HHS Public Access

Author manuscript

Nat Cell Biol. Author manuscript; available in PMC 2023 July 05.

Published in final edited form as:

Nat Cell Biol. 2023 January ; 25(1): 145–158. doi:10.1038/s41556-022-01045-0.

MLL3 Loss Drives Metastasis by Promoting a Hybrid Epithelial-Mesenchymal Transition State

Jihong Cui[1,2], **Chi Zhang**[2], **Ji-Eun Lee**[3], **Boris A. Bartholdy**[2], **Dapeng Yang**[4], **Yu Liu**[1,2], **Piril Erler**[1,2], **Phillip M Galbo Jr**[5], **Dayle Q. Hodge**[1,2], **Danwei Huangfu**[4], **Deyou Zheng**[5,6], **Kai Ge**[3], **Wenjun Guo**[1,2,7,*]

[1]Ruth L. and David S. Gottesman Institute for Stem Cell and Regenerative Medicine Research, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[2]Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[3]National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20814, USA

[4]Developmental Biology Program, Sloan Kettering Institute, 1275 York Avenue, New York, NY 10065, USA

[5]Departments of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[6]Departments of Neurology and Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[7]Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, NY 10461, USA

Abstract

Phenotypic plasticity associated with the hybrid epithelial-mesenchymal transition (EMT) is crucial to metastatic seeding and outgrowth. However, the mechanisms governing the hybrid EMT state remain poorly defined. We showed that deletion of the epigenetic regulator MLL3, a tumor suppressor frequently altered in human cancer, promoted the acquisition of hybrid EMT in breast cancer cells. Distinct from other EMT regulators that only mediate unidirectional changes, MLL3 loss enhanced responses to stimuli inducing EMT and mesenchymal-epithelial transition in epithelial and mesenchymal cells, respectively. Consequently, MLL3 loss greatly increased metastasis by enhancing metastatic colonization. Mechanistically, MLL3 loss led to increased IFNγ signaling, which contributed to the induction of hybrid EMT cells and enhanced metastatic capacity. Furthermore, BET inhibition effectively suppressed the growth of *MLL3*-mutant primary

*Correspondence: wenjun.guo@einsteinmed.edu.

Author Contributions Statement

J.C. and W.G. conceived the study; J.C. performed the experiments, acquired and analyzed data; C.Z and D.Z. carried out the bioinformatics analyses of ChIP-seq data. B.A.B. analyzed ATAC-seq data. Y.L. analyzed the transcriptomics data; J.E.L and K.G. generated MLL3 reagents and performed ChIP-seq experiments; D.Y and D.H performed the histone mark ChIP-seq experiments. P.E. generated MaSC models; P.G. and D.Z. performed patient data analysis; D.Q.H. generated CRISPR reagents; all authors contributed to data interpretation; J.C. and W.G. wrote the manuscript with inputs from all other authors.

Conflicts of Interest: All authors declare no potential conflicts of interest.

Competing Interests Statement

The authors declare no competing interests.

tumors and metastases. These results uncovered MLL3 mutation as a key driver of hybrid EMT and metastasis in breast cancer that could be targeted therapeutically.

Keywords

MLL3; metastasis; hybrid EMT; breast cancer; epigenetic vulnerability

Introduction

Metastasis is the leading cause of cancer death [1–3]. The ability to undergo reversible cellular and phenotypic changes is crucial for disseminating cancer cells to adapt to the changing microenvironments and stresses during metastasis. One prominent form of cellular plasticity is the epithelial-mesenchymal transition (EMT) and its reverse process mesenchymal-epithelial transition (MET) [4, 5]. While the EMT facilitates cancer cells to invade, intravasate, and survive in circulation, cancer cells need to undergo MET to colonize efficiently at distant organs [6, 7, 8, 9]. Cancer cells with a hybrid epithelial and mesenchymal (hybrid EMT) property can efficiently seed metastatic colonies and subsequently outgrow [10–13].

Initially proposed by mathematical modeling [14], hybrid EMT cells were shown to enrich tumor-initiating cells compared to fully epithelial or mesenchymal cancer cells [15, 16]. Distinct hybrid EMT states were also identified in mouse and human tumor models and shown to be more metastatic than other cells [17–19]. Despite the importance of hybrid EMT states in cancer, how cells acquire and maintain such states remains poorly understood. Thus far, EMT and MET regulators are only known to drive epithelial or mesenchymal phenotypic changes unidirectionally. It remains unclear whether a single driver can promote EMT and MET at different stages during metastasis.

The histone methyltransferase *MLL3* (also known as *KMT2C*) is frequently mutated in human cancer [20–25]. Most *MLL3* mutations in breast cancer are gene deletions or protein-truncating point mutations that are likely loss-of-function. Functional studies have shown that *MLL3* loss accelerated tumor initiation and growth in mouse tumor models and human cancer cell lines [26–29]. However, the function of *MLL3* mutations in tumor metastasis remains poorly understood. In the current study, we uncovered a role of *MLL3* loss in promoting the acquisition of hybrid EMT and metastasis.

Results

MLL3 loss enhances breast tumor metastasis

MLL3 is frequently mutated in a wide range of human cancer (Extended Data Fig. 1a) [20–22, 30]. In breast cancer, *MLL3* is altered by both point mutations and gene deletions in major subtypes (Extended Data Fig. 1b). Breast cancers with *MLL3* alterations have a significantly poorer prognosis than *MLL3* wild-type (WT) breast cancers (Fig. 1a). Furthermore, *MLL3* expression was significantly downregulated in breast tumors compared to normal mammary tissues (Extended Data Fig. 1c) [22, 31]. Interestingly, *MLL3* expression

was further reduced in distant metastases compared to primary tumors, suggesting a potential role of *MLL3* loss in metastasis (Fig. 1b–c, Extended Data Fig. 1d) [32].

To test this, we deleted *MLL3* with lentiviral CRISPR in luciferase2-tdTomato (Luc2-Tomato)-labeled MDA-MB-231 cells, which expresses WT *MLL3* (Extended Data Fig. 2a). These *MLL3*-mutant MDA-MB-231 cells developed significantly larger primary tumors than *MLL3*-WT cells (transduced with a non-targeting CRISPR vector) after orthotopic injection (Fig. 1D–E and Extended Data Fig. 2b). Furthermore, the *MLL3*-mutant tumors generated much higher metastatic burdens in the lung (11-fold) and the liver (10-fold) than *MLL3*-WT tumors, even when normalized by primary tumor burden (Fig. 1f–g and Extended Data Fig. 2c–d). The MLL3-WT lung metastases expressed significantly lower levels of MLL3 than the primary tumors, consistent with the result that loss of MLL3 enhanced metastasis (Extended Data Fig. 2e–f). *MLL3* deletion in MDA-MB-435s cells also increased spontaneous metastasis to the lung (6-fold) and liver (3-fold), even though it did not affect primary tumor growth (Fig. 1h–k and Extended Data Fig. 2g–j). A marked increase of spontaneous metastasis by *MLL3* deletion was observed in another TNBC cell line SUM159 (Fig. 1l–o and Extended Data Fig. 2k–l). Together, these results showed that *MLL3* loss promotes spontaneous metastasis in multiple cancer cell models.

Loss of *MLL3* promotes metastatic colonization

Circulating tumor cell (CTC) frequency reflects the capacity of tumor cells to invade, intravasate, and survive in circulation. The numbers of CTCs were comparable between mice bearing *MLL3*-WT and -mutant MDA-MB-231 tumors, suggesting that *MLL3* deletion did not affect these early steps of metastasis by MDA-MB-231 cells (Fig. 2a). By examining the frequency of Tomato$^+$ cancer cells extravasating out of FITC lectin-labeled blood vessels after tail vein injection, we found that *MLL3* deletion also did not affect the rate of extravasation (Fig. 2b–c).

We then measured the ability of cancer cells to colonize the lung. One week after tail vein injection, the number of cancer cells and clone sizes were similar between the *MLL3*-WT and -mutant groups, suggesting *MLL3*-deletion did not affect the initial seeding/survival of extravasated cells (Fig. 2d and Extended Data Fig. 3a). However, *MLL3*-mutant metastatic nodules grew significantly larger and generated a greater metastatic burden than the WT control at later time points (Fig. 2d–e and Extended Data Fig. 3b). Consistently, we found that *MLL3*-mutant metastases have higher frequencies of Ki67$^+$ cells than WT metastases (Fig. 2f and Extended Data Fig. 3c), while no difference was observed for cleaved caspase 3$^+$ cells (Extended Data Fig. 3d). *MLL3* loss also greatly increased metastatic colonization of SUM159 cells (Fig. 2g and Extended Data Fig. 3e).

To assess lung colonization in immunocompetent hosts, we generated mouse mammary stem cell (MaSC) organoids carrying *Trp53* CRISPR deletion and *Pik3ca^{H1047R}* mutation and then further deleted *Mll3* with lentiviral CRISPR to create *Mll3*-mutant or -WT *Trp53*^{null}/*Pik3ca*^{H1047R} MaSCs. After tail vein injection in syngeneic FVB/n mice, two independent lines of *Mll3*-mutant MaSCs generated a dramatically higher metastatic burden than *Mll3*-WT cells (Fig. 2h–j and Extended Data Fig. 3f). Together, these results suggested that *MLL3* loss increased metastatic colonization.

MLL3 loss facilitates hybrid EMT in mesenchymal cells

While MDA-MB-231 primary tumor cells had a vimentin $(VIM)^+$/E-cadherin $(Ecad)^-$ mesenchymal phenotype, a fraction of cancer cells in spontaneous lung metastases showed a hybrid EMT phenotype co-expressing VIM and Ecad (Fig. 3a). Similar $VIM^+/Ecad^+$ cells were observed in MDA-MB-231 experimental metastases, indicating that MDA-MB-231 cells undergo partial MET to acquire hybrid EMT features at distant metastatic sites (Fig. 3a). We then used the CD44 and CD104 cell-surface markers to measure the acquisition of hybrid EMT, which can be identified as $CD44^+CD104^{high}$, whereas fully mesenchymal cells are $CD44^+CD104^{low}$ [15, 16]. The *MLL3*-mutant MDA-MB-231 lung metastases contained ~2-fold more $CD44^+CD104^{high}$ cells than the control metastases two weeks after tail vein injection (Fig. 3b–c). *Mll3* deletion also increased the frequency of $VIM^+/Ecad^+$ hybrid EMT cells in *Trp53*^{null}/*Pik3ca*^{H1047R} MaSC organoids and lung metastases (Fig. 3d–e).

The MET occurring in distant metastases is likely driven by stromal signals in the metastatic niche, although such signals are yet to be clearly defined in the lung [33]. One signaling pathway that promotes MET is protein kinase A (PKA) [34]. Therefore, we tested whether *MLL3* loss can facilitate PKA-induced MET. PKA activator forskolin treatment generated significantly more (3-fold) $CD44^+CD104^{high}$ hybrid EMT cells in *MLL3*-mutant cells than the WT control (Fig. 3f). Interestingly, forskolin upregulated the epithelial marker E-cadherin only in *MLL3*-mutant cells (Fig. 3g). Together, these results showed that loss of *MLL3* facilitated the acquisition of the hybrid EMT state.

We next assessed whether the hybrid EMT cells increased by *MLL3* deletion have higher metastatic capacity. $CD44^+/CD104^{high}$ hybrid EMT cells and $CD44^+/CD104^{low}$ mesenchymal cells were sorted from forskolin-treated WT or *MLL3*-mutant MDA-MB-231 cells and assessed by experimental metastasis. The *MLL3*-mutant hybrid EMT cells showed markedly higher lung metastasis-initiating ability than the other cell types (Fig. 4a–c). The *MLL3*-mutant hybrid cells also have significantly increased metastasis-initiating ability to multiple other organs, including the bone, liver, and kidney (Fig. 4d–e and Extended Data Fig. 4a–d). Furthermore, *MLL3*-mutant hybrid cells showed significantly higher tumor-initiating ability and faster tumor growth than the other cell types (Fig. 4f, Extended Data Fig. 4e and Supplementary Table 1). These results suggest that the induction of hybrid EMT by *MLL3* loss increases multi-organ metastatic and tumor-initiating abilities.

MLL3 loss facilitates EMT and metastasis in luminal cancer

To determine whether *MLL3* loss promotes epithelial-mesenchymal plasticity in epithelial cells, we deleted *MLL3* in the luminal breast cancer cell line MCF7 by lentiviral CRISPR. While *MLL3* deletion did not lead to a noticeable EMT phenotype in adherent culture, it significantly increased the sphere-forming ability, a property associated with cells undergoing an EMT (Fig. 5a) [35]. *MLL3*-mutant MCF7 spheres contained substantially higher frequencies of cells lacking the epithelial marker CD24 (Fig. 5b). These $CD24^-$ cells had a higher sphere-forming ability than $CD24^+$ cells (Extended Data Fig. 5a). Furthermore, the *MLL3*-mutant spheres markedly upregulated vimentin and EMT transcription factors SLUG and SNAIL while maintaining the expression of the epithelial markers E-cadherin and CD104, suggesting these cells underwent a partial EMT (Fig. 5c and Extended Data

Fig. 5b). *MLL3* deletion in another luminal breast cancer cell line T47D similarly increased sphere formation and the frequency of $CD24^-$ and VIM^+ cells (Fig. 5d–e and Extended Data Fig. 5c).

We further tested whether *MLL3* loss potentiates the EMT in response to TGF-β in HMLE immortalized human mammary epithelial cells, a well-characterized model for TGF-β-induced EMT [35]. *MLL3*-mutant HMLE cells underwent EMT more efficiently than *MLL3*-WT cells responding to low concentrations of TGF-β1 and generated mesenchymal cells at faster kinetics, as measured by the induction of $CD44^{high}CD24^{low}$ mesenchymal cells (Fig. 5f–g and Extended data Fig. 5d).

In vivo, MLL3 deletion significantly accelerated MCF7 tumor onset and growth (Fig. 5h–i). Furthermore, the *MLL3*-mutant tumors exhibited decreased E-cadherin expression and disorganized adherent junctions, consistent with the partial EMT phenotype in *MLL3*-mutant MCF7 spheres (Fig. 5j). Furthermore, *MLL3*-mutant MCF7 cells generated significantly more metastases than the WT cells after tail vein injection (Fig. 5k–l). These results showed that *MLL3* loss enabled luminal cancer cells to acquire the hybrid EMT property and become more metastatic.

MLL3 loss activates the interferon-γ pathway

To understand the mechanisms by which *MLL3* deletion promotes hybrid EMT, we compared transcriptomes of *MLL3*-WT and -mutant MCF7 cells. Several inflammatory signaling pathways, including interferon-γ (IFNγ), interferon-α (IFNα), and tumor necrosis factor-α (TNFα) signaling, were among the top gene signatures enriched by *MLL3* deletion (Fig. 6a and Extended Data Fig. 6a–b). *MLL3* deletion also increased the IFNγ gene signature in MDA-MB-231 cells (Fig. 6b). Furthermore, the hybrid EMT cells sorted from forskolin-treated MDA-MB-231 cells expressed a significantly higher level of IFNγ gene signature than the mesenchymal cells (Fig. 6c and Extended Data Fig. 6c). Ingenuity Pathway Analysis (IPA) identified IFNγ as the upstream cytokine regulator in the hybrid EMT cells (Fig. 6d). Several interferon regulatory factors (IRF1, IRF2, IRF3, IRF5, and IRF7) and STAT1 were also identified as active transcription factor regulators in the hybrid EMT cells (Fig. 6e). These results suggest that the hybrid EMT state is associated with higher IFNγ pathway activity, which is upregulated by *MLL3* loss.

IFNγ signaling is required for hybrid EMT and metastasis

We tested the functional role of inflammatory cytokines and found that IFNγ treatment, but not IFNα or TNFα, robustly increased the frequency of $CD24^-$ cells in MCF7 adherent culture, suggesting a unique role of IFNγ in hybrid EMT induction (Extended Data Fig. 6d). Furthermore, *MLL3*-mutant MCF7 cells downregulated the epithelial marker CD24 more efficiently than WT cells upon IFNγ treatment, supporting that *MLL3* loss sensitizes the response to IFNγ (Extended Data Fig. 6e). Notably, IFNγ treatment in control MCF7 spheres increased the frequency of $CD24^-$ cells to that of *MLL3*-mutant spheres (Fig. 6f). Conversely, inhibition of IFNγ signaling using an *IFNGR1* blocking antibody decreased the frequency of $CD24^-$ cells in *MLL3*-mutant spheres (Fig. 6g). Lentiviral CRISPR deletion of *IFNGR1* similarly reduced the frequency of $CD24^-$ cells and sphere-forming efficiency in

MLL3-mutant cells (Fig. 6h–i and Extended Data Fig. 6f). *IFNGR1* deletion also abolished the effect of *MLL3* loss on accelerating MCF7 tumor onset (Fig. 6j). Additionally, *IFNGR1* deletion diminished EMT phenotypes in *MLL3*-mutant MCF7 tumors, as measured by EMT markers and tumor histology (Fig. 6k and Extended Data Fig. 6g–i).

