetastatic cell lines were included in the analysis (Supplementary Fig. 10b).

We further investigated the enhancer profiles of these ten paired samples and verified that common target genes of Met-VELs in cell lines often overlapped with common Met-VEL targets in samples of primary tumors (Supplementary Table 2).

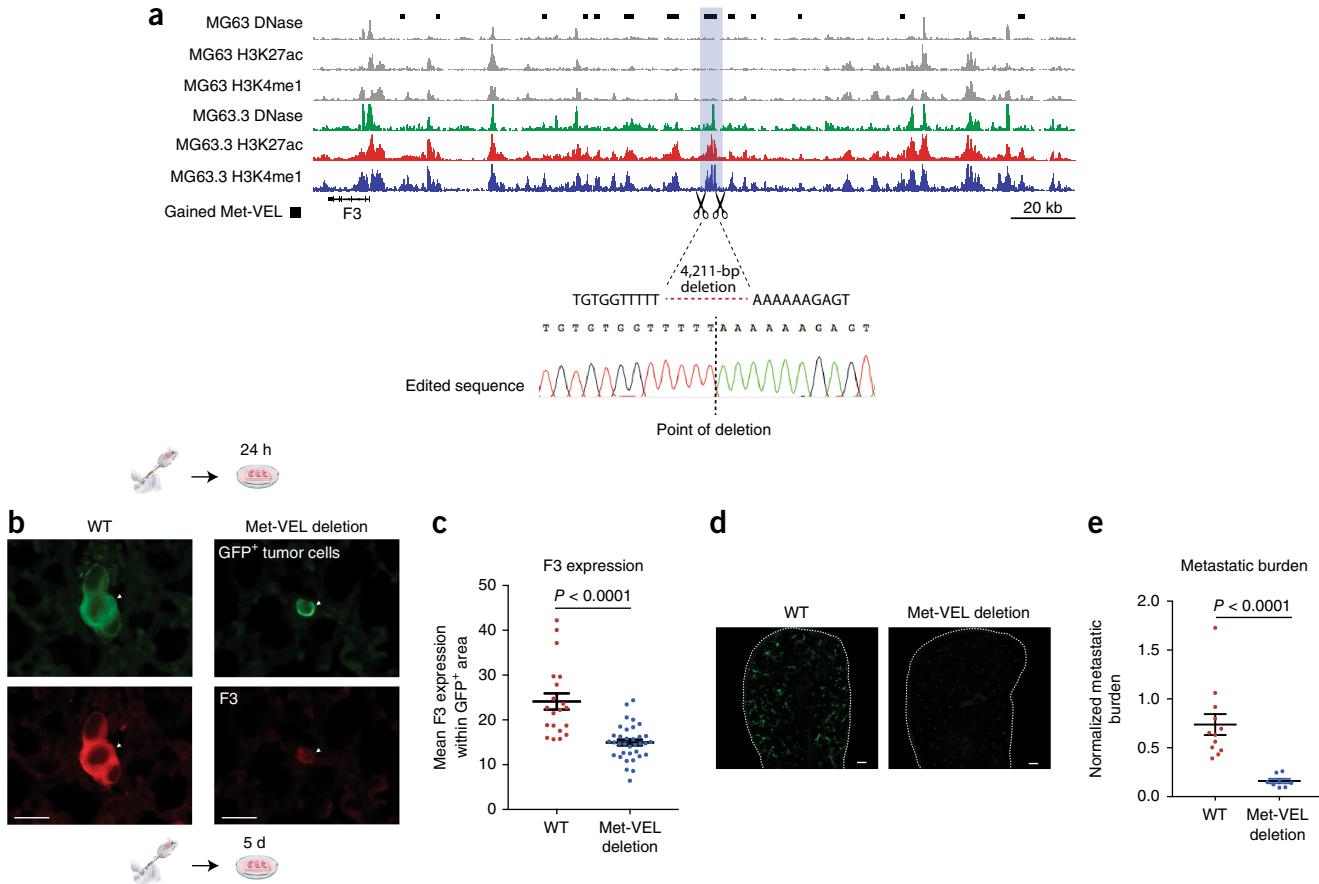
We next tested the functional contribution of *F3* expression to the metastatic phenotype. We cloned two shRNAs targeting *F3*, which were not included in the RNAi assay, into the tetracycline-inducible LT3REPIR construct. Relative to that in noninduced control cells, *F3* expression was reduced by 44–63% 40 h after induction of each shRNA (Supplementary Fig. 11). *F3* knockdown with these shRNAs did not affect the *in vitro* growth rate of metastatic MG63.3 or MNNG cells (Supplementary Fig. 12a) but significantly reduced metastatic outgrowth of these cells in *ex vivo* lung culture (Supplementary Fig. 12b,c), supporting the metastasis-specific role for *F3* in this setting.

F3 knockdown also significantly reduced metastatic outgrowth of osteosarcoma cells *in vivo* (Supplementary Fig. 12d,e) and substantially prolonged survival of mice injected with metastatic osteosarcoma cells (Fig. 4e). To further test whether F3 knockdown reduced *in vivo* growth of metastatic cells generally or specifically in the context of metastatic

outgrowth of cells in lung, we performed a spontaneous metastasis experiment using an orthotopic injection model. We found that F3 knockdown did not reduce primary tumor development or growth (Fig. 4f) but significantly inhibited metastasis, reducing the average metastatic burden by 3.6-fold (Fig. 4g,h). Although extensive GFP<sup>+</sup> metastatic



**Figure 4** F3 mediates lung metastasis of osteosarcoma. (a) Representative *in vivo* images (2.5× magnification) of GFP<sup>+</sup> MG63 (parental) and MG63.3 (metastatic) cells in mouse lung 24 h following injection of  $1 \times 10^6$  cells in tail vein. Sections were stained for tumor (GFP, green), F3 (red), and DAPI (blue). Arrowheads indicate individual tumor cells within lung. Scale bars, 20  $\mu$ m. (b) Quantification of mean red pixel intensity in the images represented in a within the GFP<sup>+</sup> (tumor) area in MG63 (parental) and MG63.3 (metastatic) cells 24 h after injection into tail vein ( $n = 15$  images per condition). Data are presented as means  $\pm$  s.e.m. (c) Representative images of immunohistochemical staining for F3 in human osteosarcoma lung metastases (M), primary tumor (P), and omission control (OC). The tissue microarray contained 18 scoreable lung metastases of similar quality to those displayed. (d) Percentage of lung metastases with various levels of F3 positivity. (e) Kaplan–Meier plot of untreated (red) and doxycycline-treated (blue) mice injected with  $5 \times 10^4$  GFP<sup>+</sup> MG63.3 cells transduced with the shF3B construct (assessed in Supplementary Fig. 11;  $n = 5$  mice per condition) via tail vein. The  $P$  value was calculated using the Gehan–Breslow–Wilcoxon test. (f) Primary tumor growth in untreated (red) and doxycycline-treated (blue) mice that received an orthotopic injection of  $8 \times 10^5$  GFP<sup>+</sup> MG63.3 cells transduced with the shF3B construct. Data are presented as means  $\pm$  s.e.m. ( $n = 5$  mice per condition).  $P$  values were calculated using a Student's *t*-test; n.s., nonsignificant. (g) Representative images (2.5× magnification) of *in vivo* metastatic lesions in lung 21 d after measurable tumor formation in untreated (left) and doxycycline-treated (right) mice that received orthotopic injection of  $8 \times 10^5$  GFP<sup>+</sup> MG63.3 cells transduced with the shF3B construct. kd, knockdown. Arrowheads indicate individual tumor cells within lung. Scale bar, 500  $\mu$ m. (h) Quantification of lung metastatic burden 21 d after measurable tumor formation in untreated (red) and doxycycline-treated (blue) mice that received orthotopic injection of  $8 \times 10^5$  GFP<sup>+</sup> MG63.3 cells transduced with shF3B construct ( $n = 5$  mice per condition, 5 images per mouse).  $P$  values were calculated using the Mann–Whitney test. Each point represents one mouse, and midlines represent means. (i) The amount of FXa formed in the *in vitro* assay by MG63 (left) and MG63.3 (right) cells treated with 25  $\mu$ g/ml IgG control, 10H10, or 5G9 antibodies for 20 min before the addition of FVIIa and FX to a final concentration of 100 nM. FXa formation was assessed 30 min after adding FX.  $P$  values were calculated using Student's *t*-test with Welch's correction. (j) Representative images (2.5× magnification) of *in vivo* metastatic lesions in lungs of mice 14 d after injection of  $5 \times 10^5$  MG63.3 cells with 500  $\mu$ g of IgG, 5G9, or 10H10 antibodies via the tail vein. Scale bar, 500  $\mu$ m. (k) Quantification of lung metastatic burden 14 d after injection of  $5 \times 10^5$  MG63.3 cells with 500  $\mu$ g of IgG, 5G9, or 10H10 antibodies via tail vein ( $n = 5$  mice per condition, 5 images per mouse).  $P$  values were calculated using the Mann–Whitney test. Each point represents one mouse, and midlines represent means.