We then assessed the function of IFNγ signaling in hybrid EMT induction and metastasis in mesenchymal cells. Treating MDA-MB-231 cells with IFNγ increased the frequency of $CD44^{+}CD104^{high}$ hybrid EMT cells (Fig. 6l). Conversely, *IFNGR1* CRISPR deletion in *MLL3*-mutant MDA-MB-231 cells reduced forskolin-induced hybrid EMT cells (Extended Data Fig. 7a). When *MLL3*-WT MDA-MB-231 cells were pretreated with IFNγ (1 or 10 ng/ml) for four days *in vitro*, they acquired markedly increased metastatic ability (Fig. 6m). Conversely, *IFNGR1* deletion in *MLL3*-mutant MDA-MB-231 cells reduced lung colonization to the level of MLL3-WT cells (Fig. 6n). Furthermore, the percentages of hybrid EMT cells among disseminated tumor cells in the lung were significantly reduced by *IFNGR1* deletion (Fig. 6o–q and Extended Data Fig. 7b–d). *IFNGR1* deletion also dramatically decreased primary tumor growth and spontaneous metastasis by *MLL3*-mutant MDA-MB-231 cells (Fig. 6r–s and Extended Data Fig 7e–f). Together, these results indicate that IFNγ signaling is required for the induction of hybrid EMT and efficient metastatic colonization, acting downstream of *MLL3* loss.

IFγ treatment significantly increased the expression of EMT transcription factors (TFs) SLUG, SNAIL, and ZEB1 in MCF7 and MDA-MB-231 cells, recapitulating the effect of *MLL3* deletion (Extended Data Fig. 7g–h). Conversely, *IFNGR1* deletion in *MLL3*-mutant cells reduced the levels of these EMT TFs in both cell lines. Concordant with the role of IFNγ in EMT TF upregulation, we observed a significant correlation between the IFNγ response gene signature and the EMT TF signature in the TCGA BRCA cancer samples (Extended Data Fig. 7i). High IFNGR1 levels are associated with poorer survival in TNBC, although with moderately improved survival in ER^{+} cancer[36]. While the underlying causes of these distinct associations remain unknown, this data supports the role of IFNγ in promoting aggressive tumor phenotypes.

MLL3 loss increases H3K27ac in IFNγ response gene enhancers

To understand how *MLL3* loss may enhance IFNγ signaling at the chromatin level, we performed ATAC-seq and ChIP-seq for enhancer and promoter histone marks comparing *MLL3*-WT and *MLL3*-mutant MDA-MB-231 cells. HOMER analysis[37] of ATAC-seq peaks showed that the motifs of IFNγ downstream TFs were enriched in the more accessible chromatin regions in *MLL3*-mutant cells, while the global chromatin landscape did not change significantly (Fig. 7a and Extended Data Fig. 8a). ChIP-seq for H3K4me1, H3K27ac and H3K4me3 showed that *MLL3* loss did not alter the abundance and global distribution of these histone marks at the genome-wide scale (Extended Data Fig 8a–b), agreeing with previous studies in MCF7 cells [38, 39].

Next, we focused on the effect of *MLL3* loss on enhancers of IFNγ pathway genes. We analyzed the H3K27ac peaks by ROSE [40, 41] and found that *MLL3* loss increased H3K27ac intensity in the enhancers of a subset of IFNγ response genes in both MCF7 and MDA-MB-231 cells (Fig. 7b–d). In contrast, H3K4me1 levels at these H3K27ac-up regions were

not changed in MDA-MB-231 cells while slightly increased in MCF7 cells upon *MLL3* loss (Extended Data Fig. 9a–b). Of note, *MLL3* knockdown in MCF7 cells increased H3K27ac levels and established a super-enhancer at the *IRF1* gene, a key mediator of IFNγ signaling (Fig. 7e–f). In MDA-MB-231 cells, *MLL3* deletion markedly increased H3K27ac levels at the *IRF1* enhancer upon IFNγ stimulation, although it did not affect the H3K27ac levels in the unstimulated state (Extended Data Fig. 9c–e). Interestingly, *MLL3* deletion significantly increased the overall enrichment of H3K27ac in enhancers in both MDA-MB-231 and MCF7 cells (Extended Data Fig. 9f).

We next examined whether the MLL3 complex binds to the IFNγ response gene loci. We performed ChIP-seq using an antibody recognizing both MLL3 and MLL4 and another antibody against UTX, a subunit shared by MLL3 and MLL4 complexes. Both MLL3/4 and UTX bind to enhancers of IFNγ response genes, including *IRF1* and *STAT1*. Furthermore, the intensity of MLL3/4 and UTX peaks was significantly reduced after *MLL3* deletion, supporting direct occupancy of the MLL3 complex in these regions (Extended Data Fig 9a–b, g).

STAT1 phosphorylation and the subsequent auto-upregulation are critical steps in the IFNγ signaling pathway [42, 43]. We found that H3K27ac levels at the *STAT1* enhancer/promoter regions increased significantly in *MLL3*-mutant MCF7 cells (Fig. 7e). Furthermore, phospho-STAT1 and total STAT1 expression were higher in these cells upon IFNγ treatment in both adherent and sphere cultures (Fig. 7g and Extended Data Fig. 10a–b). STAT1 occupancy at its own promoter was significantly higher in *MLL3*-mutant MCF7 cells than WT cells upon IFNγ treatment (Fig. 7h). *MLL3* deletion in MDA-MB-231 cells also led to a more robust upregulation of phospho-STAT1 and total STAT1 by IFNγ (Extended Data Fig. 10c–d). These results point to a consistent effect of *MLL3* loss on activating enhancers of a sub-cluster of IFNγ response genes, which leads to the upregulation of key mediators of IFNγ signaling, such as STAT1.

MLL3-deficient tumor cells are sensitive to BET inhibition

To identify molecular vulnerabilities for *MLL3*-mutant cells, we performed a chemical screen comparing the response of syngeneic WT and *Mll3*-mutant MaSCs to a collection of epigenetic compounds. We found that the BET (bromodomain and extra-terminal) protein inhibitor (+)-JQ1 inhibited *Mll3*-mutant cells much more efficiently than WT cells (Fig. 8a). Furthermore, *Mll3*-mutant organoid cells treated with JQ1 had a greatly diminished ability to re-seed secondary organoids in the absence of further JQ1 treatment, suggesting inhibition of the self-renewal ability (Fig. 8b). Using primary cultures of *Mll3*-WT and *Mll3*-mutant tumors induced by constitutively active *Pik3ca* (*Pik3ca**) [29], we showed that *Mll3*-mutant tumor cells are more sensitive (10-fold) to two distinct BET inhibitors, JQ1 and GSK525762 (Fig. 8c). Furthermore, while parental MCF7 cells are relatively resistant to BET inhibition [44], *MLL3* deletion increased their sensitivity to JQ1 by 15-fold. Similar effects were found for GSK525762 (Fig. 8d).

In vivo, JQ1 significantly inhibited the growth of *Mll3*-mutant *Pik3ca** tumors but did not affect *Mll3*-WT tumors (Fig. 8e). Similar differences in JQ1 responses were observed between *MLL3*-WT and -mutant MDA-MB-231 tumors (Fig. 8f). Furthermore, JQ1

significantly inhibited the growth of *MLL3*-mutant MDA-MB-231 metastases in both the lung and bone but did not affect *MLL3*-WT metastases (Fig. 8g–h). The proliferation in *MLL3*-mutant metastatic lesions was significantly inhibited by JQ1, as shown by Ki67 immunostaining (Fig. 8i). TNBC cells, including MDA-MB-231, show increased sensitivity to BET inhibitors, albeit resistance frequently occurs [45–49]. Our results show that *MLL3* loss sensitizes tumor cells to BET inhibition, suggesting a potential strategy for targeting *MLL3*-mutant cancers, including metastatic diseases.

Discussion

Carcinoma cells with hybrid epithelial and mesenchymal features exhibit remarkable plasticity, stemness, and metastasis-initiating ability [10–13]. Here, we discovered that *MLL3* loss is an important epigenetic mechanism promoting the hybrid EMT state. While previous work suggests *MLL3* loss promotes EMT in gastric cancer and metastasis in lung cancer [50, 51], how MLL3 mechanistically controls EMT and metastasis remains poorly understood. We found that, distinct from other known EMT regulators that drive uni-directional epithelial-mesenchymal transitions, *MLL3* deletion enhanced responses to signals that induce either EMT or MET, which allow tumor cells to acquire hybrid EMT states, leading to efficient metastatic colonization. This function of MLL3 is distinct from that of the related MLL protein KMT2D, whose loss causes a fully mesenchymal phenotype and, consequently, fails to increase metastatic potential [52].

We found a critical role of IFNγ signaling in the induction of hybrid EMT, acting downstream of *MLL3* loss. IFNγ signaling has a complex role in cancer. Although IFNγ signaling is critical for anti-tumor immunity and can induce growth arrest and cell death [53, 54], IFNγ signaling can also lead to resistance to immune checkpoint blockade and increase the infiltration of immunosuppressive cells to promote metastasis [36, 54]. Our study identified a cell-intrinsic function of IFNγ in EMT and metastasis. Interestingly, normal and cancerous mammary stem cells can evade the cytostatic and cytotoxic effects of IFN [55]. Likely, hybrid EMT cells can similarly circumvent the inhibitory effect of IFNγ and utilize it to sustain the hybrid EMT state.

MLL3 loss increased H3K27ac levels in the enhancers of a subset of IFNγ genes. This effect is unexpected, as *MLL3* catalyzes the formation of H3K4me1, which primes enhancers for H3K27ac modification [56]. Interestingly, increased H3K27ac at the enhancers of IFNγ pathway genes were not correlated with H3K4me1 changes. This suggests that *MLL3* may act through a methyltransferase-independent mechanism to regulate IFNγ genes. Consistent with this notion, a study in mouse ES cells showed that the adaptor function of *MLL3* plays a more prominent role in transcriptional regulation than the enzymatic function [57]. An alternative scenario is that MLL3 loss downregulates a transcription repressor for IFNγ response genes, thereby increasing their enhancer activation manifested as increased H2K27ac levels. Future studies are required to investigate these possibilities.

Our results also indicate the importance of targeting the hybrid EMT state rather than simply inhibiting EMT in cancer therapy. We found that inducing MET in *MLL3*-mutant MDA-MB-231 cells with the PKA activator forskolin generated hybrid EMT cells that are

highly metastatic. Although sustained activation of PKA can eventually convert the cells to a fully epithelial state inhibiting metastasis [58], EMT-inhibiting strategies may inadvertently enhance metastasis in highly plastic cancer cells by generating hybrid EMT cells. Thus, identifying the vulnerability of hybrid EMT cells will be particularly useful for blocking metastasis. We found that *MLL3*-deficient breast cancer cells were uniquely sensitive to BET inhibition. Interestingly, *MLL3*-knockdown AML cells are also susceptible to the BET inhibitor JQ1 [27]. Together, our data show that BET inhibition can be an effective approach for targeting the highly plastic *MLL3*-mutant cancer cells in various cancer types.

Methods

All experiments performed in this study were approved by the Institutional Animal Care and Use Committee (IACUC) and Institutional Biosafety Committee of Albert Einstein College of Medicine.

Cell lines and adherent cell culture

MCF7, MDA-MB-231, SUM159, T47D and MDA-MB-435s cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Corning) with 10% fetal bovine serum (FBS, HyClone) and 1% (v/v) antibiotic-antimycotic solution (HyClone). HMLE cells were maintained in 1:1 mix of MEGM (Lonza) and DMEM/F-12 (Corning) supplemented with 5 ng/ml epidermal growth factor (EGF, Sigma-Aldrich E9644), 2.5 μg/ml insulin (Sigma-Aldrich, 91077C), 0.25 μg/ml hydrocortisone (Sigma-Aldrich, H0888). For TGFβ1 treatment, HMLE cells were cultured in MEGM and DMEM/F12 media (1:1) supplemented with insulin, EGF, hydrocortisone, 5% FBS and different concentrations of TGFβ1. The medium was refreshed every 3 days.

All cell lines, except HMLE, were authenticated by short tandem repeat (STR) analysis, matching to the ATCC database. HMLE cells, which lack preexisting STR data, were authenticated by morphology and flow cytometric profile. Of cell lines used in this work, only MDA-MB-435s is in the list of commonly misidentified cell lines maintained by ICLAC. MDA-MB-435s was obtained from ATCC and authenticated by STR on November 21, 2017 with 100% match to ATCC database (Extended Data Fig. 2m). We used MDA-MB-435s at the beginning of this study as a cell line that can form spontaneous metastasis with reasonable efficiency. One report suggests that MDA-MB-435s is a breast cancer cell line that expresses melanoma genes due to lineage infidelity [59]. However, its exact origin is still controversial. Therefore, after showing the effect of MLL3 deletion in promoting metastasis in MDA-MB-435s, we used other bona fide breast cancer cell lines to confirm our findings and dissect the mechanisms.

Tumor sphere culture

Tumor sphere culture was performed as described in previous studies [60, 61] with slight modification. MCF7 were grown in DMEM:F12 supplemented with 2% (v/v) B27 (50x, Gibco, 17504044), 20 ng/ml EGF (Sigma-Aldrich E9644), 20 ng/ml basic fibroblast growth factor (bFGF, EMD Millipore, GF003.), 4 μg/ml heparin (Sigma, H4784), 10 μg/ml insulin (Sigma-Aldrich, 91077C), 0.5 μg/ml hydrocortisone (Sigma-Aldrich, H0888). 1,000 cells/

well were plated in ultra-low attachment 96-well plates and cultured for 7 days. The spheres greater than 100 μm in diameter were counted.

Mammary Organoid Culture

Wild-type and *Mll3*-null MaSC organoids were generated and cultured as described in [29]. 500 cells/well were seeded in ultra-low attachment 96-well plates (Corning Life Sciences) with organoid medium, containing Advance DMEM/F12 (Life Technologies) supplemented with 1% L-glutamine, 5% heat-inactivated FBS (Sigma-Aldrich, F2442), 5% Matrigel (Corning Life Sciences, 354234), 10 ng/ml EGF (Sigma-Aldrich, E9644), 20 ng/ml bFGF (EMD Millipore, GF003), 4 μg/ml heparin (Sigma-Aldrich, H4784) and 5 μM Y-27632 (Cayman Chemical Company, 10005583). After 7-day culture, organoids greater than 100 μm in diameter were counted.

Lentiviral CRISPR-mediated gene knockout

The sgRNAs targeting *MLL3* were designed using the CRISPR Design Tool from the Zhang Laboratory (http://crispr.mit.edu) and cloned into the lentiCRISPRv2 vector (Addgene #52961, a gift from Feng Zhang). The sgRNA targeting sequences are as follows: sg*MLL3* #3-GGGGAAGCGCCATCTTTGCG, sg*MLL3* #9-GTCGGAGGAGGACAAGAGCG, sgNT-GCGAGGTATTCGGCTCCGCG. Cells were transduced with the lentiviral vectors and selected by puromycin. For *IFNGR1* targeting, the sgRNA sequence targeting *IFNGR1* or sgNT was cloned into the lentiCRISPRv2-GFP vector (a gift from David Feldser, Addgene #82416). Cells were transduced and sorted based on GFP. The sg*IFNGR1* sequence is GTATAAAATACTCACGCAGA.

Animals

NOD/SCID, NSG, C57BL/6 and FVB/n mice were imported from the Jackson Laboratory (Bar Harbor, ME) and bred in-house. Mice were housed in microisolator cages and a pathogen-free condition. Mice were on a 12-12 hour light-dark cycle, with water and food ad libitum. Temperature in the facility was maintained at 21 °C. 4-8 weeks female mice were used in the study. All animal experiments were performed according to protocols approved by the IACUC of Albert Einstein College of Medicine (Protocol 20170607 and 00001266). For primary tumors, the maximal tumor burden approved by our IACUC is 2 cm in diameter and all tumors in this study did not exceed the limit.