**Figure 5** Deletion of a single gained Met-VEL blunts F3 expression and mitigates lung metastasis of osteosarcoma cells. **(a)** Integrative Genomics Viewer (IGV) browser view of the region targeted for deletion with TALENs. The schematic shows the strategy for 4,211-bp deletion. Sanger sequencing shows the resulting clonal homozygous deletion, with gray indicating the nonmetastatic line and other colors indicating the various marks. **(b)** Representative images (40x magnification) of WT and Met-VEL-deleted MG63.3 cells in lung 24 h after initiation of the *ex vivo* lung metastasis experiment. Scale bars, 50  $\mu$ m. **(c)** Quantification of mean red pixel intensity (F3 expression) within GFP<sup>+</sup> (tumor) area for WT and Met-VEL-deleted MG63.3 cells 24 h after initiation of the *ex vivo* lung metastasis assay. Each point represents one mouse, and data are presented as mean  $\pm$  s.e.m. ( $n =$  at least 20 images per condition, sections taken from  $n = 3$ –5 mice per condition). The  $P$  value was calculated using a Mann–Whitney test. **(d)** Representative images (2.5x magnification) of WT (left) and Met-VEL-deleted (right) lung sections at day 5 of *ex vivo* lung culture. Lung sections are outlined with a dashed white line. Scale bar, 500  $\mu$ m. **(e)** Quantification of metastatic burden for GFP<sup>+</sup> WT and Met-VEL-deleted MG63.3 cells at day 5 of *ex vivo* lung culture. Each point represents one mouse, and data are presented as mean  $\pm$  s.e.m. from at least 8 sections per condition ( $n = 4$  sections per mouse for  $n = 2$  or 3 mice) normalized to the same section at day 0. The  $P$  value was calculated using the Mann–Whitney test.

lesions were observed in control mice, lungs of mice in the F3-knockdown group were virtually devoid of metastatic lesions, with only rare single GFP<sup>+</sup> cells observable in most cases (Supplementary Fig. 13).

F3 is known to both induce blood coagulation through mediating the generation of the active form of factor X (FXa) and promote cell survival and proliferation upon binding to activated factor VII via intracellular signaling mechanisms<sup>28</sup>. To determine the relative contribution of each of these functions to the prometastatic role of F3, we used monoclonal antibodies generated to inhibit each function independently<sup>37</sup>. As expected, we found that MG63.3 cells produced more FXa in an *in vitro* assay than MG63 cells did (Fig. 4i). We confirmed that anticoagulant monoclonal antibody 5G9 robustly inhibited this activity, whereas monoclonal antibody 10H10, designed to prevent intracellular signaling, did not. We next tested the antimetastatic effects of these antibodies *in vivo*. We co-injected MG63.3 cells with each antibody or with IgG control into tail veins of mice and found that both inhibited metastasis, with monoclonal antibody 5G9 showing a more pronounced effect (Fig. 4j,k). These results indicate that both the intracellular signaling and procoagulant functions of F3

contribute to metastatic progression, but the procoagulant activity of F3 is especially critical for metastatic success. Collectively, these findings suggest that Met-VELs regulate expression of genes with critical functions during metastatic progression, such as F3.

To directly test the role of Met-VELs in mediating the metastatic phenotype, we employed transcription activator-like effector nuclease (TALEN) genome editing to excise one of the gained Met-VELs predicted to regulate F3 in metastatic cells. We targeted a Met-VEL located in a particularly robust DNase hypersensitivity site containing high levels of both H3K4me1 and H3K27ac (Fig. 5a). This site also showed high ChIP-seq enrichment for both FOS and FOSL1 AP-1 complex members. We generated a cell clone with homozygous deletion of this Met-VEL, which was verified using Sanger sequencing (Fig. 5a). Edited cells and unedited control cells were then seeded to mouse lungs via intravenous tail vein injection, and growth of the cells was monitored in the lungs using the *ex vivo* metastasis assay. Quantification of F3 levels at 24 h following injection showed that F3 expression in the edited cells was reduced by 34% relative to unedited control cells (Fig. 5b,c). By day 5, lungs seeded with the F3 Met-VEL-deleted cells were nearly devoid of

tumor cells, whereas extensive GFP<sup>+</sup> metastatic lesions were observed in lungs seeded with the unedited cells (**Fig. 5d**). Quantification showed that deletion of this gained Met-VEL decreased metastatic burden by 78% relative to unedited cells (**Fig. 5e**).

## DISCUSSION

Although many of the genes responsible for metastatic progression have been identified across tumor types, the underlying mechanisms regulating expression of these genes are not well defined. Our studies demonstrate that altered enhancer activity is a fundamental mechanism through which tumor cells regulate gene expression during the dynamic process of metastasis and thereby acquire metastatic traits. Through epigenomic profiling experiments, we identified enhancers that distinguished human osteosarcoma lung metastases from matched primary tumors and verified that these differences were also present in near-isogenic metastatic/nonmetastatic paired human osteosarcoma cell lines. Subsets of these enhancer changes occurred in nonrandom clusters, indicating that they were positively selected during the process of metastatic derivation. These results demonstrate that the metastatic phenotype is accompanied by a shift in the enhancer epigenome, similar to the enhancer shifts that occur as cells transition through successive stages of embryonic development<sup>8–10</sup> or during conversion of a healthy cell to a malignant state<sup>12,14,15</sup>. Our findings suggest that the evolutionary selective forces encountered by tumor cells during metastasis act to shape the enhancer landscape of metastatically successful cancer cell populations. The result of this selection is a population of cells possessing all of the traits necessary to overcome the barriers to metastatic colonization in distant tissues. Indeed, we show that many genes previously associated with metastasis become dysregulated through alterations in enhancer activity.

We provide multiple lines of evidence that acquired enhancer changes in metastatic osteosarcoma cells are functional and relevant to the osteosarcoma metastatic phenotype in experimental models and human tissues. First, we show that Met-VEL-associated genes are dynamically regulated as metastatic cells engage the lung microenvironment and proliferate. Second, we demonstrate that metastatic cell outgrowth in the lung can be mitigated through BET inhibition, and that this effect is associated with selective suppression of genes that are normally activated by Met-VELs in the lung. Third, through *in vivo* functional RNAi-based assays, we demonstrate that the metastatic capacity of the osteosarcoma cells can be diminished through targeted inhibition of individual Met-VEL-associated genes and associated AP-1 family TFs that likely regulate Met-VEL transcriptional programs. Using a fully *in vivo* model of spontaneous metastasis, we further verify that one such Met-VEL-associated gene, *F3*, is a clinically relevant, bona fide metastasis dependency gene that is essential for metastatic colonization but does not have an apparent role in primary tumor growth. Interrupting the signaling and procoagulant functions of *F3* was sufficient to inhibit metastasis, shedding light on the biological role of this gene in the metastatic progression of osteosarcoma. Our genomic Met-VEL-deletion experiments demonstrate that the loss of function of a single gained enhancer at the *F3* locus was sufficient to impair metastatic colonization and subsequent outgrowth in mice, indicating that enhancer activation contributed to acquisition of the metastatic phenotype in these cells.

Our current model for lung colonization by metastatic osteosarcoma cells involves a requirement for upregulation of *F3* expression via the aberrant activation of its enhancers. It is not yet clear whether neutralizing *F3* after colonization would lead to regression of overt metastases. The current standard of care for osteosarcoma involves multiple cycles

of neoadjuvant combination chemotherapy before surgical removal of the primary tumor<sup>18</sup>. As such, there is a period of several months between the time of diagnosis and primary tumor control. This treatment period is critical, as there is continued potential for cells to leave the primary tumor and seed the lung. We believe that the most clearly indicated therapeutic setting for *F3* inhibition would be during this period of neoadjuvant therapy. The findings described here may present a rationale for testing *F3* inhibition in such a setting. Further studies are required to determine whether there may be clinical benefit from continued *F3* inhibition after primary tumor control is achieved.

Collectively, our findings indicate that altered enhancer activity is a driver of gene expression that is critical for tumor cells, allowing them to overcome the barriers of distant tissue colonization during metastasis. It is well established that primary tumor formation is driven by a combination of genetic and epigenetic events<sup>38</sup>. With respect to metastasis, studies have shown that primary and matched metastatic tumors are broadly similar at the genetic level with no recurrent mutations identified in metastases that were not present in the primary tumor<sup>3,4,39–44</sup>. These studies suggest that primary tumors are likely already genetically equipped with the ability to metastasize. Further, this implies that epigenetic processes may mediate the shift in cell state that accompanies metastatic progression, as proposed by others<sup>17,45–48</sup>. In accordance with this epigenetic hypothesis, we show that osteosarcoma metastasis is accompanied by a shift in epigenetic state at enhancer elements. Although our findings are not mutually exclusive with genetic theories of metastatic progression, we find that positive selection of enhancer activity is a fundamental component of the metastatic phenotype.

Our findings add to the growing body of evidence implicating epigenomic processes in metastasis. McDonald *et al.*<sup>17</sup> recently showed that pancreatic cancer metastasis is associated with widespread changes to heterochromatin, defined by H3K9 and H4K20 methylation. Denny *et al.*<sup>49</sup> found that differences in chromatin accessibility correlated with metastatic progression in NFIB-driven small cell lung cancer. Most recently, Roe *et al.*<sup>50</sup> showed that aberrant enhancer activity mediated by FOXA1 promotes metastasis in pancreatic cancer. Collectively, these studies demonstrate that chromatin changes drive the metastatic phenotype across various cancers. Our findings are well aligned with these studies but further implicate enhancer dysregulation as the basis of metastatic trait acquisition and demonstrate that such dysregulation presents an opportunity for the development of targeted antimetastatic therapies, as illustrated at the *F3* locus.

## METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

*Note:* Any Supplementary Information and Source Data files are available in the [online version of the paper](#).

## ACKNOWLEDGMENTS

The authors thank the members of the Tesar, Kaplan, and Helman laboratories for their input throughout the course of the project as well as B. Decker for his input on the manuscript text. Additional support was provided by the Genomics Core Facility of the Case Western Reserve University School of Medicine's Genetics and Genome Sciences Department and the Case Comprehensive Cancer Center (P30CA043703). This work was supported by the Liddy Shriver Sarcoma Initiative (P.C.S., C.K., J.J.M.), the QuadW Foundation (P.C.S.), Sarcoma Foundation of America (P.C.S.), St. Baldrick's Foundation (A.Y.H.), Alex's Lemonade Stand Foundation (A.Y.H.), Hyundai Hope-on-Wheels Program (A.Y.H.), Pediatric Cancer Research Foundation (A.Y.H.), CCCC AYA Oncology Pilot Grant (A.Y.H.), National Institutes of Health (NIH) grants F30

CA186633 (J.J.M.), F30 CA183510 (T.E.M.), T32 GM007250 (J.J.M., T.E.M., S.H.), R01CA193677 (P.C.S.), R01CA204279 (P.C.S.), R01CA160356 (P.C.S.), F31CA192874 (F.A.), R21CA218790 (A.Y.H.), NIH Intramural Visiting Fellow Program 15335 (M.M.L.), and NIH Intramural Research Program (C.K.).

#### AUTHOR CONTRIBUTIONS

J.J.M., C.K., and P.C.S. conceived the overall experimental design. J.J.M., C.F.B., and G.D. generated ChIP-seq, RNA-seq, and DHS-seq data. J.J.M., A.S., S.H., and P.C.S. completed analyses of ChIP-seq, RNA-seq, and DHS-seq data. J.J.M. and T.E.M. designed and completed the shRNA screening experiment and subsequent analysis. T.E.M. completed functional enrichment analysis of RNA-seq data. J.J.M. and I.B. generated 4C-seq data. J.J.M. and A.S. analyzed 4C-seq data. J.J.M., A.M., and I.B. and M.M.L. completed the *in vivo* and *ex vivo* metastasis experiments. J.J.M., J.T.M., and F.A. designed and completed the orthotopic metastasis experiments. J.J.M., D.R.C., and A.P.W.F. designed and completed the TALEN deletion experiments. M.Y.K. completed the *in vitro* F3 experiments. M.G., A.R., and P.P. provided subject tumor samples and clinical data. B.P.R. assessed F3 staining in subject tissue microarrays. A.D., A.Y.H., P.S.M., L.J.H., H.V., J.S., C.K., and P.C.S. provided the technical expertise and facilities to complete the experiments. J.J.M. and P.C.S. analyzed all data and wrote the paper. All authors provided intellectual input, edited, and approved the final manuscript.

#### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Valastyan, S. & Weinberg, R.A. Tumor metastasis: molecular insights and evolving paradigms. *Cell* **147**, 275–292 (2011).
2. Chambers, A.F., Groom, A.C. & MacDonald, I.C. Dissemination and growth of cancer cells in metastatic sites. *Nat. Rev. Cancer* **2**, 563–572 (2002).
3. Gundem, G. *et al.* The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**, 353–357 (2015).
4. Hong, M.K. *et al.* Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat. Commun.* **6**, 6605 (2015).
5. Bos, P.D. *et al.* Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005–1009 (2009).
6. Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537–549 (2003).
7. Minn, A.J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
8. Factor, D.C. *et al.* Epigenomic comparison reveals activation of “seed” enhancers during transition from naïve to primed pluripotency. *Cell Stem Cell* **14**, 854–863 (2014).
9. Gifford, C.A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
10. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642–654 (2013).
11. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
12. Akhtar-Zaidi, B. *et al.* Epigenomic enhancer profiling defines a signature of colon cancer. *Science* **336**, 736–739 (2012).
13. Cohen, A.J. *et al.* Hotspots of aberrant enhancer activity punctuate the colorectal cancer epigenome. *Nat. Commun.* **8**, 14400 (2017).
14. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
15. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
16. Ramaswamy, S., Ross, K.N., Lander, E.S. & Golub, T.R. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**, 49–54 (2003).
17. McDonald, O.G. *et al.* Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nat. Genet.* **49**, 367–376 (2017).
18. Kansara, M., Teng, M.W., Smyth, M.J. & Thomas, D.M. Translational biology of osteosarcoma. *Nat. Rev. Cancer* **14**, 722–735 (2014).
19. Huang, Y.M., Hou, C.H., Hou, S.M. & Yang, R.S. The metastasectomy and timing of pulmonary metastases on the outcome of osteosarcoma patients. *Clin. Med. Oncol.* **3**, 99–105 (2009).
20. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
21. Ren, L. *et al.* Characterization of the metastatic phenotype of a panel of established osteosarcoma cells. *Oncotarget* **6**, 29469–29481 (2015).
22. Zentner, G.E., Tesar, P.J. & Scacheri, P.C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* **21**, 1273–1283 (2011).
23. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
24. Huang, H., Bhat, A., Woodnutt, G. & Lappe, R. Targeting the ANGPT-TIE2 pathway in malignancy. *Nat. Rev. Cancer* **10**, 575–585 (2010).
25. Clayton, P.E., Banerjee, I., Murray, P.G. & Renehan, A.G. Growth hormone, the insulin-like growth factor axis, insulin and cancer risk. *Nat. Rev. Endocrinol.* **7**, 11–24 (2011).
26. Pinski, J. *et al.* Inhibition of growth of human osteosarcomas by antagonists of growth hormone-releasing hormone. *J. Natl. Cancer Inst.* **87**, 1787–1794 (1995).
27. Li, N. *et al.* Phosphodiesterase 10A: a novel target for selective inhibition of colon tumor cell growth and β-catenin-dependent TCF transcriptional activity. *Oncogene* **34**, 1499–1509 (2015).
28. van den Berg, Y.W., Osanto, S., Reitsma, P.H. & Versteeg, H.H. The relationship between tissue factor and cancer progression: insights from bench and bedside. *Blood* **119**, 924–932 (2012).
29. Mendoza, A. *et al.* Modeling metastasis biology and therapy in real time in the mouse lung. *J. Clin. Invest.* **120**, 2979–2988 (2010).
30. Corradin, O. *et al.* Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res.* **24**, 1–13 (2014).
31. Leaner, V.D. *et al.* Inhibition of AP-1 transcriptional activity blocks the migration, invasion, and experimental metastasis of murine osteosarcoma. *Am. J. Pathol.* **174**, 265–275 (2009).
32. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
33. Lamoureux, F. *et al.* Selective inhibition of BET bromodomain epigenetic signalling interferes with the bone-associated tumour vicious cycle. *Nat. Commun.* **5**, 3511 (2014).
34. Puissant, A. *et al.* Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
35. Bandopadhyay, P. *et al.* BET bromodomain inhibition of MYC-amplified medulloblastoma. *Clin. Cancer Res.* **20**, 912–925 (2014).
36. Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
37. Versteeg, H.H. *et al.* Inhibition of tissue factor signaling suppresses tumor growth. *Blood* **111**, 190–199 (2008).
38. You, J.S. & Jones, P.A. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* **22**, 9–20 (2012).
39. Jones, S. *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl. Acad. Sci. USA* **105**, 4283–4288 (2008).
40. Liu, W. *et al.* Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat. Med.* **15**, 559–565 (2009).
41. Campbell, P.J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
42. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
43. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
44. Moelans, C.B. *et al.* Genomic evolution from primary breast carcinoma to distant metastasis: Few copy number changes of breast cancer related genes. *Cancer Lett.* **344**, 138–146 (2014).
45. Kerbel, R.S., Frost, P., Liteplo, R., Carlow, D.A. & Elliott, B.E. Possible epigenetic mechanisms of tumor progression: induction of high-frequency heritable but phenotypically unstable changes in the tumorigenic and metastatic properties of tumor cell populations by 5-azacytidine treatment. *J. Cell. Physiol. Suppl.* **3**, 87–97 (1984).
46. Rodenkirch, D.I. Epigenetic contributions to cancer metastasis. *Clin. Exp. Metastasis* **26**, 5–18 (2009).
47. Javaid, S. *et al.* Dynamic chromatin modification sustains epithelial-mesenchymal transition following inducible expression of Snail-1. *Cell Rep.* **5**, 1679–1689 (2013).
48. Latil, M. *et al.* Cell-type-specific chromatin states differentially prime squamous cell carcinoma tumor-initiating cells for epithelial to mesenchymal transition. *Cell Stem Cell* **20**, 191–204.e5 (2017).
49. Denny, S.K. *et al.* Nfib promotes metastasis through a widespread increase in chromatin accessibility. *Cell* **166**, 328–342 (2016).
50. Roe, J.S. *et al.* Enhancer reprogramming promotes pancreatic cancer metastasis. *Cell* **170**, 875–888.e20 (2017).