Metastasis assays

For spontaneous metastasis, female NOD/SID mice at 6-8 weeks of age were injected with 1×10^6 cells suspended in 50 μl 75% PBS/25% Matrigel into the left inguinal mammary fat pad. Tumor growth was monitored weekly by caliper measurements ($V = (4/3) \times \pi \times (L/2) \times (L/2) \times (D/2)$) and metastasis burden was monitored longitudinally by bioluminescence imaging with an IVIS Spectrum system (PerkinElmer, Living image version: 4.3.1.0.15880). At the endpoint, organs were dissected post-mortem and imaged for bioluminescence *ex vivo*. For experimental metastasis, 3-5 x 10^5 cells were injected via tail vein into 4-6 weeks old female NSG or FVB/n mice. Metastatic burdens were monitored by bioluminescence imaging, flow cytometry or tissue histology. For experimental metastasis of FACS sorted

$CD44^{+}CD104^{high}$ and $CD44^{+}CD104^{-/low}$ MDA-MB-231 cells, 2×10^4 sorted cells were injected per mouse.

Orthotopic tumor cell implantation

Female NSG or NOD/SID mice at 6-8 weeks of age were injected with 1×10^6 tumor cells suspended in 50 μl of 75% PBS/25% Matrigel (Corning Life Sciences, 354234) into the inguinal mammary fat pad. For MCF7 experiments, mice received 200 nM 17β-estradiol (Sigma-Aldrich, E8875) in the drinking water as previously published [62]. Tumor onset was monitored by weekly palpation. Tumor volume was measured by caliper and calculated as $V = (4/3) \times \pi \times (L/2) \times (L/2) \times (D/2)$. For *in vivo* limiting dilution assay, 20,000, 4,000, and 400 of FACS sorted $CD44^{+}CD104^{high}$ and $CD44^{+}CD104^{-/low}$ sgNT or sgMLL3 MDA-MB-231 cells were injected into the third and fourth mammary fat pad of NOD/SCID mice. Tumor onset and progression were monitored by weekly palpation for 80 days.

Circulating tumor cell and extravasation analyses

100 μl blood samples from tumor-bearing mice were collected via retro-orbital venous plexus. After removal of red blood cells using lysis buffer, tdTomato+ tumor cells were quantified by flow cytometer. For extravasation, blood vessels were visualized via tail vein injection of FITC-labeled Ricinus communis agglutinin I (RCA I, Vector Laboratories, FL-1081) 15 min before animal sacrifice. Frozen tissue sections were processed for imaging tumor cells and blood vessels. Percentages of tumor cells outside of blood vessels were scored in at least ten random fields in each mouse.

Western blot

Cells were lysed with the RIPA buffer containing 25 mM Tris-HCL, pH 7.5, 150 mM NaCl, 1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA and 1.0% NP-40 on ice. Samples were resolved by SDS-PAGE and transferred to PVDF membranes. The western blots were performed with antibodies against MLL3 (Polyclonal, Millipore Sigma, ABE1851, 1:1000), E-cadherin (clone 36, BD Biosciences, cat# 610182, 1:1000), Vimentin (clone D21H3, Cell Signaling Technology, cat# 5741S, 1:1000), Slug (clone C19G7, Cell Signaling Technology, cat#9585S, 1:1000), Snail (clone C15d3, Cell Signaling Technology, cat#3879S, 1:1000), Zeb1 (clone D80D3, Cell Signaling Technology, cat#3396S, 1:1000), Stat1 (clone D1K9Y, Cell Signaling Technology, cat#14994S, 1:1000), phosphor-Stat1(Tyr701) (clone 58D6, Cell Signaling Technology, cat#9167S,1:1000), β-actin (clone C4/actin, BD Biosciences, cat#612656, 1:10,000), H3K27ac (clone D5E4, Cell Signaling Technology, cat#8173S, 1:1000), H3K4me1 (clone D1A9, Cell Signaling Technology, cat#5326S, 1:1000), H3K4me3 (clone C42D8, Cell Signaling Technology, cat#9751S, 1:1000), or Histone H3 (clone 1B1B2, Cell Signaling Technology, cat#14269S, 1:10,000). Original scans of the western blot results are provided in the Source Data for individual figures.

Lung metastatic tumor cell isolation

Lungs were perfused with PBS through right ventricle, dissected and minced. The tissues were then digested with 300 units/ml collagenase type 3 (Worthington Biochemical, LS004182), 100 units/ml hyaluronidase (Worthington Biochemical, LS002592), 5 μM

Y-27632 (Cayman Chemical Company, 10005583-10) and 10 μg/ml DNase I (Roche, 104159) in the DMEM/F-12 medium at 37°C for 1 hours on an orbital shaker. After primary digestion, the cell pellet was washed with 1x PBS. RBCs were lysed by RBC lysis buffer (eBioscience, 00-4300-54). Cells was further digested with 0.05% trypsin-EDTA for 5 minutes and 1 unit/ml neutral protease (dispase) (Worthington Biochemical, LS02109) plus 100 μg/ml DNase I (Roche, 104159) for 5 minutes. The digested cells were then filtered through 40 μm cell strainers to obtain single cells.

Flow cytometry and sorting

Multiparametric flow cytometric analyses were performed on an LSRII equipped with FACS Diva 6.1 software (BD Biosciences) and analyzed with FlowJo software (FlowJo, LLC). Dead cells were excluded by FSC, SSC and 49,6-diamino-2-phenylindole (DAPI, Sigma) staining. Cell sorting was performed using a MoFlow Astrios Cell Sorter (Beckman Coulter). Fluorophore-conjugated anti-CD44 (clone G44-26, cat#559942) and anti-CD24 (clone ML5, cat#555428) antibodies were obtained from BD Bioscience. Biotinylated-CD104 antibody (ITGB4, clone 439-9B, cat#13-1049-82) was purchased from eBioscience. Anti-human IFNGR1 antibody (Clone: GIT208, cat#I443) was purchased from Leinco Technologies. Unless otherwise specified, all antibodies were used at a 1:100 dilution. FACS gating strategies are provided in Supplementary Figure.

Immunofluorescence and immunohistochemistry

Formalin-fixed and paraffin-embedded tissue sections or OCT-embedded frozen sections were stained with standard protocols. The following primary antibodies were used: E-cadherin (clone 36, BD Biosciences, cat#610182, 1:200), Vimentin (clone D21H3, Cell Signaling Technology, cat#5741S, 1:200), CD44 (clone G44-26, BD Bioscience, cat#559942, 1:50), CD104 antibody (clone 439-9B, eBioscience, cat#14-1049-82, 1:200), Ki67 (Polyclonal, EMD Millipore, cat#AB9260, 1:200), cleaved caspase 3 (clone 5A1E, Cell Signaling Technology, cat#9664S, 1:200), anti-human mitochondria antibody (clone 113-1, EMD Millipore, cat# MAB1273, 1:200), Snail (clone C15d3, Cell Signaling Technology, cat#3879S, 1:200), Slug (clone C19G7, Cell Signaling Technology, cat#9585S, 1:200), Keratin 8 (Developmental Studies Hybridoma Bank, cat#AB_531826, 1:60), MLL3 (Polyclonal, Millipore Sigma, ABE1851, 1:100), α-smooth muscle actin (clone 1A4, Biolegend, cat#904601, 1:300). The ImmPRESS HRP goat anti-mouse IgG Polymer Reagent (Vector Laboratories, MP-7402) and the DAB/HRP detection IHC kit (Vector Laboratories, SK-4100) were used for immunohistochemistry. Fluorescent slides were imaged with an AXIO Examiner D1 microscope (Zeiss) with a confocal scanner unit (Yakagawa). SlideBook software 6.0 was used for the acquisition of images. Immunohistochemistry slides were scanned with a P250 High Capacity Slide Scanner (3DHISTECH).

Chemical screening and Cell viability assay

MaSCs were screened with an epigenetic compound library (Cayman #11076). 1,000 cells were seeded in 96-well plate. After 24 h, cells were treated with 0.8 μM compounds or 0.1% DMSO for 3 days. For treatment of organoid cultures, 1,000 cells were seeded in ultra-low attachment 96-well plate and the test compounds were applied at the same

time. Cell viability was determined by the cell counting kit-8 (CCK-8, Dojindo, CK04-13) assay or alamar blue assay (Invitrogen, DAL1025) after 72 h treatment according to the manufacturer's instructions. The optical density (OD) was measured by synergy H4 hybrid microplate reader (Biotec Inc). The cell viability (%) was calculated as the following: $(OD_{compound} - OD_{background})/(OD_{vehicle} - OD_{background}) \times 100\%$.

In vivo JQ1 treatment

Small tumor fragments (~1 mm^3) of *Mll3*-WT/*Pik3ca** and *Mll3*-mutant/*Pik3ca** tumors were orthotopically transplanted into 6-8 weeks female C57BL/6 mice. For MDA-MB-231 primary tumor, female NOD/SID mice at 6-8 weeks of age were injected with 1×10^6 cells suspended in 50 μl 75% PBS/25% Matrigel into the inguinal mammary fat pads. When tumors reached ~250 mm^3, mice received daily i.p. injections with 50 mg/kg JQ1 or vehicle control for 14 days. JQ1 was kindly provided by Dr. James E Bradner from Dana-Farber Cancer Institute (Boston, MA, USA). Tumor volume was measured by caliper and calculated as $V = (4/3) \times \pi \times (L/2) \times (L/2) \times (D/2)$ every two days.

To assess the efficacy of JQ1 in the metastatic setting, Luc2-tomato-labeled *MLL3*-WT or -mutant MDA-MB-231 cells were injected via tail vein into 4-6 weeks old female NSG mice. When metastatic burden reached $6\text{-}8\times10^9$ photon flux as measured by bioluminescence imaging, mice received daily i.p. injections with 50 mg/kg JQ1 or vehicle control for 14 days. Metastatic burdens were monitored by bioluminescence imaging weekly.

Quantitative RT-PCR

Total RNA was isolated from cells using the Quick-RNA microprep RNA kit (Zymo Research, R1050) or Direct-zol RNA Miniprep Plus kit (Zymo Research, R2072) and cDNA synthesis was performed by the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814). Real-time PCR was performed with SYBR Green PCR Master Mix (Applied Biosystems, 4368708) with *GAPDH* as internal controls in an Applied Biosystems 7900HT real-time PCR machine with QuantStudio 6 Real-Time PCR system. Primer sequences are shown in Supplementary Table 2.

Microarray and RNA-seq gene expression analyses

Microarray gene expression profiling was done using total RNA on Affymetrix human Gene 2.0ST chips by the Einstein Genomic Core Facility according to the manufacturer's instruction. For RNA-seq, total RNA was isolated using Direct-zol RNA Miniprep Plus kit. RNA-seq libraries were prepared and sequenced by Genewiz. RNA-seq reads were aligned using STAR (version 2.4.0) to the reference human genome (GRCh38.p13) with Ensembl annotation (gencode.v33). DESeq2 (v1.30.1) R package was applied to analyze the differential expression. Upstream regulators were analyzed by IPA (QIAGEN).

Histone mark ChIP-seq

ChIP-seq was performed as previously described [63, 64]. The following antibodies were used for immunoprecipitation: anti-H3K4me1 (clone D1A9, Cell Signaling Technology, cat#5326S, 2 μg per sample), H3K4me3 (clone C42D8, Cell Signaling Technology,

cat#9751S, 2 μg per sample), and H3K27ac (Polyclonal, Active Motif, cat#39133, 2 μg per sample). Around 2 x 10^7 cells and two biological repeats per genotype were used for ChIP-seq experiment. Crosslinking was performed with 1% formaldehyde (Sigma-Aldrich, F1635) at 37°C for 10 min. A final concentration of 125 mM glycine was added for 5 minutes at room temperature to quench formaldehyde crosslinking. Cells were washed with cold PBS and harvested on ice. Cells were lysed with 1ml of SDS lysis buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH8) in the presence of protease inhibitor (Roche, cat#11836170001) and phosphatase inhibitor (Thermo, cat#78427). Cells were sonicated using the Branson Sonifier 250 (Branson Ultrasonics 101063588) with a 20% amplitude setting for 5 min 30 s with 10 s of on and off intervals. After sonication, 1 ml of ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH8, 167 mM NaCl) with proteinase and phosphatase inhibitor was added to prevent SDS precipitation. After centrifugation at 3,000 x g for 10 min at 4°C, 2 ml of supernatant were transferred to a 15 ml falcon tube and diluted with 3 ml of ChIP dilution buffer containing proteinase and phosphatase inhibitor.

50 μl of Dynabeads (Thermo Fisher Scientific 10009D), which were previously blocked with 1% BSA, were added to samples and incubated at 4°C with rotation for 1 hour. After pre-clearing, Dynabeads beads were removed. The supernatant was transferred to new 15 ml falcon tubes and 100 μl out of 5 ml volume was collected as 2% input. Antibodies were added to the pre-cleared samples for overnight incubation at 4°C with rotation. 200 μl of Dynabeads that were previously blocked with 1% BSA, were added to the sample and incubated for 6 hours at 4°C with rotation. After incubation, beads were collected and washed at 4°C with rotation following the procedure: 5 min with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH8, 150 mM NaCl), twice 5 min with high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH8, 500 mM NaCl) and twice 5 min with TE buffer (10 mM Tris-HCl pH8, 1 mM EDTA). After washing, beads were resuspended in 250 μl of elution buffer (1% SDS, 0.1 M NaHCO3) and incubated at 60°C for 30 min using a thermomixer. The supernatant was transferred to new tubes and 5 M NaCl was added for overnight de-crosslinking at 65°C. 10 μl of 0.5 M EDTA, 20 μl of 1 M Tris-HCl (pH 6.5) and 1 μl of proteinase K (20 mg/ml, NEB, P8107S) were added and incubated at 45°C for 1 hour for de-crosslinking. ChIP-DNA was isolated by using the QIAquick PCR purification kit (QIAGEN, 28104). Libraries were prepared using the NEBNext® Ultra II DNA Library Prep Kit (NEB, E7103S) and NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1; NEB, E7335S). Samples were pooled and submitted to MSKCC Integrated Genomics Operation core for sequencing. Libraries were run over one lane of a HiSeq 4000 in a 50bp/50bp paired-end run, using the HiSeq 3000/4000 SBS Kit (Illumina).

UTX and MLL4 ChIP-seq

ChIP was performed as described previously with some modifications[65]. Briefly, around 2 x 10^7 *MLL3*-WT or -mutant MCF7 and MDA-MB-231 cells were crosslinked with 1% formaldehyde for 10min, quenched by 125 mM glycine for 10 min and rinsed twice with 1X PBS. Fixed cells were incubated in the lysis buffer containing 5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40 and protease inhibitors on ice for 10 min and centrifuged at 850 x g

for 5 min at 4°C. Nuclear pellets were resuspended with cold TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), and sonicated into 200-500 bp fragments at 4°C. The experimental chromatins mixed with 20 ng drosophila spike-in chromatin (Active Motif, cat#53083) were incubated overnight at 4°C with Dynabeads Protein A (Invitrogen, 10002D), which had been pre-coated with 1 μg Drosophila-specific histone variant H2Av spike-in antibody (Active Motif, cat#61686) and 8 μg primary target antibody. Anti-UTX (UTX #2 antibody) and anti-MLL4 (MLL4 #3 antibody) were both home-made, which were described previously [66, 67]. The beads were washed twice with RIPA buffer, twice with RIPA buffer containing 300 mM NaCl, twice with LiCI buffer (50 mM Tris-Cl pH 7.5/250 mM LiCl/0.5% NP-40/0.5% Na-DOC) and twice with PBS. DNA was then eluted with 100 μl elution buffer (0.1 M $NaHCO_3$, 1%SDS and 20 μg proteinase K) and reverse crosslinked at 65°C overnight. DNA was purified using QIAquick PCR purification kit (Qiagen) and quantified. The ChIP DNA and input DNA were used to construct libraries using NEBNext Ultra™ II DNA Library Prep kit for Illumina (NEB, E7645) following the manufacturer's instructions. The final libraries were sequenced on Illumina NovaSeq 6000.

ChIP-seq data analysis

The raw reads were assessed for quality using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), then trimmed using Trimmomatic (v0.40) [68] with Illumina adaptor file and default values. The trimmed paired-end reads were then aligned to the human reference genome (GRCh38) with bwa-mem (arXiv:1303.3997v2), then sorted and indexed with samtools (v1.15.1) [69]. The resulted alignment BAM files were then processed through PICARD (https://github.com/broadinstitute/picard) tools for 'MarkDuplicates'. The marked alignments were then used for DeepTools2 for clustering and heatmaps [70]. Firstly, BigWig files were generated from the Markduplicated bam files using bamCoverage with RPKM normalization. The BigWig files were then passed to computeMatrix, which generates a matrix of scores of ChIP-seq read density across selected genomic regions (e.g., peaks), which was used for generating heatmaps. The regions of interest, including IFNγ signature genes and enhancers, were curated from UCSC table browser and BEDOPS v2.4.40 [71]. Peaks were called using MACS2[72] with a p value of $1e^{-3}$. Super-enhancers were identified using ROSE based on H3K27ac peaks and ChIP-seq signals. The peak annotation was also taken from the ROSE output [40, 41]. IFNγ response genes are grouped into two clusters according to their H3K27ac intensity at enhancers or stitched enhancer regions between *MLL3*-mutant and -WT cells. The H3K27ac intensity was calculated from ChIP samples minus the input control, and then averaged for replicates. If the log (fold change) of calculated intensity (sgMLL3/sgNT) is larger than 0, the gene was classified into Cluster 1, otherwise into Cluster 2. The gene list and relevant H3K27ac intensity can be found in Source Data.

STAT1 and H3K27ac ChIP-qPCR

Around 3-4 x 10^6 cells were crosslinked with 1% PFA at room temperature for 10 minutes, quenched with 125 mM Glycine, and rinsed twice with PBS. Cells were lysed and sonicated at 4°C using a Bioruptor Sonicator (Diagenode, UCD-200TM) into 100-600 bp fragments. For STAT1 ChIP, Dynabeads Protein A (Invitrogen, 10002D) was pre-coated with 10 μg anti-STAT1 (clone D1K9Y, Cell Signaling Technology, cat#14994S) or Rabbit

IgG (clone DA1E, Cell Signaling Technology, cat#3900S) antibody. For H3K27ac ChIP, Dynabeads Protein A (Invitrogen, 10002D) was pre-coated with 5 μg anti-H3K27ac (clone D5E4, Cell Signaling Technology, cat#8173S) or Rabbit IgG (clone DA1E, Cell Signaling Technology, cat#3900S) antibody. The samples were incubated with antibody pre-coated beads at 4°C overnight. ChIP samples were then washed, eluted, and reverse crosslinked by overnight incubation at 65 °C. ChIP-DNA was purified by Phenol:Chloroform:Isoamyl Alcohol (Sigma-Aldrich, 77677-100ML) and used for qPCR analysis. ChIP enrichment was calculated relative to input samples. Primers used for ChIP-qPCR are listed in Supplementary Table 3.

ATAC-seq and data analysis

100,000 FACS sorted $CD44^{+}CD104^{high}$ and $CD44^{+}CD104^{-/low}$ cells were used for ATAC-seq assay. All subsequent processing using standard DNA extraction and library preparation protocols were performed by Novogene. Briefly, nuclei are extracted from samples, and the nuclei pellet is resuspended in the Tn5 transposase reaction mix. The transposition reaction is incubated at 37°C for 30 min. Equimolar Adapter1 and Adatper 2 are added after transposition, PCR is then performed to amplify the library. After the PCR reaction, libraries are purified with the AMPure beads. After library quality control, sequencing is performed for different libraries according to the concentration and the demand of data amount on Illumina NovaSeq platform. For data analyses, after investigating the quality of raw data by FASTQC, raw reads were trimmed with BBduk.sh, with the parameters minlen=50 qtrim=r trimq=6 ktrim=r k=25 mink=11. Alignment to the human genome was done with bowtie2 (v. 2.4.1) with the parameters _mm _no-mixed _no-discordant _very-sensitive -X 1000, duplicates were removed with samblaster, and reads from chrM were removed with sambamba. Differential ATAC-seq peaks were called using Genrich (v. 0.6; https://github.com/yejg2017/Genrich), and blacklisted regions were filtered out in the process. Motif analysis on peaks was performed with Homer (v4.11).

Statistics and reproducibility

Data are represented as mean ± SEM. Statistical tests were performed with GraphPad Prism V8.0 (San Diego, CA). Data distribution was assumed to be normal, but this was not formally tested. A two-tailed Student’s t-test was used for comparisons of continuous variables between two groups. One-way ANOVA with Tukey or Dunnett test was used when three or more groups were compared. Two-way ANOVA with multiple comparisons test was used to check the effect of two factors on a dependent variable. Chi-square likelihood ratio test was used to check the limiting dilution assay. Fisher’s right tailed exact test was used in IPA. Kaplan-Meier survival curves were analyzed by using GraphPad Prism V8.0 or 9.0 (San Diego, CA) and p values were calculated using the log-rank test. Statistical details (sample size and specific test performed) for each experiment are denoted in the corresponding figure or figure legend. No statistical method was used to predetermine the sample size. Sample sizes were determined based on previous experimental experience. No data were excluded from the analyses. The *in vitro* cell experiments were not randomized. For *in vivo* tumor onset, progression, and metastasis experiments, mice were randomly assigned to experimental groups. For experiments with chemical inhibitors, samples were randomly allocated to control and experimental groups. For tumor onset and progression

experiments, investigators were blind to group information. For metastasis experiments, investigators were not blind to group information, however, data reported for metastasis experiments are not subjective but rather based on quantitative total photo flux by IVIS. Investigators were not blinded to cell genotypes during in vitro experiments, because experiments required frequent intervention by investigators to maintain cell lines.

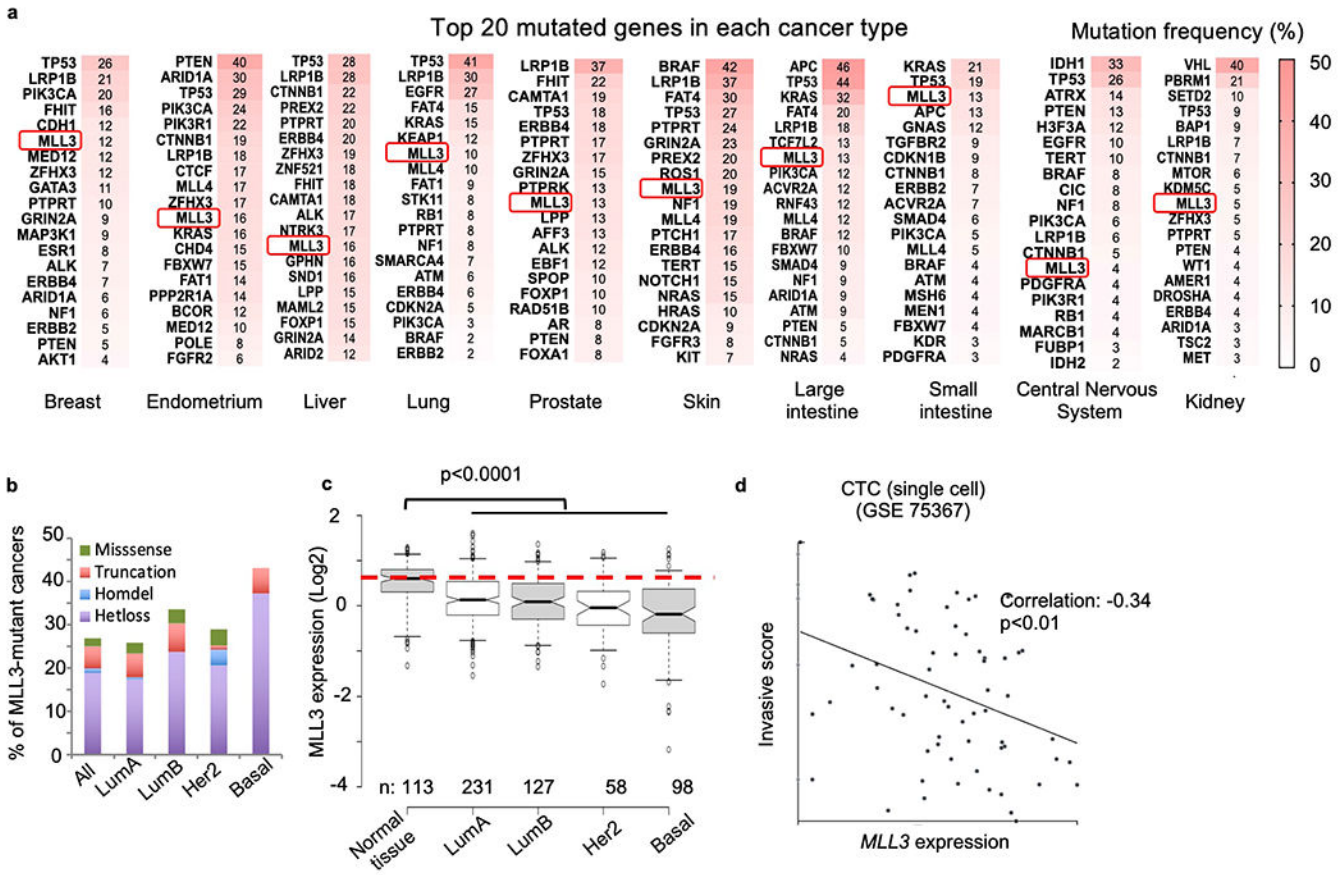
Data Availability

RNA–seq, ChIP-seq, ATAC-seq and microarray data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE171447. The human reference genome (GRCh38) was used as a reference genome for the alignment. The human breast cancer data were derived from the TCGA Research Network: http://cancergenome.nih.gov/. The data-set derived from this resource that supports the findings of this study is available in the details of Source data. Other source data for all figures are provided with the paper. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Code Availability

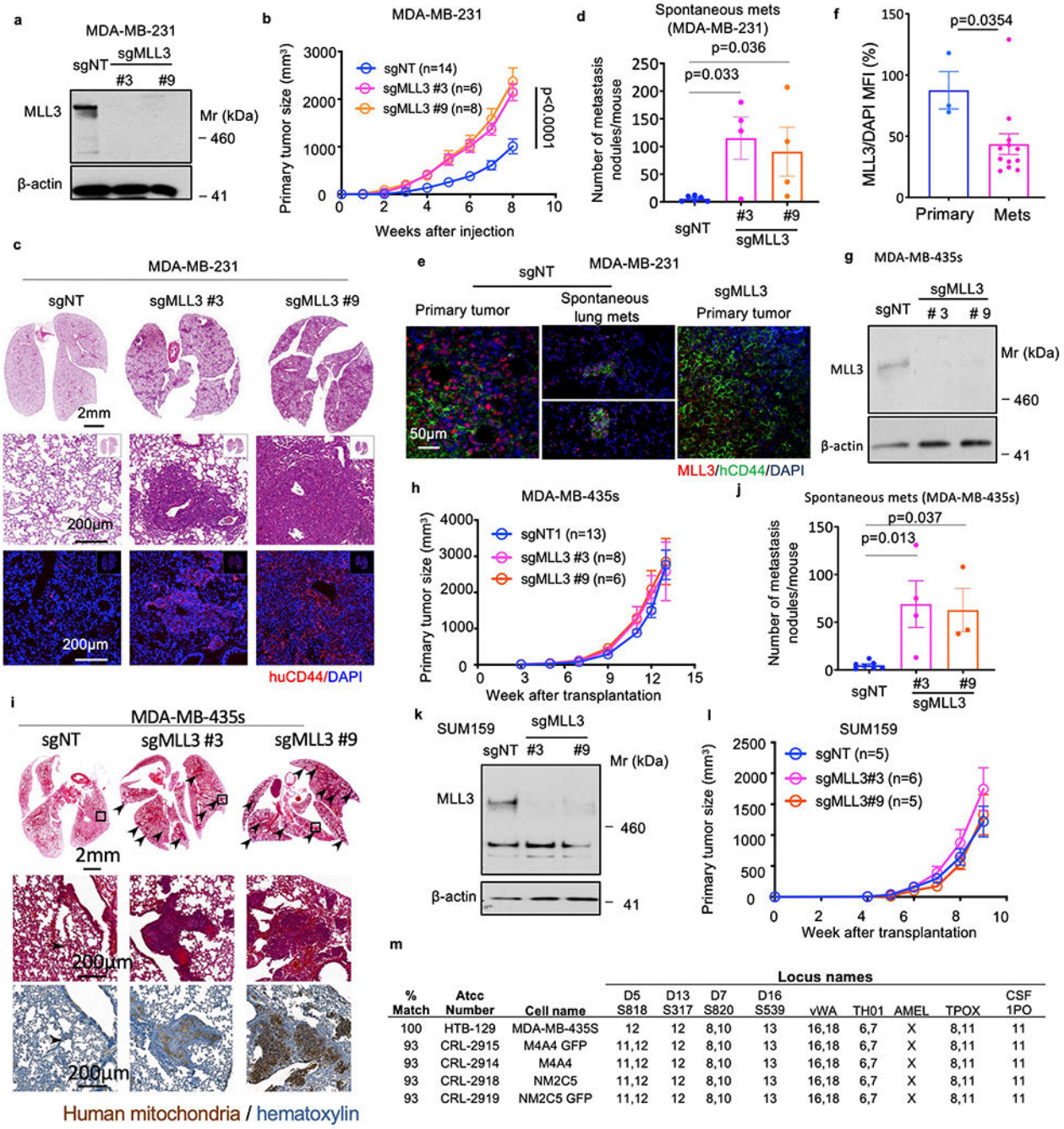
The code for in-house scripts used in this study was deposited into the following link: https://github.com/zcmit/EinsteinMed/tree/main/Cui_etal.2022.

Extended Data



Extended Data Fig.1. *MLL3* is one of the most frequently altered genes in breast cancer and other cancer types.

a, Top 20 mutated genes in various cancer types as described in COSMIC v91. The mutation frequencies of individual genes are shown. b, Frequencies of different *MLL3* alterations in various breast cancer subtypes. Data were obtained from the TCGA dataset via cBioPortal. c, *MLL3* expression levels (Log2 of normalized RNA-seq values) in normal breast tissues and various breast cancer subtypes. Data were obtained from the TCGA BRCA dataset via UCSC Cancer Browser. Values from minimum to maximum are shown by the box and whiskers. Box plots indicate median (middle line), 25th and 75th percentile (box) and 5th and 95th percentile (whiskers). d, Correlation analysis of invasion ability and *MLL3* expression levels in single breast CTCs (GSE75367, analyzed using CancerSEA). *P* values were determined by two-tailed Student's t-test (c). *P* values for d were calculated by CancerSEA. Source data are provided.