## ONLINE METHODS

**Cell culture.** Human osteosarcoma cell lines were obtained and cultured as previously described<sup>21</sup>. MG63.3 cells were derived from MG63.2 through metastatic selection in mice as previously described<sup>51</sup>.

The metastatic properties of these clonally related parental and metastatic cell lines have been thoroughly characterized in multiple murine models of metastasis<sup>21</sup>. A sample size of five cell line pairs was chosen on the basis of similar studies completed by our lab in the past in order to capture the spectrum of methods of metastatic derivation and to sufficiently power comparative analyses of results.

The purity and authenticity of all lines used in these studies were independently confirmed through short tandem repeat (STR) profiling performed by the International Cell Line Authentication Committee. Mycoplasma testing was routinely performed with Mycoalert Mycoplasma detection kit (Lonza).

**Mouse studies.** All mice were housed and handled in accordance with protocols approved by the Case Western Reserve University Institutional Animal Care and Use Committee (IACUC) or the National Cancer Institute IACUC, depending on the location of the performed studies. The number of mice included in each of the described studies was based on extensive past experience in the development and use of murine models of metastases by our group. Each study was designed to minimize unnecessary mouse use, optimize statistical power, and account for known variance in each model system. Within each experiment, mice of the same strain, sex, and age were used for all conditions. At the initiation of each experiment, mice were randomly assigned to cages, and all mice in a given cage received equivalent treatment (for example, doxycycline). Researchers were not blinded to the group assignments of mice, as no subjective measurements were used.

**Human subjects.** Osteosarcoma primary tumors and lung metastatic tumors were obtained from the Laboratory of Experimental Oncology, Rizzoli Institute, Bologna, Italy with approval from Rizzoli Institute Ethics Committee and with informed written consent. A waiver was granted for informed consent from subjects deceased at the time of data collection, adhering to the European Union General Data Protection Regulation. Estimated tumor cellularity for the samples ranged from 50–90%.

**Ex vivo lung metastasis assay. Procedure for RNA isolation.** GFP<sup>+</sup> tumor cells ( $5 \times 10^5$ ) were injected into the tail vein of 8- to 10-week-old female SCID/Beige mice (Charles River). Within 15 min of tumor cell injection, mice were euthanized via CO<sub>2</sub> inhalation, and lungs were insufflated with a mixed agarose media solution. Lung sections for *ex vivo* culture were generated as previously described<sup>29</sup> and were incubated at 37 °C in humidified conditions of 5% CO<sub>2</sub>. Culture media was changed, and lung sections were flipped every 2 d. Tumor cell RNA was harvested at 24 h and 14 d after injection from one mouse for each condition. Lung sections were chopped into fine pieces and incubated in 3 ml Hank's balanced salt solution (HBSS) with 1 mg/ml collagenase at 37 °C for 30 min. EDTA was added to attain a final concentration of 10 mM, and the solution was placed on ice to stop digestion. Digested material was homogenized via passing it through an 18-gauge needle three to five times using a 10-ml syringe. Homogenate was passed through a 70-μm cell strainer (Corning Life Sciences) and centrifuged at 500g for 5 min at 4 °C. Supernatant was aspirated, and cells were resuspended in 5 ml ammonium-chloride-potassium (ACK) lysing buffer for 3 min at RT to lyse red blood cells. Lysis was stopped through addition of 10 ml HBSS, and cells were centrifuged at 500g for 5 min at 4 °C. Supernatant was aspirated, and cells were resuspended in 2–3 ml 0.5 mM EDTA-PBS and placed on ice. Immediately before sorting, cells were passed through a 40-μm cell strainer (Corning Life Sciences).

Cells were using FACS to isolate GFP<sup>+</sup> tumor cells and then were immediately centrifuged at 500g for 5 min at 4 °C. Supernatant was aspirated, cells were lysed in 1 ml TRIzol reagent (Life Technologies), and RNA was extracted with 200 μl chloroform. The organic phase was isolated, and 700 μl of ethanol was added. RNA was purified from this solution using the RNeasy Micro Kit (Qiagen).

**Procedure for F3 knockdown studies.** F3-knockdown cells were pretreated with 5 μg/ml doxycycline for 40 h in standard culture. Cells expressing shRNA targeting *F3* were sorted using FACS to isolate dsRed<sup>+</sup>GFP<sup>+</sup> fraction. Noninduced

control cells were sorted to isolate the GFP<sup>+</sup> fraction.  $2 \times 10^5$  cells were injected into the tail vein of each mouse. For the F3 knockdown condition, 5 μg/ml doxycycline was added to agarose media solution used to insufflate lungs as well as culture media. The medium was changed, and fresh doxycycline was added every 2 d. A total of 8 lung sections were imaged for each condition (4 sections per mouse, 2 mice per condition).

**Procedure for JQ1 studies.**  $5 \times 10^5$  tumor cells were injected into the tail vein of each mouse. For JQ1-treated cultures, medium was supplemented to a final concentration of 250 nM JQ1 by adding 10 mM DMSO stock solution. Vehicle-treated culture medium was supplemented with DMSO volumes matching JQ1 treatment. The medium was changed, and fresh JQ1 or DMSO was added every 2 d. A total of 8 lung sections were imaged for each condition (4 sections per mouse, 2 mice per condition).

**Assessment of metastatic burden.** Lung sections were imaged by inverted fluorescent microscopy (Leica DM IRB) at a magnification of 2.5×. Two or three images per lung section were taken to capture the entire surface of each section. Image analysis was performed using ImageJ software to quantify the total GFP<sup>+</sup> area per lung section. The metastatic burden was calculated by normalizing the total GFP<sup>+</sup> area to the GFP<sup>+</sup> area for each section on day 0. The values reported represent the mean normalized tumor burden for all sections for each condition (8 sections per condition).