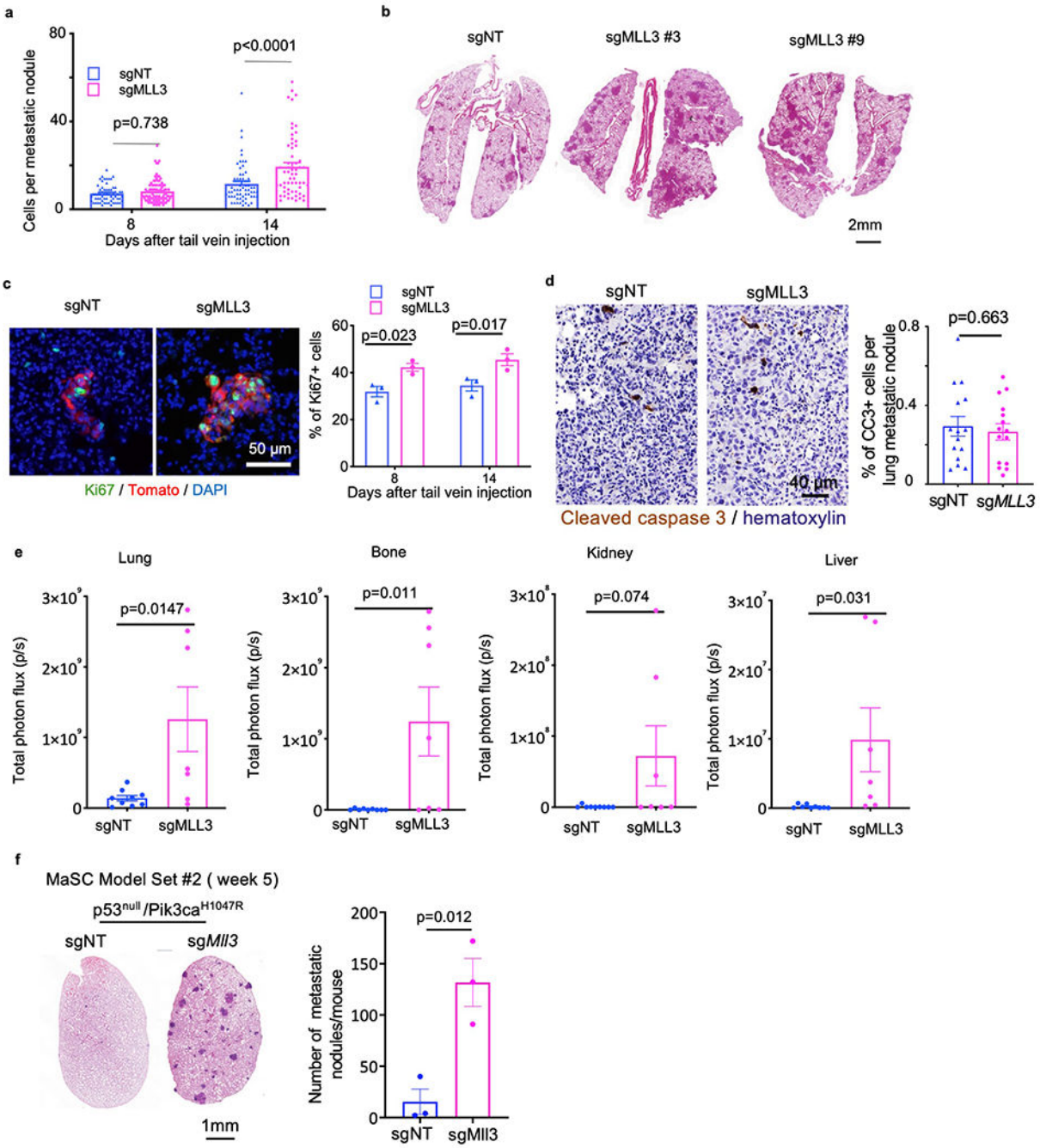
m

			Locus names								
% Match	Atcc Number	Cell name	D5 S818	D13 S317	D7 S820	D16 S539	vWA	TH01	AMEL	TPOX	CSF 1PO
100	HTB-129	MDA-MB-435S	12	12	8,10	13	16,18	6,7	X	8,11	11
93	CRL-2915	M4A4 GFP	11,12	12	8,10	13	16,18	6,7	X	8,11	11
93	CRL-2914	M4A4	11,12	12	8,10	13	16,18	6,7	X	8,11	11
93	CRL-2918	NM2C5	11,12	12	8,10	13	16,18	6,7	X	8,11	11
93	CRL-2919	NM2C5 GFP	11,12	12	8,10	13	16,18	6,7	X	8,11	11

Extended Data Fig. 2. *MLL3* deletion enhances breast tumor metastasis.

a, MLL3 western blot of indicated MDA-MB-231 cells ($n = 3$ independent experiments). **b**, MDA-MB-231 orthotopic tumor growth. 1×10^6 sgNT ($n = 14$) and sgMLL3 (#3, $n = 6$; #9, $n = 8$ xenografts) cells were injected into each NOD/SCID mouse. **c,** Representative H&E and human CD44 staining of lung sections from mice bearing indicated MDA-MB-231 orthotopic tumors at week 13 after injection (sgNT, $n = 6$, sgMLL3#3, $n = 4$; sgMLL3#9, $n = 4$ mice). **d,** Number of metastatic nodules in the lung as generated in **b** and **c**. **e,** Representative MLL3 immunofluorescence of MDA-MB-231 primary tumor and

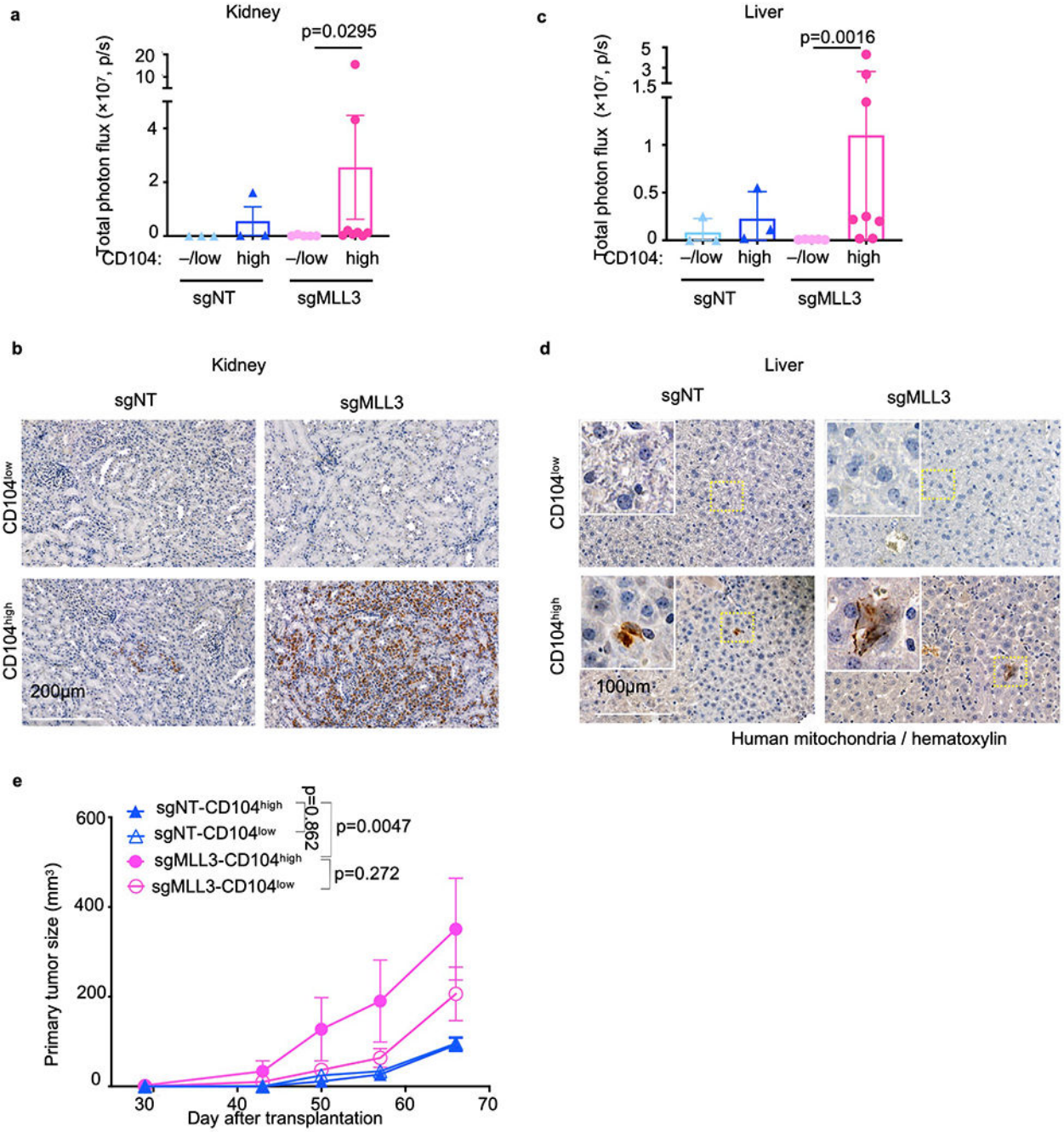
spontaneous lung metastases. Human-specific CD44 staining identifies tumor cells. **f,** MLL3 expression level in MDA-MB-231 primary tumors and spontaneous lung metastases, as shown in **e** (primary tumor, $n = 3$ tumors; lung metastases, $n = 12$ lung nodules). **g,** MLL3 western blot ($n = 3$ independent experiments). **h,** MDA-MB-435s primary tumor growth rate. 1×10^6 cells were injected into each NOD/SCID mouse. **i,** Representative H&E and human mitochondria staining of lung sections from mice bearing indicated MDA-MB-435s orthotopic tumors at week 13 after injection. Arrows point to examples of metastases. **j,** Number of metastatic nodules in the lung as generated in **h** and **i** (sgNT, $n = 7$; sgMLL3 (#3), $n = 4$; sgMLL3 (#9), $n = 3$ mice). **k,** MLL3 western blot ($n = 2$ independent experiments). **l,** SUM159 primary tumor growth rate (sgNT, $n = 5$; sgMLL3 (#3), $n = 6$; sgMLL3 (#9), $n = 5$ tumors). **m,** STR report of MDA-MB-435s cell authentication. All data are represented as mean ± SEM. *P* values were determined by two-way RM ANOVA with correction using Geisser-Greenhouse method (**b**), one-way ANOVA with Dunnett's multiple comparisons test (**d**, **j**) or two-tailed Student's t-test (**f**). Source data are provided.



Extended Data Fig. 3. Loss of *MLL3* promotes metastatic colonization.

a, The average number of tumor cells in each lung metastatic nodule from the mice injected with sgNT or sgMLL3 MDA-MB-231 cells at 8 or 14 days post tail vein injection (8 days: sgNT, $n = 62$; sgMLL3, $n = 93$; 14 days: sgNT, $n = 61$; sgMLL3, $n = 60$ nodules from 3 mice per group). **b,** H&E staining of lung sections from mice 4 weeks after tail vein injection with sgNT or sgMLL3 MDA-MB-231 cells. **c,** Representative Ki67 immunofluorescence staining of lung metastases in mice at day 8 or day 14 after tail-vein injection with sgNT or sgMLL3 MDA-MB-231 cells. The graph shows percentages of

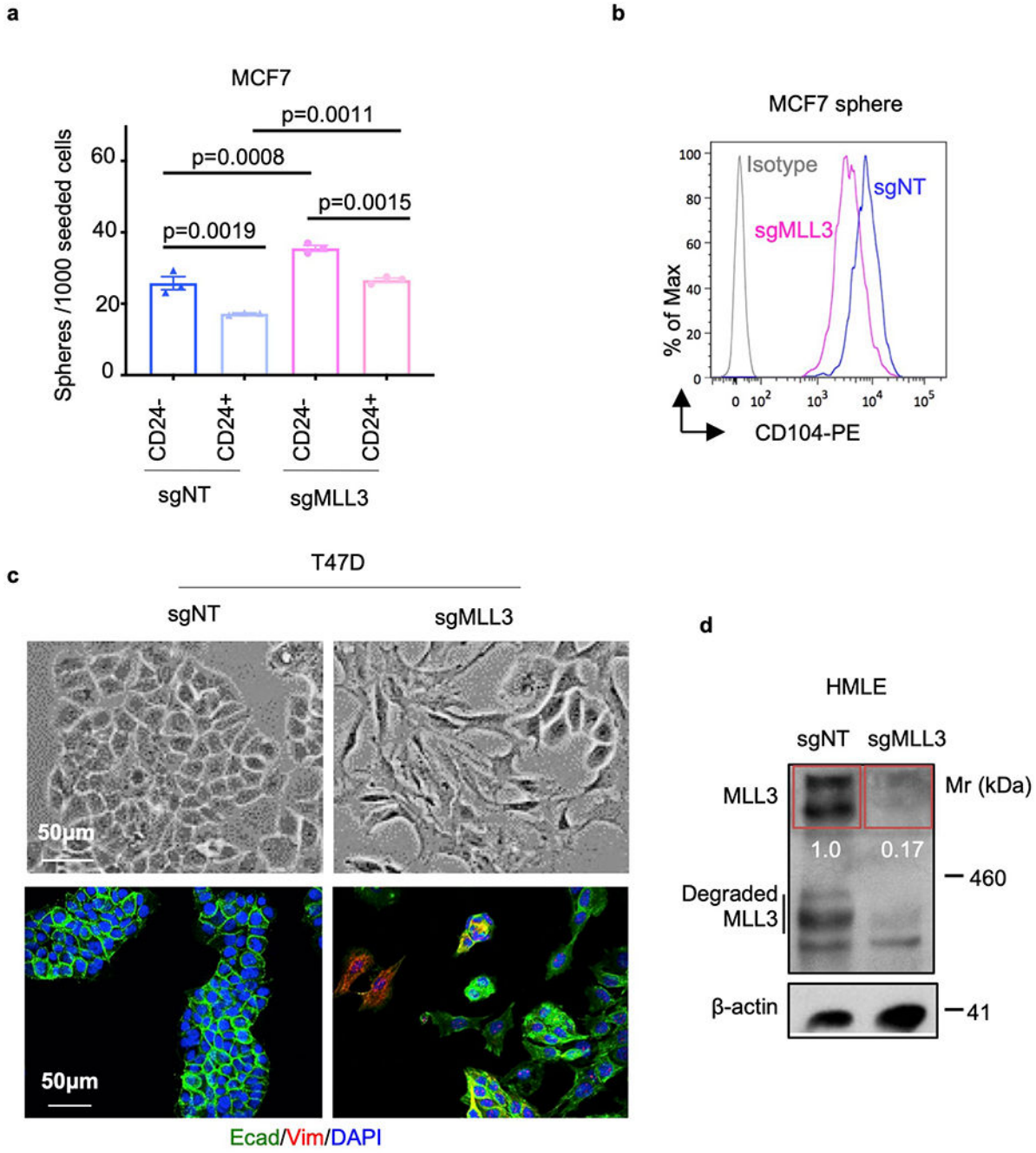
Ki67^{+} tumor cells in lung metastases from each mouse (n = 3 mice). **d,** Representative images and quantification of the percentage of cleaved caspase 3-positive cells in size-matched lung metastatic nodules formed by sgNT or sgMLL3 MDA-MB-231 cells 6 weeks after tail-vein injection (n = 15 tumor regions of 3 mice per group). **e,** Quantification of *ex vivo* bioluminescence signals of lung, bone, kidney and liver metastasis burdens in animals injected with sgNT (n = 9) or sgMLL3 (n = 7 mice) SUM159 cell 4 weeks after tail vein injection. **f,** Lung colonization of sgNT and sgMll3 *p53*^{null}/*Pik3ca*^{H1047R} cells injected by tail vein (Set 2). Representative H&E staining and number of metastatic nodules in the left lung lobe of each animal (n = 3 per group) were shown. 500,000 luciferase-labeled cells were injected into FVB mice via the tail vein and analyzed at week 5. All data are represented as mean ± SEM. P values were determined by two-way ANOVA with Šídák's multiple comparisons test (**a**, **c**), or two-tailed Student's t-test (**d-f**). Source data are provided as a source data file.



Extended Data Fig. 4. Loss of *MLL3* greatly potentiates the metastasis-initiating ability of hybrid E/M cells.

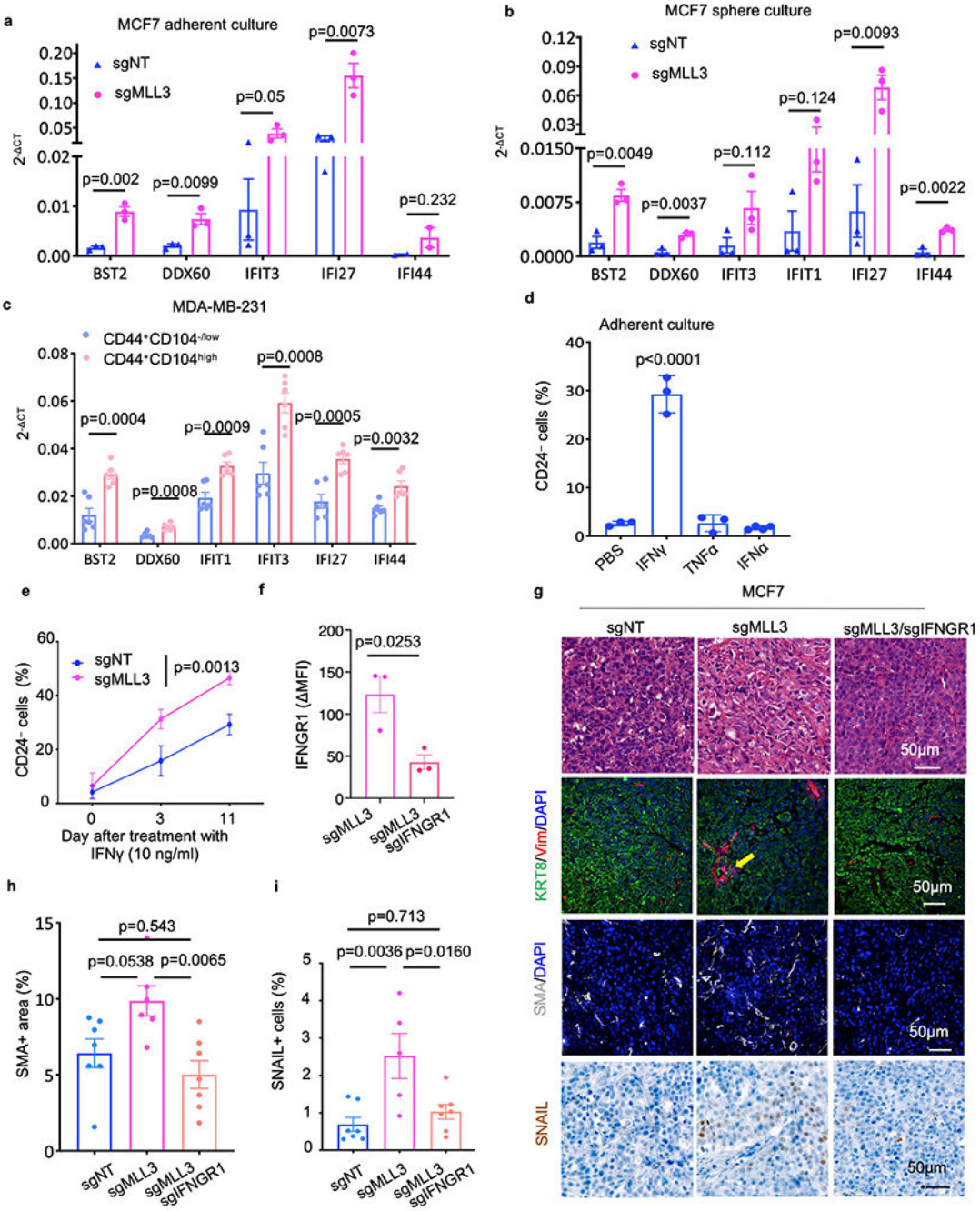
a, Bioluminescence signals of kidney metastatic burdens in animals injected with the indicated cell types via the tail vein (sgNT-CD44^{+}CD104^{low}, n = 3, sgNT-CD44^{+}CD104^{high}, n = 3, sgMLL3-CD44^{+}CD104^{low}, n = 5, sgMLL3-CD44^{+}CD104^{high}, n = 8 mice). b, Anti-human mitochondria staining showing metastatic lesions in the kidney. c, Bioluminescence signals of liver metastasis burdens in animals injected with the indicated cell types via the tail vein. d, Anti-human mitochondria staining showing metastatic lesions in the liver. e, Tumor growth curve of CD104^{high}/CD44^{+} and CD104^{low}/CD44^{+} WT(sgNT) or

MLL3-mutant (sgMLL3) MDA-MB-231 cells in NSG mice. 2×10^5 cells were injected orthotopically into mammary gland fat pads of female NSG mice ($n = 6$ for each group). All data are represented as mean ± SEM. *P* values were determined by two-tailed Student's t-test (a, c), or two-way ANOVA with correction using Geisser-Greenhouse method (e). Source data are provided.



Extended Data Fig. 5. *MLL3* loss enables epithelial cells to gain mesenchymal features.

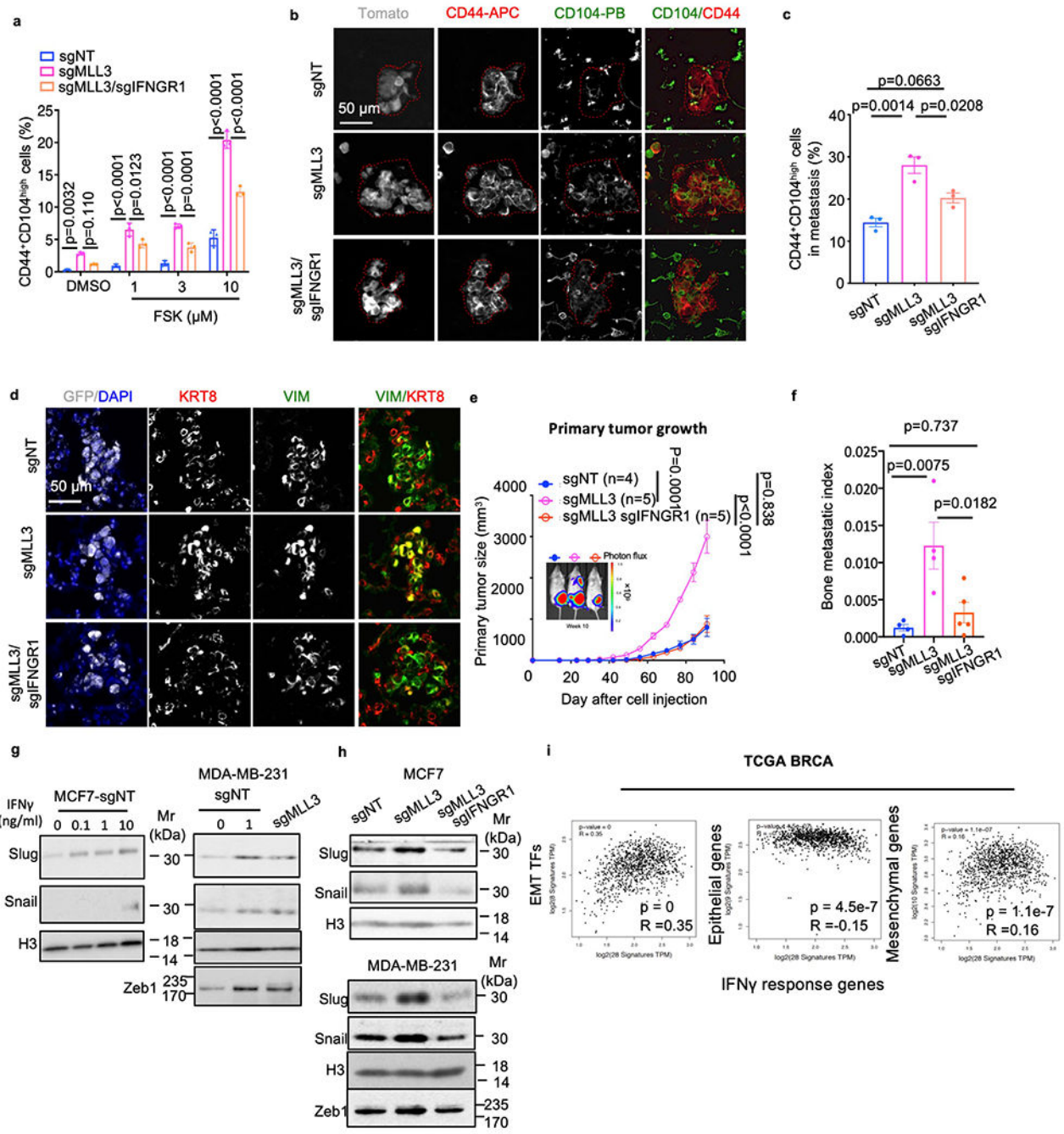
a, Sphere-forming efficiency of $CD24^-$ and $CD24^+$ cells sorted from sgNT or sgMLL3 MCF7 spheres ($n = 3$ biological samples for each group). **b,** Flow cytometric analysis of CD104 expression in sgNT or sgMLL3 MCF7 spheres. **c,** Phase-contrast images and E-cadherin and vimentin immunofluorescence staining in sgNT and sgMLL3 T47D cells ($n = 3$ independent experiments). **d,** *MLL3* western blot of sgNT and sgMLL3 HMLE cells. MLL3 is a very large protein (> 500 kDa), thus prone to degradation during western blot ($n = 2$ independent experiments). All data are represented as mean ± SEM. *P* values were determined by one-way ANOVA with Tukey's multiple comparisons test (**a**). Source data are provided.



Extended Data Fig. 6. *MLL3* loss upregulates the interferon-γ pathway.

a, Expression of ISGs in the WT and *MLL3*-mutant MCF7 cells in adherent culture, as measured by qRT-PCR ($n = 3$ biological samples). **b,** Expression of ISGs in the WT and *MLL3*-mutant MCF7 cells in sphere culture, as measured by qRT-PCR ($n = 3$ biological samples). **c,** Expression of ISGs in the $CD44^{+}CD104^{high}$ and $CD44^{+}CD104^{low}$ *MLL3*-WT MDA-MB-231 cells, as measured by qRT-PCR ($n = 3$ biological samples with 2 technique replicates). **d,** Effect of inflammatory cytokines on the frequency of $CD24^{-}$ MCF7 cells. Cells were treated with PBS, IFNγ (10 ng/ml), TNFα (10 ng/ml) for 11 days, or IFNα

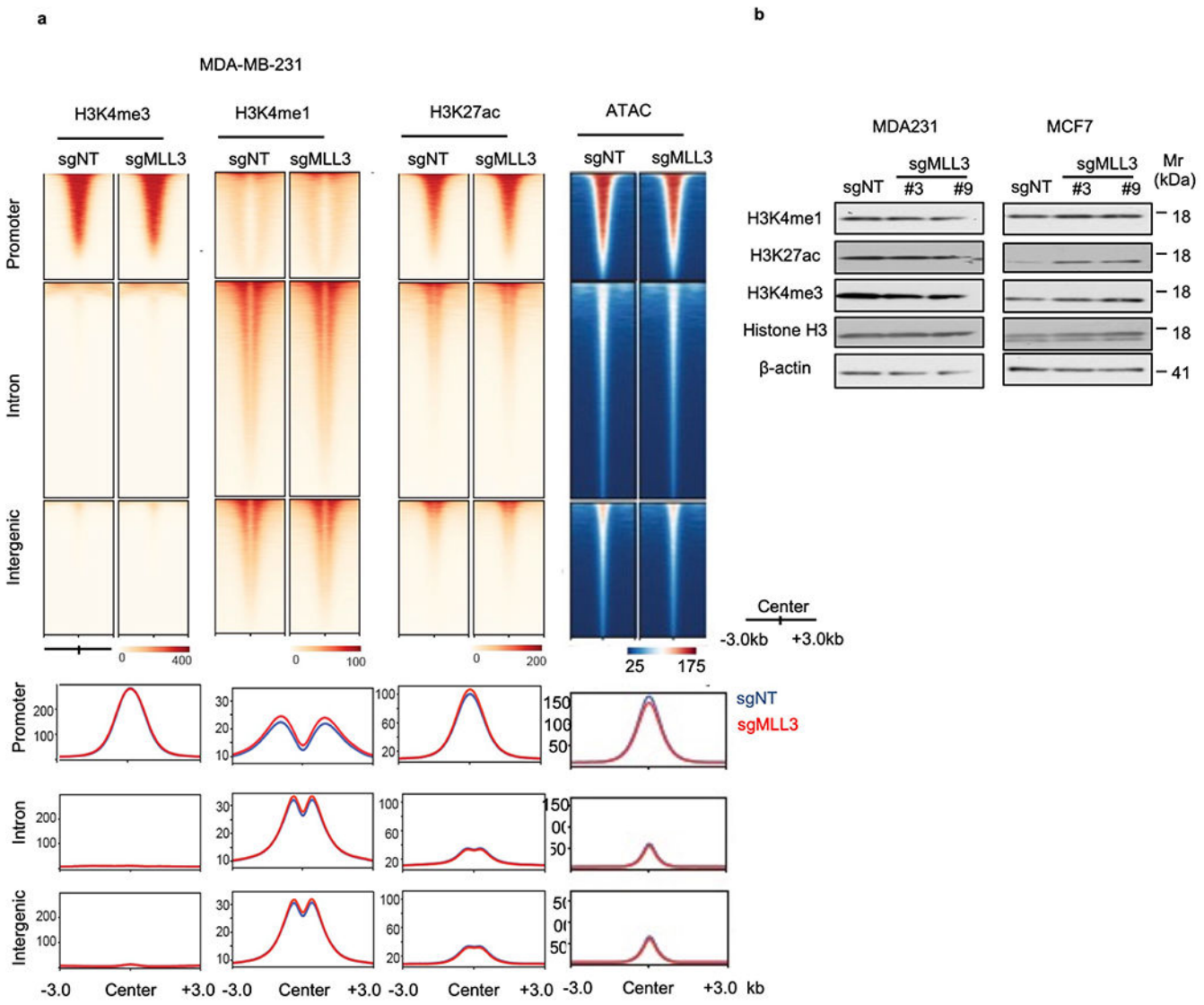
(2,000 unit/ml) for 17 days in adherent culture (n = 3 biological samples). **e,** Kinetics of $CD24^-$ cell induction by IFNγ in sgNT or sgMLL3 MCF7 cells (n = 3 biological samples in each group). **f,** IFNGR1 flow cytometry measuring the CRISPR deletion efficiency (n = 3 independent experiments). **g,** Representative images of H&E staining, Keratin 8(KRT8)/Vimentin(Vim), smooth-muscle actin (SMA) immunofluorescence, and SNAIL immunohistochemistry of sgNT, sgMLL3, and sgMLL3/sgIFNGR1 MCF7 tumors. **h,** Quantification of SMA expression in sgNT, sgMLL3 and sgMLL3/sgIFNGR1 MCF7 tumors (sgNT, n = 7; sgMLL3, n = 6; and sgMLL3/sgIFNGR1, n = 7 tumors). **i,** Quantification of $SNAIL^+$ cells in sgNT, sgMLL3 and sgMLL3/sgIFNGR1 MCF7 tumors (sgNT, n = 7; sgMLL3, n = 5; and sgMLL3/sgIFNGR1, n = 7 tumors). All data are represented as mean ± SEM. P values were determined by two-tailed Student's t-test (**a**-**c**, **f**), one-way ANOVA with Dunnett's or Turkey's multiple comparisons test (**d**, **h**-**i**), or Ordinary two-way ANOVA (**e**). Source data are provided.



Extended Data Fig. 7. *IFNGR1* deletion abolished the enhancement of breast cancer metastasis by *MLL3* loss.

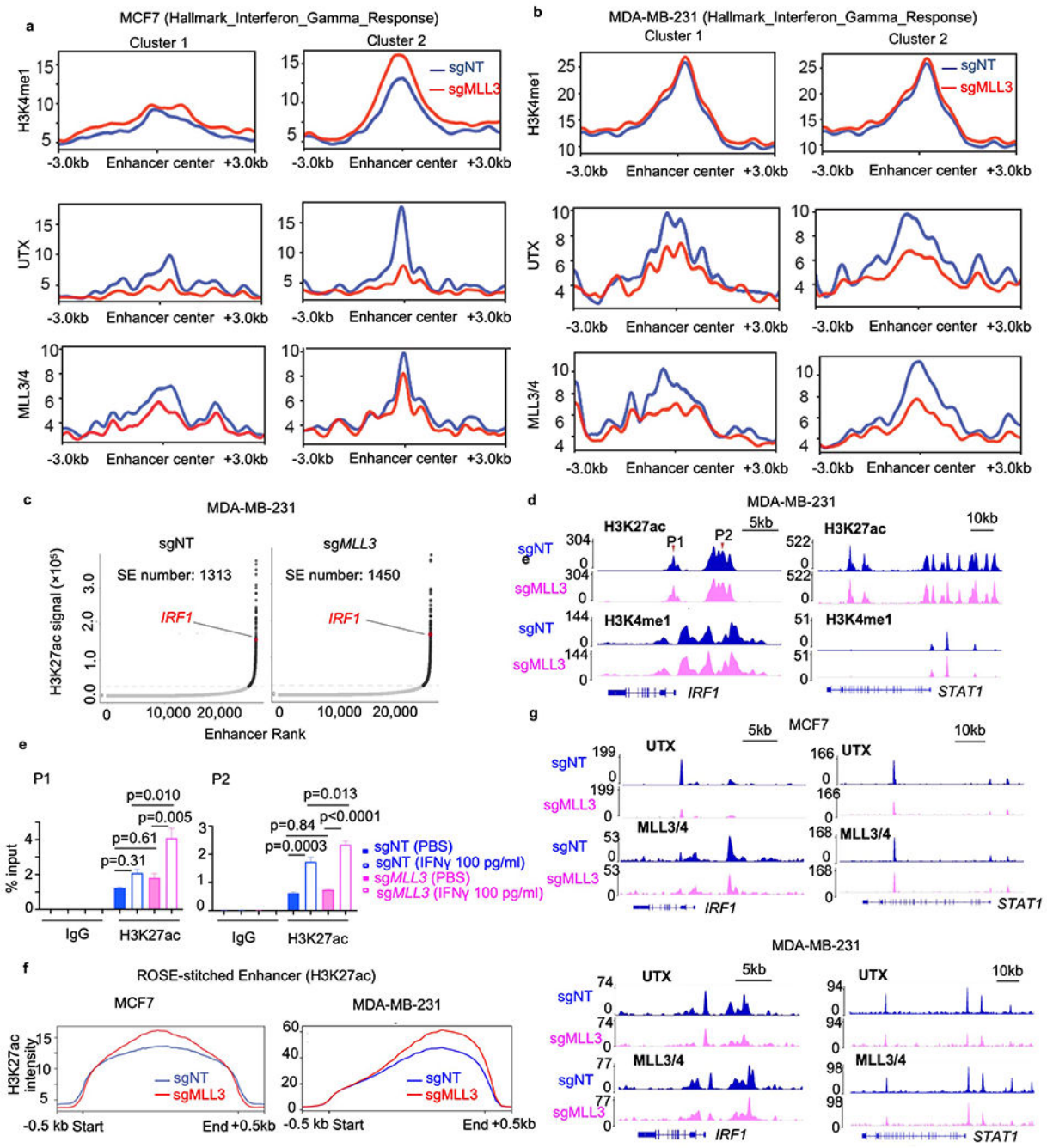
a, The induction of $CD44^{+}CD104^{high}$ cells by forskolin in indicated MDA-MB-231 cells (n = 3 biological samples in each group). **b,** Representative images of CD104 and CD44 staining in lung metastases formed by indicated MDA-MB-231 cells at day 14 post tail-vein injection. **c,** Quantification of CD104 and CD44 co-staining of cells in lung metastases generated as in **b** (n = 3 mice in each group). **d,** Representative images of GFP, Keratin 8 (KRT8), and vimentin (Vim) co-staining in lung metastases generated as in **b**. **e,** Growth