**In vitro RNA isolation.** To match the conditions used to isolate cells from *ex vivo* lung sections, cells growing *in vitro* were trypsinized, exposed to the same mechanical and enzymatic digestion conditions, and sorted through FACS as described above.  $5 \times 10^5$  GFP<sup>+</sup> cells were collected, and RNA was isolated as described above.

**ChIP-seq.** ChIPs were performed using  $5 \times 10^6$  to  $10 \times 10^6$  cross-linked cells, and sequencing libraries were prepared as previously described<sup>52</sup>. The following antibodies were used for ChIP: rabbit anti-H3K4me1 (Abcam no. 8895), rabbit anti-H3K27ac (Abcam no. 4729), rabbit anti-c-Fos (Santa Cruz no. sc-52), and rabbit anti-FOSL1/Fra-1 (Santa Cruz no. sc-605). ChIP-seq libraries were sequenced on the HiSeq 2000 or 2500 platform at the Case Western Reserve University Genomics Core Facility. Analysis was performed as previously described<sup>8</sup>.

**Met-VEL analysis.** H3K4me1 ChIP-seq peaks were filtered to remove all peaks overlapping ENCODE blacklisted regions for functional genomics analysis (<https://sites.google.com/site/anshulkundaje/projects/blacklists/>) as well as peaks  $\pm$  1 kb from transcription start sites of all annotated RefSeq genes to exclude promoters. Resulting peak lists of parental and metastatic cell line pairs were merged, and RPKM values within merged peaks were calculated. Gained and lost Met-VELs were called as peaks with RPKM values in metastatic cell lines that were increased or decreased threefold relative to parental cell lines, respectively. To determine the fraction of differentially active enhancers in different cell types (**Supplementary Fig. 2**), H3K4me1 ChIP-seq peaks for each pair of samples were filtered for ENCODE blacklisted regions and promoters, concatenated, and merged. Peak RPKMs were calculated for each sample in a pair and floored to 0.3. Differentially active enhancers were defined as those showing a threefold change in H3K4me1 signal in one sample relative to the other. The fractions of differentially active enhancers for the osteosarcoma tumors, and cell line panels were based on averages for each group.

**Met-VEL clustering analysis.** Global Met VEL distribution was assessed by calculating Met-VEL counts in 200-kb sliding windows across all chromosomes. Met-VEL islands were defined as regions bordered by 200-kb windows with Met-VEL counts of 0. 200-kb windows with maximum Met-VEL counts in each Met-VEL island were identified. To test for nonrandom Met-VEL distribution, the same analysis was performed on 1,000 Met-VEL-size-matched H3K4me1 peak lists randomly sampled from all H3K4me1 peaks in the cell line being analyzed to account for global enhancer distribution biases. Metastatic cell line H3K4me1 peaks were sampled to assess gained Met-VEL clustering. Parental cell line H3K4me1 peaks were sampled to assess lost Met-VEL clustering. The sampled lists were used to define expected distributions of random VEL acquisition in each cell line. Expected distributions were compared to observed distributions to test the null hypothesis of random Met-VEL acquisition.

A *P* value threshold of 0.05 was used to reject the null hypothesis in support of nonrandom acquisition of Met-VELs. 200-kb windows with Met-VEL counts exceeding these thresholds were called as Met-VEL clusters.

**Super-enhancers.** Metastatic cell line- and parental cell line-specific SEs were identified from their H3K27ac profiles using the ROSE software (<https://github.com/BradnerLab/pipeline/>). Analysis was performed as previously described<sup>8</sup>.

**RNA-seq.** Gene expression profiles of cell lines grown *in vitro* were compared to expression profiles of the same cell lines at various time points during metastatic colonization using the *ex vivo* pulmonary metastasis assay. RNA quality was assessed by 2200 TapeStation Instrument (Agilent). PolyA<sup>+</sup> RNA was isolated using the Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer's protocol. RNA-seq libraries were sequenced on the Illumina HiSeq 2000 or 2500 platform at the Case Western Reserve University Next Generation Sequencing Core Facility. Expression analysis was performed as previously described, aligning to the hg19 genome build<sup>8</sup>. FPKMs were quantile normalized across all samples.

**Prediction of gene targets of enhancers.** Enhancer-gene assignments were made as previously described<sup>30</sup>. To identify Met-VEL gene targets in subject tissues that we did not have RNA-seq data for, we used the GREAT software package<sup>53</sup> to generate an initial list of candidate Met-VEL gene targets. We further filtered this list by only including genes within the same topologically associating domain as an enhancer with a H3K27ac peak at its promoter using a *P* value cutoff of  $P < 1 \times 10^{-5}$ . We then assessed overlaps in predicted Met-VEL gene targets between cell lines and subject tissues using the resulting gene lists.

**Gene ontology analysis.** *Gained Met-VEL gene lists.* Met-VEL gene lists were imported into gProfiler<sup>54</sup> to generate enrichment scores for all GO, Kyoto encyclopedia of genes and genomes (KEGG), and REACTOME gene sets according to the recommended settings for gProfiler (<http://baderlab.org/Software/EnrichmentMap/GProfilerTutorial/>). Cytoscape (v3.2.1) and the Enrichment Map<sup>55</sup> plug-in were used to generate networks for enriched gene sets with a false discovery rate cutoff of  $<0.05$ .

*Lost Met-VEL gene lists.* For GO analysis, the genes associated with Met-VELs were analyzed using DAVID (<http://david.abcc.ncifcrf.gov/home.jsp>). A *P* value of  $10^{-3}$  was used as the threshold for significant enrichment of an ontologic category. Categories significantly enriched for genes associated with gained or lost Met-VELs in two or more pairs are reported, limiting overlapping lists to the three top-scoring categories in each cell line (i.e., the categories with the lowest *P* values).

**DHS-seq.**  $6.4 \times 10^6$ – $56 \times 10^6$  cells from each cell line were sequenced for DNase hypersensitivity as previously described<sup>56</sup>. A ribose 5-phosphate was added to linker 1B to increase ligation efficiency. After DNase concentrations were optimized for each line, a total of approximately  $1 \times 10^6$  cells from optimally digested conditions were processed for sequencing. Libraries were sequenced on the HiSeq 2500 platform at the Case Western Reserve University Genomics Core Facility. Analysis was performed as previously described<sup>8</sup>.

**Chromosome conformation capture sequencing.** Chromosome conformation capture sequencing (4C-seq) sample preparation was performed as previously described<sup>57</sup>. NlaIII served as a primary restriction enzyme, DpnII as a secondary 4-bp cutter. The primer sequences used for 4C-seq analysis are as follows: forward, 5'-CTGTAAACTTGCCAAGCATG-3'; reverse, 5'-GGTCGCACAGGAGAAAGG-3'. Amplified sample libraries were pooled and spiked with 40% PhiX viral genome sequencing library to increase sample diversity. Multiplexed sequencing was performed on the MiSeq platform. Demultiplexing was performed using an in-house algorithm, and all reads were hard trimmed to 36 bp. Clipping of the primer sequences and data processing were performed using 4Cseqpipe Version 0.7 ([http://compgenomics.weizmann.ac.il/tanay/?page\\_id=367](http://compgenomics.weizmann.ac.il/tanay/?page_id=367)). The viewpoint reads were aligned to a fragmented genome, as determined by the restriction site positions of the chosen primary and secondary restriction enzymes. A running linearly weighted mean,

calculated in 2- to 50-kb sliding windows, was used for signal smoothing of each genomic bin (size 16 bp). Contact enrichment sites along the chromosomal axis were visually inspected.

**Motif analysis.** To identify TF motifs enriched in Met-VEL peaks, enhancers were centered on DNase hypersensitivity sites and the SeqPos module of the Cistrome tool was used to scan a 1-kb window for enriched curated motifs<sup>58</sup>. Significantly enriched motifs in each cell line were then filtered using RNA-seq data, and only expressed TFs were used for downstream analysis. Expressed TFs with enriched motifs in 3 out of 3 metastatic/parental cell line pairs (MG63/3/MG63; MNNG/HOS; 143B/HOS) are presented in the results.