rates and representative bioluminescence images of NOD/SCID mice at 10 weeks after orthotopic injection of sgNT (n = 4), sgMLL3 (n = 5), or sgMLL3/sgIFNGR1 (n = 5 mice) MDA-MB-231 cells. **f,** Quantification of spontaneous bone metastasis in sgNT (n = 4), sgMLL3 (n = 4), and sgMLL3/sgIFNGR1 (n = 5) MDA-MB-231 tumor-bearing mice as generated in **e** at week 13. Metastatic index = bone photon flux/primary tumor photon flux. **g,** Western blot of EMT TFs in sgNT MCF7 cells treated with indicated levels of IFNγ for 24h (Left) or in sgNT and sgMLL3 MDA-MB-231 cells treated with 1 ng/ml IFNγ for 4 days (right) (n = 2 independent experiments). **h,** Western blot of EMT TFs in indicated cells (n = 2 independent experiments). **i,** Correlation analysis of IFNγ response genes and EMT TFs, epithelial signature genes, or mesenchymal signature genes in the BRCA TCGA patient cohort by using GEPIA2 (n = 1083 patient samples). Gene lists were provided in Supplementary Table 4. All data are represented as mean ± SEM. P values were determined by two-way ANOVA with Šídák's multiple comparisons test (**a**) or with correction using Geisser-Greenhouse method (**e**), one-way ANOVA with Turkey's multiple comparisons test (**c**, **f**), or Pearson's correlation coefficient analysis (**i**). Source data are provided.



Extended Data Fig. 8. The effect of *MLL3* loss on genome-wide changes of H3K27ac, H3K4me1 and H3K4me3 in MCF7 and MDA-MB-231 cells.

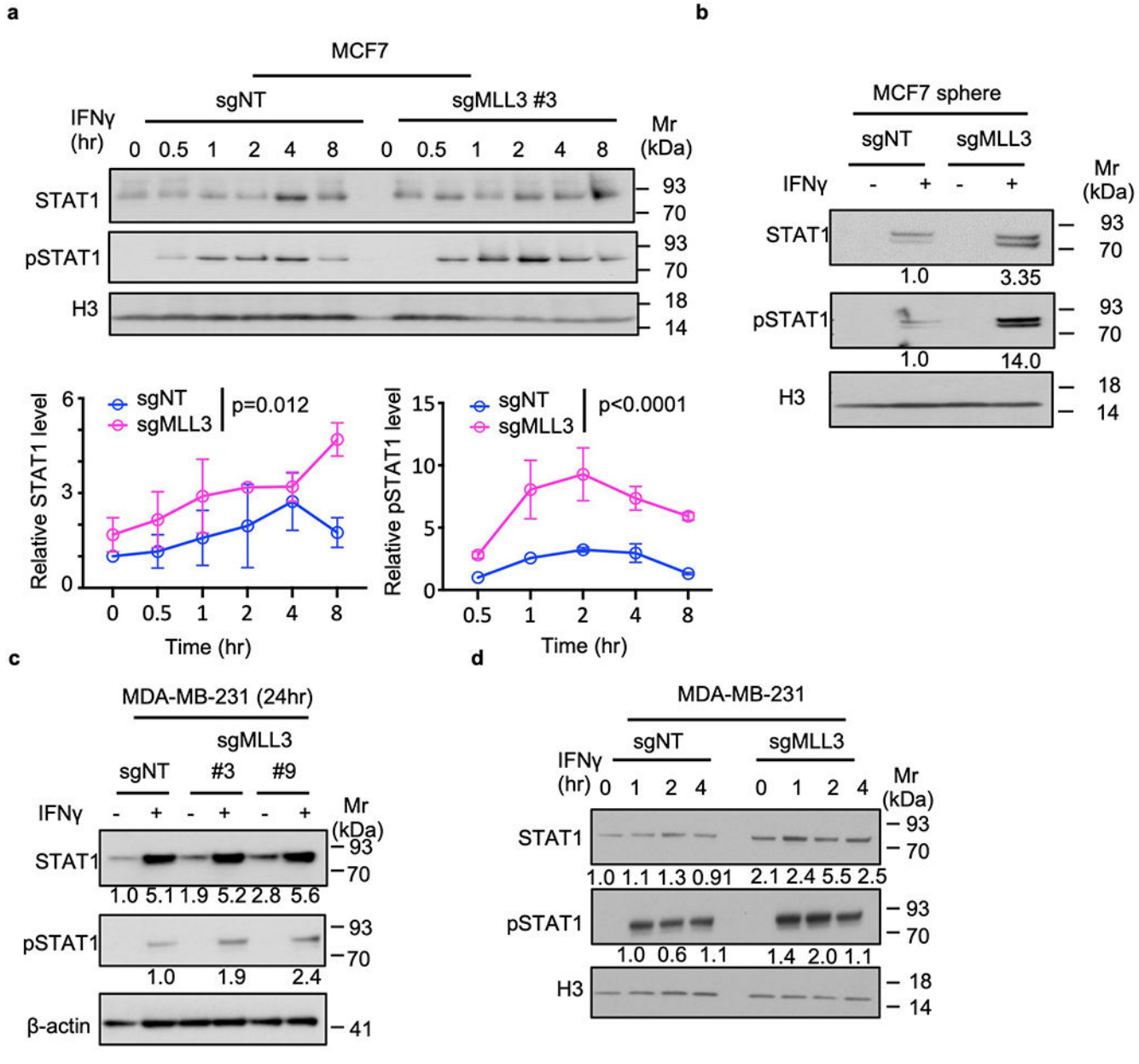
a, Heatmaps of ChIP-seq signals for the indicated histone marks and ATAC-seq signals in a 6kb window grouped by localization at the promoter, intron, and intergenic regions in sgNT or sgMLL3 MDA-MB-231 cells. **b,** Western blot analysis of protein levels of H3K27ac, H3K4me1, and H3K4me3 in sgNT or sgMLL3 MCF7 and MDA-MB-231 cells ($n = 2$ independent experiments).



Extended Data Fig. 9. *MLL3* loss increases H3K27ac levels in the enhancers of a subset of IFNγ pathway genes.

a, H3K4me1, UTX and MLL3/4 binding signals at enhancer regions of IFNγ response genes in sgNT and sgMLL3 MCF7 cells based on ChIP-seq. **b,** H3k4me1, UTX and MLL3/4 binding signals at enhancer regions of IFNγ response genes in sgNT and sgMLL3 MDA-MB-231 cells based on ChIP-seq. **c,** Hockey-stick plots generated by ROSE showing ranked enhancers based on H3K27ac ChIP-seq signal intensity in sgNT and sgMLL3 MDA-MB-231 cells (SE: super enhancer). **d,** H3K27ac and H3K4me1 binding signals at *IRF1*

and *STAT1* gene loci in sgNT and sgMLL3 MDA-MB-231 cells. **e,** ChIP-qPCR analysis of H3K27ac occupancy at the *IRF1* enhancer regions in sgNT and sgMLL3 MDA-MB-231 cells treated with PBS or 100 pg/ml IFNγ for 4h. ChIP samples were analyzed by qPCR and normalized to input. The regions detected by PCR were labeled in D (P1 and P2) ($n = 3$ technique replicates each experiment, similar results were repeated 3 times). **f,** The average signal intensity of H3K27ac at enhancer regions identified by ROSE in sgNT and sgMLL3 MCF7 and MDA-MB-231 cells. **g,** UTX and MLL3/4 binding signals at *IRF1* and *STAT1* gene loci in sgNT and sgMLL3 MCF7 and MDA-MB-231 cells, respectively. All data are represented as mean± SEM. *P* values were determined by one-way ANOVA with Turkey's multiple comparisons test (**e**). Source data are provided.



Extended Data Fig. 10. *MLL3* loss leads to increased phosphorylation and expression of STAT1 upon IFNγ stimulation.

a, Western blot analysis of pSTAT1 and STAT1 in sgNT and sgMLL3 MCF7 cells treated with 10 ng/ml IFNγ for 0.5, 1, 2, 4, or 8h. Two experimental repeats were quantified. **b,** Western blot analysis of pSTAT1 and STAT1 in sgNT and sgMLL3 MCF7 spheres treated with 10 ng/ml IFNγ for 5 days (n = 3 biologically independent experiments). **c,** Western blot analysis of pSTAT1 and STAT1 in sgNT and sgMLL3 MDA-MB-231 cells treated with 100 pg/ml IFNγ for 24h (n = 2 independent experiments). **d,** Western blot analysis of pSTAT1 and STAT1 in sgNT and sgMLL3 MDA-MB-231 cells treated with 10 ng/ml IFNγ for 1, 2, or 4hrs (n = 2 biological independent experiments). All data are represented as mean ± SEM. P values were determined by two-way ANOVA with Šídák's multiple comparisons test (**a**). Source data are provided.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Flow Cytometry, Histopathology, Analytical Imaging, and Stem Cell Isolation core facilities of Albert Einstein College of Medicine for technical assistance, supported by Einstein Cancer Center Support Grant (P30 CA013330), the New York State Department of Health / NYSTEM Program (C029154), and P250 high-capacity slide scanner (1S10OD019961-01). We thank Guojia Xie for assistance with MLL3 western blot. This work is supported by the DOD BCRP grant W81XWH-16-1-0311, the NIH/NCI grant 1R01CA212424 and the Mary Kay Foundation grant 04-19 (to W.G). W.G. is a V Scholar of the V Foundation for Cancer Research. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript

References

1. Lambert AW, Pattabiraman DR & Weinberg RA Emerging Biological Principles of Metastasis. Cell 168, 670–691 (2017). [PubMed: 28187288]
2. Steeg PS Targeting metastasis. Nat Rev Cancer 16, 201–218 (2016). [PubMed: 27009393]
3. Esposito M, Ganesan S & Kang Y Emerging strategies for treating metastasis. Nature Cancer 2, 258–270 (2021). [PubMed: 33899000]
4. Nieto MA, Huang RY, Jackson RA & Thiery JP Emt: 2016. Cell 166, 21–45 (2016). [PubMed: 27368099]
5. Yang J et al. Guidelines and definitions for research on epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 21, 341–352 (2020). [PubMed: 32300252]
6. Tsai JH & Yang J Epithelial-mesenchymal plasticity in carcinoma metastasis. Genes Dev 27,2192–2206(2013). [PubMed: 24142872]
7. Brabletz T To differentiate or not--routes towards metastasis. Nat Rev Cancer 12, 425–436(2012). [PubMed: 22576165]
8. Tsai JH, Donaher JL, Murphy DA, Chau S & Yang J Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell 22, 725–736 (2012). [PubMed: 23201165]
9. Ocaña OH et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 22, 709–724 (2012). [PubMed: 23201163]
10. Williams ED, Gao D, Redfern A & Thompson EW Controversies around epithelial-mesenchymal plasticity in cancer metastasis. Nat Rev Cancer 19, 716–732 (2019). [PubMed: 31666716]
11. Yuan S, Norgard RJ & Stanger BZ Cellular Plasticity in Cancer. Cancer Discov 9, 837–851 (2019). [PubMed: 30992279]

12. Lu W & Kang Y Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis. Dev Cell 49, 361–374 (2019). [PubMed: 31063755]
13. Pastushenko I & Blanpain C EMT Transition States during Tumor Progression and Metastasis. Trends Cell Biol 29, 212–226 (2019). [PubMed: 30594349]
14. Tripathi S, Levine H & Jolly MK The Physics of Cellular Decision Making During Epithelial-Mesenchymal Transition. Annu Rev Biophys 49, 1–18 (2020). [PubMed: 31913665]
15. Kroger C et al. Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. Proc Natl Acad Sci U S A 116, 7353–7362 (2019). [PubMed: 30910979]
16. Bierie B et al. Integrin-β4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. Proc Natl Acad Sci U S A 114, E2337–e2346 (2017). [PubMed: 28270621]
17. Pastushenko I et al. Identification of the tumour transition states occurring during EMT. Nature 556, 463–468 (2018). [PubMed: 29670281]
18. Aiello NM et al. EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. Dev Cell 45, 681–695.e684 (2018). [PubMed: 29920274]
19. Pastushenko I et al. Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis. Nature 589, 448–455 (2021). [PubMed: 33328637]
20. Lawrence MS et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495–501 (2014). [PubMed: 24390350]
21. Kandoth C et al. Mutational landscape and significance across 12 major cancer types. Nature 502, 333–339 (2013). [PubMed: 24132290]
22. TCGA Network Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70 (2012). [PubMed: 23000897]
23. Ciriello G. et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell 163, 506–519 (2015). [PubMed: 26451490]
24. Bailey MH et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell 174,1034–1035 (2018). [PubMed: 30096302]
25. Razavi P. et al. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. Cancer Cell 34, 427–438.e426 (2018). [PubMed: 30205045]
26. Wang L. et al. Resetting the epigenetic balance of Polycomb and COMPASS function at enhancers for cancer therapy. Nature medicine 24, 758–769 (2018).
27. Chen C. et al. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. Cancer Cell 25, 652–665 (2014). [PubMed: 24794707]
28. Lee J. et al. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proc Natl Acad Sci U S A 106, 8513–8518 (2009). [PubMed: 19433796]
29. Zhang Z. et al. Mammary-Stem-Cell-Based Somatic Mouse Models Reveal Breast Cancer Drivers Causing Cell Fate Dysregulation. Cell Rep 16, 3146–3156 (2016). [PubMed: 27653681]
30. Biankin AV et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 491, 399–405 (2012). [PubMed: 23103869]
31. Curtis C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486, 346–352 (2012). [PubMed: 22522925]
32. Waldron L. et al. Expression profiling of archival tumors for long-term health studies. Clin Cancer Res 18, 6136–6146 (2012). [PubMed: 23136189]
33. Esposito M. et al. Bone vascular niche E-selectin induces mesenchymal-epithelial transition and Wnt activation in cancer cells to promote bone metastasis. Nat Cell Biol 21, 627–639 (2019). [PubMed: 30988423]
34. Pattabiraman DR et al. Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. Science 351, aad3680 (2016). [PubMed: 26941323]
35. Mani SA et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 133, 704–715 (2008). [PubMed: 18485877]