**In vivo RNAi high-throughput functional assay.** *Vector construction.* The Tet-ON lentiviral construct was made via modifying the previously published optimized shRNAmir, miR-E, with a pRRRL backbone<sup>36</sup>. Briefly, this construct contains an optimized 3rd-generation Tet-responsive element and reverse tetracycline transactivator to potentiate a positive feedback loop, enhancing expression of the construct upon induction and reducing construct leakiness. The version of the construct that we modified contained a constitutive Venus reporter and an induced dsRed reporter of expression (LT3REVR). The construct was modified using standard cloning techniques to replace the Venus reporter with a puromycin-resistance element (renamed LT3REPIR) so that cells already constitutively expressing GFP could be selected for transduction.

*shRNA library generation.* shRNAs targeting 33 genes were selected from the transOMIC technologies shERWOOD-UltramiR shRNA library (3 or 4 shRNAs per gene). Cloning of shRNA into the backbone construct was performed on contract by transOMIC technologies. The shRNA sequences that were included are listed in **Supplementary Table 3**, where scores indicate shERWOOD metric of predicted potency of each shRNA as assigned by previously described algorithm<sup>59</sup>. "#N/A" indicates that the shRNA failed to clone into the lentiviral backbone.

*Lentiviral production.* VSV-G-pseudotyped lentivirus was generated through standard laboratory techniques. Briefly, shRNA-LT3REPIR plasmids were co-transfected with packaging vectors psPAX2 and pCI-VSVG (Addgene) into 293FT cells using Cal-Phos Mammalian Transfection Kit (Clontech). Individual supernatants containing virus were harvested at 48 and 72 h post-transfection and filtered with a 0.45-μm PVDF membrane (Millipore).

*Lentiviral transduction and selection.* Transduction of MG63.3 was performed via 24-h exposure to lentivirus in the presence of 8 μg/ml polybrene using conditions to achieve  $>1000\times$  coverage of each shRNA in the library. Infection rate was estimated to be 0.135% and was predicted to achieve predominantly one lentiviral integration per cell. Transduced and nontransduced cells were then treated with 2 μg/ml puromycin. Transduced cells were selected until all cells in the nontransduced plate were dead (2–4 d) to obtain a pure population of transduced cells (MG63.3i).

*High-throughput in vivo functional assay.* 8- to 10-week-old female SCID/Beige (Charles River) mice were used for the *in vivo* study arm. Mice were fed Dox Diet pellets containing 200 mg per kg of pellets doxycycline (Bio-Serve) for 5 d before injection of cells. MG63.3i cells were pretreated with 5 μg/ml doxycycline for 12 h in standard culture before being delivered to mouse lungs through tail vein injection. This primed the cells, but knockdown was not achieved even at the transcript level until 24–72 h after doxycycline addition.  $1.5 \times 10^6$  cells were injected into each mouse ( $n = 15$ ). Mice were maintained on a Dox diet throughout the 21-d course of the experiment. At the conclusion of the experiment, mice were euthanized through CO<sub>2</sub> inhalation, and lungs were surgically extracted and homogenized using the Tumor Dissociation Kit, human (Miltenyi) according to the manufacturer's protocol. Mouse lung cells were depleted using the Mouse Cell Depletion Kit (Miltenyi) according to the manufacturer's protocol. Lungs from 5 mice were pooled for each replicate to achieve 1,000× engraftment coverage of each shRNA in the library. GFP<sup>+</sup>dsRed<sup>+</sup> cells were then isolated through FACS.

Three replicates of  $1.5 \times 10^6$  MG63.3i cells growing *in vitro* were induced with 5 μg/ml doxycycline and maintained on doxycycline over 21 d in culture. GFP<sup>+</sup>dsRed<sup>+</sup> cells were isolated by FACS. Sorted cell counts of *in vitro* replicates were matched to numbers of cells isolated from the *in vivo* arm ( $3 \times 10^5$ – $5 \times 10^5$ ).

DNA was isolated from three replicates of uninduced MG63.3i cells as well as *in vivo* and *in vitro* arms of the experiment.

**shRNA amplification and sequencing.** Genomic DNA was isolated and sequenced as previously described<sup>60</sup> but with slight modification. Genomic DNA was isolated by two rounds of phenol extraction using 5PRIME Phase Lock tubes (Quantabio) followed by isopropanol precipitation. Deep sequencing libraries were generated through PCR amplification of shRNA guide strands using barcoded primers that tag the product with standard Illumina adapters. For each sample, DNA from at least  $3 \times 10^5$  cells was used as template in multiple parallel 50-μl PCR reactions, each containing 1 μg template, 1× AmpliTaq Gold buffer, 0.2 mM of each dNTP, 0.3 μM of each primer and 2.5 U AmpliTaq Gold (Applied Biosystems), which were run using the following cycling parameters: 95 °C for 10 min; 35 cycles of 95 °C for 30 s, 52 °C for 45 s and 72 °C for 60 s; and 72 °C for 7 min. PCR products (340 nt) were combined for each sample, precipitated, and purified on a 2% agarose gel (QIAquick gel extraction kit, Qiagen). Libraries were sequenced on the HiSeq 2500 platform at the Case Western Reserve University Genomics Core Facility. Libraries were sequenced using a primer that reads in reverse into the guide strand (miR30EcoRISeq, 5'-TAGCCCCTGAAATTCCGAGGCAGTAGGCA-3'). To provide a sufficient baseline for detecting shRNA depletion in experimental samples, we aimed to acquire >1,000 reads per shRNA in all samples, or  $1.37 \times 10^5$  reads per sample. In practice, we achieved > $2 \times 10^6$  reads for all samples. Sequence processing was performed using two custom workflows using <https://usegalaxy.org/><sup>61</sup>. Workflow can be accessed at the following links: <https://usegalaxy.org/u/tyleremiller/w/shrna-pipeline1> and <https://usegalaxy.org/u/tyleremiller/w/shrnastep2>. For each shRNA and condition, the number of matching reads was normalized to the total read number per lane. This measure of normalized coverage was used for all downstream analyses.

**Inducible knockdown of F3.** shRNAs targeting *F3* were selected from the transOMIC technologies shERWOOD-UltramiR shRNA library and cloned into the LT3REPIR vector as described in the preceding section. Cloning of shRNA into the backbone construct was performed on contract by transOMIC technologies. The shRNA sequences that were tested are listed in **Supplementary Table 3**.

**Lentiviral production.** VSV-G-pseudotyped lentivirus was generated through standard laboratory techniques. Briefly, shRNA-LT3REPIR plasmids were cotransfected with packaging vectors pSPAX2 and pCI-VSVG (Addgene) into 293FT cells using Cal-Phos Mammalian Transfection Kit (Clontech). Individual supernatants containing virus were harvested at 48 and 72 h post-transfection and filtered with 0.45 μm PVDF membrane (Millipore).

**Lentiviral transduction and selection.** Transduction was performed via 24-h exposure to lentivirus in the presence of 8 μg/ml polybrene. Transduced and nontransduced cells were then treated with 2 μg/ml puromycin. Transduced cells were selected until all cells in the nontransduced plate were dead (2–4 d).

**Assessment of knockdown.** Optimal shRNA induction was assessed and found to occur with 5 μg/ml doxycycline treatment. MG63.3 cells transduced with shF3A, shF3B, and shF3C were treated with doxycycline for 40 h, trypsinized, and sorted to isolate the GFP<sup>+</sup>dsRed<sup>+</sup> fraction. Noninduced cells were sorted to isolate the GFP<sup>+</sup> fraction. RNA was extracted from  $1 \times 10^6$  cells and purified using the RNAeasy Micro kit (Qiagen) according to the manufacturer's protocol. RNA quality was assessed by 2200 TapeStation Instrument (Agilent). cDNA was synthesized using the High Capacity RNA-to-cDNA kit (ABI) according to the manufacturer's protocol. Knockdown efficiency was determined through RT-qPCR analysis for *F3* using optimized TaqMan Gene Expression Assay primers and TaqMan Gene Expression Master Mix (Life Technologies). The probes used for the RT-qPCR are as follows: *F3*, Hs01076029\_m1; GAPDH, Hs02758991\_g1.

RT-qPCR was performed to quantify percent knockdown with induction of each hairpin relative to noninduced controls. The shF3A and shF3B hairpins showed the highest degree of *F3* knockdown and were chosen for use in the remainder of our studies.

**In vivo experimental metastasis model.** 8- to 10-week-old female SCID/Beige (Charles River) mice were used for all experimental metastasis studies. For all *in vivo* experimental *F3* knockdown studies, mice in the *F3* knockdown group were fed Dox Diet pellets containing 2 mg per kg body weight doxycycline

(Bio-Serve) for 5 d before injection of cells. *F3* knockdown cells were pretreated with 5 μg/ml doxycycline for 24 h in standard culture. Cells with induced expression of shRNA targetting *F3* were sorted through FACS to isolate the GFP<sup>+</sup>dsRed<sup>+</sup> fraction. Control cells were sorted to isolate the GFP<sup>+</sup> fraction.

**End-point assessment of lung metastasis.**  $5 \times 10^5$  MG63.3 GFP<sup>+</sup>dsRed<sup>+</sup> or MG63.3 GFP<sup>+</sup> cells were injected into the tail vein of each mouse ( $n = 10$  mice per condition). Mice in the *F3* knockdown group were maintained on the Dox Diet throughout the experiment. On day 7 or day 14 following injection, 5 mice from each group were euthanized through CO<sub>2</sub> inhalation. Lungs were insufflated with PBS and imaged through inverted fluorescent microscopy (Leica DM IRB) at a magnification of 2.5×. Five images per lung were taken to assess metastatic burden in each mouse.

Image analysis was performed using ImageJ software to quantify the total GFP<sup>+</sup> area per image. The metastatic burden was calculated as the sum of the total GFP<sup>+</sup> area in the five images from each mouse.

**Survival analysis.**  $5 \times 10^4$  cells were injected into the tail vein of each mouse ( $n = 5$  mice per condition). Mice in the *F3* knockdown group were maintained on the Dox Diet throughout the experiment. All mice that died underwent complete necropsy examination and confirmation of metastasis.

**Orthotopic spontaneous lung metastasis model.** 8- to 10-week-old female NSG mice (Jackson) were used for the spontaneous metastasis studies. Mice in the *F3* knockdown group were given water supplemented with 2 mg/ml doxycycline hydralate (Sigma) and 2% sucrose for 5 d before injection of cells. *F3* knockdown cells were pretreated with 5 μg/ml doxycycline for 24 h in standard culture. *F3* shRNA induced cells were sorted through FACS to isolate the GFP<sup>+</sup>dsRed<sup>+</sup> fraction. Control cells were sorted to isolate the GFP<sup>+</sup> fraction.  $3 \times 10^5$  cells were injected orthotopically into the paraosseous region adjacent to the left proximal tibia. For the *F3* knockdown group, water was changed and fresh doxycycline was added twice weekly. Injection sites were monitored for tumor formation twice weekly.

Tumors became measurable on day 21 for all groups, at which time tumors were measured in two dimensions twice weekly. Tumor volume was calculated as follows: volume (mm<sup>3</sup>) =  $3.14 \times (\text{long dimension (mm)}) \times (\text{short dimension (mm)})^2$ . The experiment was terminated following 21 d of tumor measurements (42 d after injections), and mice were euthanized through CO<sub>2</sub> inhalation. Lungs were insufflated with PBS and imaged through fluorescent microscopy at a magnification of 2.5× using a Leica DM 5500B light microscope with a Leica DFC 500 camera. Five images per lung were taken to assess the metastatic burden in each mouse.

Image analysis was performed using ImageJ software. Micrometastases were defined as GFP<sup>+</sup> lesions with a diameter >25 pixels in images captured at 2.5× magnification. The number of micrometastases per image was manually counted. The total number of micrometastases for each lung was calculated as the sum of the total number of micrometastases in 5 images from each mouse.

**In vitro assay for FXa formation.** Cells growing *in vitro* were pretreated with 25 μg/ml IgG control, monoclonal antibody 10H10, or monoclonal antibody 5G9 20 min before assay. *F3* inhibitory antibodies (Mab-10H10 and Mab-5G9) were acquired from the Versteeg laboratory. Validation experiments presented in ref. 37. Cells were washed in serum-free DMEM, and Xa generation over time was measured 30 min after the addition of 1nM FVIIa and 50 nM FX using the chromogenic substrate Spectrozyme FXa.

**Assessment of F3-inhibiting antibodies on metastatic progression.**  $5 \times 10^5$  GFP<sup>+</sup> MG63.3 cells were mixed with 500 μg IgG control, monoclonal antibody 10H10, or monoclonal antibody 5G9 and injected into the tail vein of 10- to 12-week-old female SCID-beige mice ( $n > 5$  mice per group). Mice were euthanized 14 d after injection, and metastatic burden was assessed through whole-lung fluorescent imaging (5 images per mouse). Metastatic burden was quantified as the total GFP<sup>+</sup> area per mouse.

**F3 lung metastasis staining.** To assess *F3* expression in metastatic tumor cells in lung at progressive time points, mice were injected with  $1 \times 10^6$  cells (via tail vein) and were euthanized via CO<sub>2</sub> inhalation at 24 h and 15 d post-injection. Lungs were harvested, formalin-fixed, and paraffin-embedded. For *ex vivo* lung

metastasis staining, the protocol described above was followed and lung sections were fixed at 24 h and 5 d post-injection. Tissue sections of lungs were cut at a thickness of 5 µm. Prior to immunostaining, paraffin sections were dewaxed with xylenes and rehydrated with an ethanol series. For antigen retrieval, tissue sections were immersed 95 °C Target Retrieval Solution (DAKO) for 25 min. Tissue sections were permeabilized with 0.01% Triton-X in PBS for 10 min. Slides were rinsed with PBS and blocked with 4% BSA in PBS for 10 min. The following primary antibodies were used: F3, rabbit monoclonal IgG F3 antibody (ab151748, Abcam); GFP, goat polyclonal IgG GFP antibody conjugated to fluorescein isothiocyanate (FITC; ab6662, Abcam). Primary antibodies were diluted in 4% BSA (1:100), and slides were incubated in antibody solution at 4 °C overnight. Slides were rinsed and incubated with goat polyclonal IgG anti-rabbit IgG (H+L) conjugated to Alexa 594 (A-11037, Life Technologies) diluted 1:200 in 4% BSA for 1 h in a dark humidified slide chamber. Nuclei were visualized with DAPI (Sigma, 1 µg/ml). Tissue sections were mounted on slides using anti-fade mounting medium (Vectashield).

Stained sections were imaged by fluorescent microscopy at 20× or 40× using a Leica DM 5500B light microscope with a Leica DFC 500 camera. Image analysis was performed using ImageJ software. F3 expression was computed within the GFP<sup>+</sup> metastatic tumor cell area.

**Tissue array staining and scoring.** A tissue microarray that we previously developed containing 20 lung metastases from individuals with osteosarcoma<sup>62</sup> was assessed for F3 expression. 18 of these 20 samples had cores of sufficient quality on the stained slide to be scored. For staining, paraffin was removed through a 5 min incubation in a xylene bath performed twice and then rehydrated using step-down concentrations of ethanol. Antigen retrieval was performed through incubation in 1:10 dilution of Target Retrieval Solution (DAKO) in a steamer for 25 min at 95 °C. Cells were permeabilized with 0.01% Triton-X in PBS for 10 min. F3 immunohistochemical staining was performed using rabbit monoclonal IgG F3 antibody (ab151748, Abcam) and the EnVision + System-HRP (Dako) according to the manufacturer's protocol. Cover slips were mounted on slides using anti-fade mounting medium.

The array was scored by the Director of Soft Tissue Pathology at the Cleveland Clinic and Learner Research Institute who was blinded to the sample type. Cores were scored based on the percentage of tumor cells in the core with positive staining for F3 (0 = 0%; 1+ = 1–25%; 2+ = 26–50%; 3+ = >50%) and the intensity of F3 staining in positive areas (low-intensity staining or high-intensity staining). Individual cores were excluded from the analysis if no tumor was present, if the tumor was predominantly necrotic, or if the core was falling off the slide.

**Targeted deletion of an F3 Met-VEL by genome engineering.** Two TALEN dimers were designed to target the flanks of the Met-VEL in the *F3* locus, as indicated in **Figure 5a**. TALEN dimers recognized the sequences 5'-ACCAAC TCACTTGAGCTGtgtttttcttCAGTCACAATTGTGAAAT-3' and 5'-GA ATCGACTGATCAAAGCacatgaactttaaaaaaGAGTAATAAGTTACTT-3', where spacer elements are in lower case. TALEN constructs were assembled with adaptations of previously described protocols<sup>63,64</sup>. MG63.3 cells were grown in a 6-well plate format to 70% confluence and transfected with 2.5 µg plasmid

for each TALEN monomer using Lipofectamine 2000 (ThermoFisher Scientific) per the manufacturer's instructions. Cells were incubated for 48 h at 30 °C, and genomic DNA was subsequently harvested using QuickExtract DNA Extraction Solution (Epicentre) as recommended by the supplier. Efficient deletion of the *F3* Met-VEL was confirmed through agarose gel electrophoresis of PCR products generated using the primers 5'-GCAGTGCACAACCTGTACAAC-3' and 5'-TTGCCAGGGTCATTATGTT-3' (Integrated DNA Technologies) and high-fidelity AccuPrime *Taq* DNA Polymerase (ThermoFisher Scientific). Single-cell clones were derived through limiting dilution, and genotypes were confirmed as described above. Enhancer deletion of the clonal cell population used for functional metastasis experiments was confirmed through Sanger sequencing using the primer sequences listed above.

**Life Sciences Reporting Summary.** Further information on experimental design is available in the **Life Sciences Reporting Summary**.

**Data availability.** The datasets generated during and analyzed during the current study are included within the published manuscript (and supplementary information) or have been deposited in Gene Expression Omnibus (RNA-seq, ChIP-seq, DNase-seq data) under accession number [GSE74230](#). All other data are available from the corresponding authors upon reasonable request.

51. Khanna, C. *et al.* An orthotopic model of murine osteosarcoma with clonally related variants differing in pulmonary metastatic potential. *Clin. Exp. Metastasis* **18**, 261–271 (2000).
52. Schmidt, D. *et al.* ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods* **48**, 240–248 (2009).
53. McLean, C.Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
54. Reimand, J., Arak, T. & Vilo, J. g:Profiler—a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res.* **39**, W307–15 (2011).
55. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
56. Song, L. & Crawford, G.E. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb. Protoc.* <http://dx.doi.org/10.1101/pdb.prot5384> (2010).
57. van de Werken, H.J. *et al.* Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).
58. Liu, T. *et al.* Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* **12**, R83 (2011).
59. Knott, S.R.V. *et al.* A computational algorithm to predict shRNA potency. *Mol. Cell* **56**, 796–807 (2014).
60. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol.* **29**, 79–83 (2011).
61. Goecks, J., Nekrutenko, A. & Taylor, J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).
62. Osborne, T.S. *et al.* Evaluation of eIF4E expression in an osteosarcoma-specific tissue microarray. *J. Pediatr. Hematol. Oncol.* **33**, 524–528 (2011).
63. Sakuma, T. *et al.* Efficient TALEN construction and evaluation methods for human cell and animal applications. *Genes Cells* **18**, 315–326 (2013).
64. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.

For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

### ► Experimental design

#### 1. Sample size

Describe how sample size was determined.

The number of animals included in each of the described studies was based on extensive past experience in the development and use of murine models of metastasis by our group. Each study was designed to minimize unnecessary animal use, optimize statistical power, and account for known variance in each model system.

#### 2. Data exclusions

Describe any data exclusions.

None.

#### 3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

Replicate experiments were performed as described in the text. In all such cases, key results were reproducible.

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

At the initiation of each experiment mice were randomly assigned to cages and all mice in a given cage received equivalent treatment (e.g. doxycycline).

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Researchers were not blinded to the group assignments of mice as no subjective measurements were used.

Researchers were blinded to sample type for scoring F3 staining of tissue microarray experiment.

For metastatic image analyses, standardized computational approaches were used and researchers were not blinded to sample.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

## 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- Test values indicating whether an effect is present  
*Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on [statistics for biologists](#) for further resources and guidance.

## ► Software

Policy information about [availability of computer code](#)

### 7. Software

Describe the software used to analyze the data in this study.

Quantitative image analysis performed using ImageJ v2.0.0-rc-43/1.50i Build: 49b667f9aa

Super-Enhancers were called using ROSE2\_main.py

Gene-enhancer predictions were called using GREAT v3.0.0 and PreSTIGE

Gene Ontology Analysis was performed with gProfiler. Maps made with Cytoscape v3.2.1 Enrichment Map plug-in or DAVID.

4C-seq data was analyzed using 4Cseqpipe v 0.7.

Transcription factor motif analysis was performed using SeqPos module of Cistrome.

shRNA screen data was processed using custom workflows on galaxy:  
<https://usegalaxy.org/u/tyleremiller/w/shrna-pipeline1>  
<https://usegalaxy.org/u/tyleremiller/w/shrnastep2>

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

## ► Materials and reagents

Policy information about [availability of materials](#)

### 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

All unique materials are available from the authors or commercial sources.

### 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

F3 antibody for tissue microarray staining: ab151748. Validation data available on Abcam website.

F3 inhibitory antibodies (Mab-10H10 and Mab-5G9) were acquired from the Vergsteeg laboratory. Validation experiments presented in Versteeg et al., Blood 2008 (PMID:17901245).

ChIP-seq antibodies fully annotated in separate form.

## 10. Eukaryotic cell lines

- a. State the source of each eukaryotic cell line used.

Human osteosarcoma cell lines HOS, SaOS-2, MG63, MNNG, and 143B were purchased from ATCC. Hu09 and M112 cell lines were obtained from Dr. Jun Yokota (National Cancer Center Research Institute, Tokyo, Japan). LM7 was obtained from Dr. Eugenie S. Kleinerman (The University of Texas: MD Anderson Cancer Center, Houston, TX). MG63.3 cells were derived from MG63.2 (obtained from Dr. Hue Luu, University of Chicago, Chicago, IL) by metastatic selection in mice as previously described (Khanna et al., Clin Exp Metastasis 2000; PMID: 11315100).

The metastatic properties of these clonally related parental and metastatic cell lines has been thoroughly characterized in multiple murine models of metastasis (Ren et al., Oncotarget 2015; PMID: 26320182).

- b. Describe the method of cell line authentication used.

The purity and authenticity of all lines used in these studies has been independently confirmed by short tandem repeat (STR) profiling performed by the International Cell Line Authentication Committee.

- c. Report whether the cell lines were tested for mycoplasma contamination.

Mycoplasma testing was routinely performed with Mycoalert Mycoplasma detection kit (Lonza).

- d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified lines were used.

## ► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

### 11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

Ex Vivo Lung Metastasis Assay: 8-10 week old female SCID/Beige (Charles River).  
High-Throughput In Vivo Functional Assay: 8-10 week old female SCID/Beige (Charles River).  
In Vivo Experimental Metastasis Model: 8-10 week old female SCID/Beige (Charles River).  
Orthotopic Spontaneous Metastasis Model: 8-10 week old female NSG (Jackson).

Policy information about [studies involving human research participants](#)

### 12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Lung Metastasis 1: 16 year old male, no chemotherapy prior to primary tumor resection, metastasis resected 68 months after diagnosis.  
Lung Metastasis 2: 15 year old male, no chemo prior to primary tumor resection, metastasis resected 3 months after diagnosis.  
Lung Metastasis 3: 18 year old male, chemotherapy prior to primary tumor resection, metastasis resected 9 months after diagnosis.  
Lung Metastasis 4: 73 year old female, chemotherapy prior to primary tumor resection, metastasis resected 6 months after diagnosis.  
Lung Metastasis 5: 68 year old female, no chemotherapy prior to primary tumor resection, metastasis resected 51 months after diagnosis.  
Lung Metastasis 6: 23 year old female, no chemotherapy prior to primary tumor resection, metastasis resected 46 months after diagnosis.  
Lung Metastasis 7: 12 year old female, chemotherapy prior to primary tumor resection, metastasis resected 17 months after diagnosis.  
Lung Metastasis 8: 15 year old male, no chemotherapy prior to primary tumor resection, metastasis resected 28 months after diagnosis. Same patient as Lung Metastasis 2.  
Lung Metastasis 9: 33 year old male, no chemotherapy prior to primary tumor resection, metastasis resected 30 months after diagnosis.  
Lung Metastasis 10: 33 year old male, no chemotherapy prior to primary tumor resection, metastasis resected 30 months after diagnosis. Same patient as Lung Metastasis 10.