36. Singh S. et al. Loss of ELF5-FBXW7 stabilizes IFNGR1 to promote the growth and metastasis of triple-negative breast cancer through interferon-γ signalling. Nat Cell Biol 22, 591–602 (2020). [PubMed: 32284542]
37. Heinz S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589 (2010). [PubMed: 20513432]
38. Gala K. et al. KMT2C mediates the estrogen dependence of breast cancer through regulation of ERalpha enhancer function. Oncogene 37, 4692–4710 (2018). [PubMed: 29755131]
39. Jozwik KM, Chernukhin I, Serandour AA, Nagarajan S & Carroll JS FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. Cell Rep 17, 2715–2723 (2016). [PubMed: 27926873]
40. Whyte WA et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013). [PubMed: 23582322]
41. Loven J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334 (2013). [PubMed: 23582323]
42. Platanias LC Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat Rev Immunol 5, 375–386 (2005). [PubMed: 15864272]
43. Cheon H & Stark GR Unphosphorylated STAT1 prolongs the expression of interferon-induced immune regulatory genes. Proc Natl Acad Sci U S A 106, 9373–9378 (2009). [PubMed: 19478064]
44. Shu S. et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. Nature 529, 413–417 (2016). [PubMed: 26735014]
45. Ge JY et al. Acquired resistance to combined BET and CDK4/6 inhibition in triple-negative breast cancer. Nature communications 11, 2350 (2020).
46. Sahni JM et al. Bromodomain and Extraterminal Protein Inhibition Blocks Growth of Triple-negative Breast Cancers through the Suppression of Aurora Kinases. The Journal of biological chemistry 291, 23756–23768 (2016). [PubMed: 27650498]
47. Schafer JM et al. Targeting MYCN-expressing triple-negative breast cancer with BET and MEK inhibitors. Science translational medicine 12 (2020).
48. Shu S. et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. Nature 529, 413–417 (2016). [PubMed: 26735014]
49. Shu S. et al. Synthetic Lethal and Resistance Interactions with BET Bromodomain Inhibitors in Triple-Negative Breast Cancer. Mol Cell 78, 1096–1113.e1098 (2020). [PubMed: 32416067]
50. Cho SJ et al. KMT2C Mutations in Diffuse-Type Gastric Adenocarcinoma Promote Epithelial-to-Mesenchymal Transition. Clin Cancer Res 24, 6556–6569 (2018). [PubMed: 30108106]
51. Na F. et al. KMT2C deficiency promotes small cell lung cancer metastasis through DNMT3A-mediated epigenetic reprogramming. Nat Cancer 3, 753–767 (2022). [PubMed: 35449309]
52. Zhang Y. et al. Genome-wide CRISPR screen identifies PRC2 and KMT2D-COMPASS as regulators of distinct EMT trajectories that contribute differentially to metastasis. Nat Cell Biol 24, 554–564 (2022). [PubMed: 35411083]
53. Zitvogel L, Galluzzi L, Kepp O, Smyth MJ & Kroemer G Type I interferons in anticancer immunity. Nat Rev Immunol 15, 405–414 (2015). [PubMed: 26027717]
54. Minn AJ & Wherry EJ Combination Cancer Therapies with Immune Checkpoint Blockade: Convergence on Interferon Signaling. Cell 165, 272–275 (2016). [PubMed: 27058661]
55. Celiá-Terrassa T. et al. Normal and cancerous mammary stem cells evade interferon-induced constraint through the miR-199a-LCOR axis. Nat Cell Biol 19, 711–723 (2017). [PubMed: 28530657]
56. Meeks JJ & Shilatifard A Multiple Roles for the MLL/COMPASS Family in the Epigenetic Regulation of Gene Expression and in Cancer. Annual Review of Cancer Biology 1, 425–446 (2017).
57. Dorighi KM et al. Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. Mol Cell 66, 568–576 e564 (2017). [PubMed: 28483418]

58. Ognjenovic NB et al. Limiting Self-Renewal of the Basal Compartment by PKA Activation Induces Differentiation and Alters the Evolution of Mammary Tumors. Dev Cell 55, 544–557.e546 (2020). [PubMed: 33120014]

Methods-only references

59. Sellappan S. et al. Lineage infidelity of MDA-MB-435 cells: expression of melanocyte proteins in a breast cancer cell line. Cancer Res 64, 3479–3485 (2004). [PubMed: 15150101]

60. Dontu G. et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev 17, 1253–1270 (2003). [PubMed: 12756227]

61. Cui J. et al. New use of an old drug: inhibition of breast cancer stem cells by benztropine mesylate. Oncotarget 8, 1007–1022 (2017). [PubMed: 27894093]

62. Levin-Allerhand JA, Sokol K & Smith JD Safe and effective method for chronic 17beta-estradiol administration to mice. Contemp Top Lab Anim Sci 42, 33–35 (2003). [PubMed: 14615958]

63. Dixon G. et al. QSER1 protects DNA methylation valleys from de novo methylation. Science 372 (2021).

64. Yang D. et al. CRISPR screening uncovers a central requirement for HHEX in pancreatic lineage commitment and plasticity restriction. Nat Cell Biol 24, 1064–1076 (2022). [PubMed: 35787684]

65. Lee JE et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife 2, e01503 (2013). [PubMed: 24368734]

66. Hong S. et al. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. Proc Natl Acad Sci U S A 104, 18439–18444 (2007). [PubMed: 18003914]

67. Wang C. et al. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. Proc Natl Acad Sci U S A 113, 11871–11876 (2016). [PubMed: 27698142]

68. Bolger AM, Lohse M & Usadel B Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014). [PubMed: 24695404]

69. Li H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009). [PubMed: 19505943]

70. Ramirez F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–165 (2016). [PubMed: 27079975]

71. Neph S. et al. BEDOPS: high-performance genomic feature operations. Bioinformatics 28, 1919–1920 (2012). [PubMed: 22576172]

72. Zhang Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137 (2008). [PubMed: 18798982]

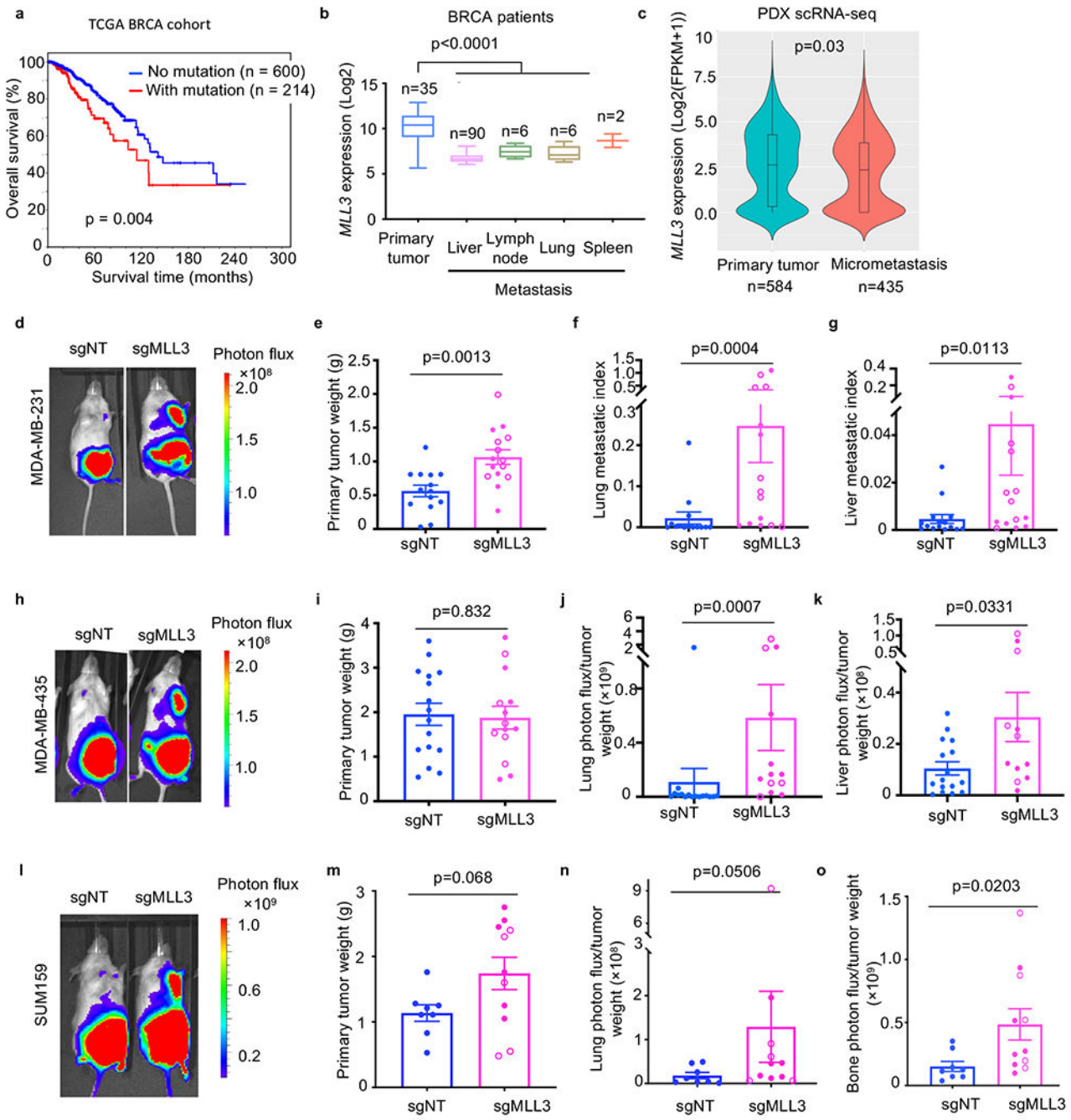


Fig 1. MLL3 deletion enhances breast tumor metastasis.

a, Overall survival of breast cancer patients with or without *MLL3* alterations (TCGA BRCA dataset). **b**, *MLL3* levels in metastases and primary breast tumors (GSE32490). **c**, *MLL3* levels in single tumor cells of primary tumors and the matching micrometastases in PDX models (GSE123837). **d**, Bioluminescence images of NOD/SCID mice 13 weeks after injection with indicated MDA-MB-231 cells. **e**, Weight of sgNT ($n = 14$) or sgMLL3 ($n = 15$) MDA-MB-231 primary tumors at week 13. **f** and **g**, Spontaneous lung (**f**) and liver (**g**) metastasis in sgNT ($n = 14$) or sgMLL3 ($n = 15$) MDA-MB-231 tumor-bearing

mice at week 13. Metastatic index = distant organ photon flux/primary tumor photon flux. **h**, Bioluminescence images of NOD/SCID mice at 12 weeks after injection with the indicated MDA-MB-435s. **i**, Primary tumor weight of sgNT (n = 16) or sgMLL3 (n = 14) MDA-MB-435s tumors at week 12. **J** and **k**, Spontaneous lung (**j**) and liver (**k**) metastasis in sgNT (n = 16) or sgMLL3 (n = 12) MDA-MB-435s tumor-bearing mice at week 12. Metastatic index = distant organ photon flux / tumor weight. **l**, Bioluminescence images of NOD/SCID mice at 9 weeks after injection with the indicated SUM159 cells. **m**, Primary tumor weight of sgNT (n = 8) or sgMLL3 (n = 11) SUM159 tumors at week 9. **n** and **o**, Spontaneous lung (**n**) and bone (**o**) metastasis in sgNT (n = 8) or sgMLL3 (n = 11) SUM159 tumor-bearing mice at week 9. Metastatic index = distant organ photon flux / tumor weight. All data are represented as mean± SEM. In **b-c**, minimum to maximum values are shown by the whiskers (**b**) or violin graph (**c**), the center line indicates the median, and the bounds of boxes represent the first and third quartiles. In **e-g**, **i-k**, and **m-n**, closed and open circles in the sgMLL3 groups represent two different sgRNAs. P values were determined by log-rank test (**a**), one-way ANOVA with Dunnett's test (**b**), or two-tailed Student's t-test (**c**, **e-g**; **i-k**; **m-o**). Source data are provided.

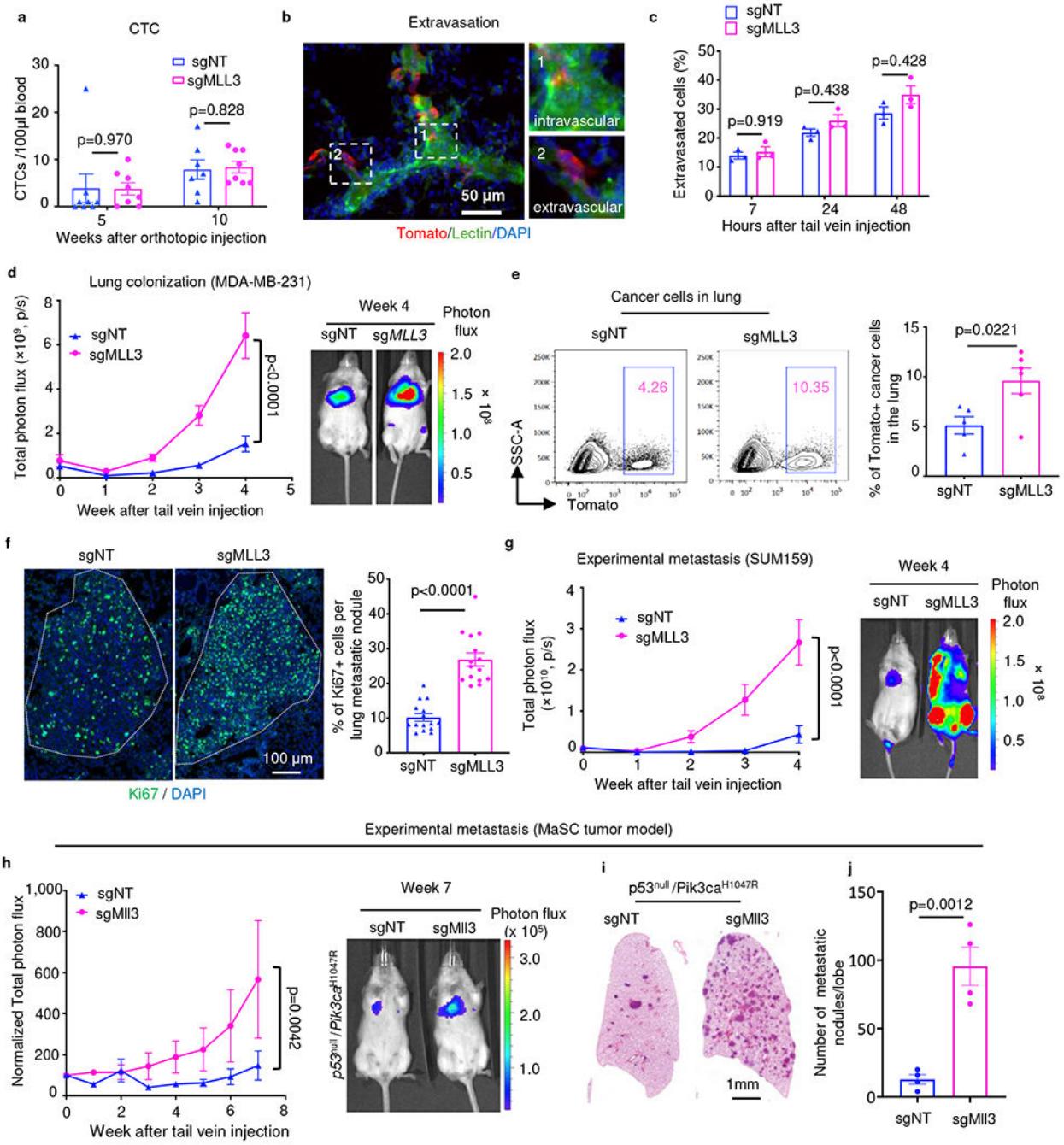


Fig 2. *MLL3* deletion promotes metastatic colonization.

a, Frequencies of circulating tumor cells (CTCs) in the peripheral blood of sgNT or sgMLL3 MDA-MB-231 tumor-bearing mice at 5 or 10 weeks post injection (Week 5: sgNT, $n = 8$; sgMLL3, $n = 8$; Week 10: sgNT, $n = 7$; sgMLL3, $n = 8$ mice). **b**, A representative image of lung cryosections showing examples of extravascular and intravascular MDA-MB-231 cells (Tomato^{+}) ($n = 3$ mice). **c**, Percentages of extravasated cancer cells were quantified on lung cryosections as shown in B ($n = 3$ mice per group). **d**, Lung colonization of sgNT and sgMLL3 (#3) MDA-MB-231 cells measured by the experimental metastasis assay.

500,000 Luc2-Tomato-labeled cells were injected via the tail vein (sgNT, $n = 6$; sgMLL3, $n = 5$ mice). **e**, Flow cytometric analysis of tumor cell frequencies in the lungs 14 days after tail vein injection with sgNT ($n = 5$ mice) or sgMLL3 ($n = 6$ mice) MDA-MB-231 cells. **f**, Representative images and quantification of $Ki67^{+}$ tumor cells in size-matched lung metastatic nodules formed by sgNT or sgMLL3 MDA-MB-231 cells 6 weeks after tail-vein injection ($n = 10$ nodules from 3 mice per group). **g**, Metastatic colonization of sgNT and sgMLL3 (#3) SUM159 cells measured by the experimental metastasis assay. 500,000 Luc2-Tomato-labeled cells were injected via the tail vein (sgNT, $n = 9$; sgMLL3, $n = 7$ mice). **h**, Lung colonization of sgNT and sgMll3 $p53^{null}/Pik3ca^{H1047R}$ cells measured by the experimental metastasis assay. 500,000 luciferase-labeled cells were injected into FVB mice via the tail vein ($n = 8$ mice per group). **i** and **j**, Representative H&E and quantification of metastatic nodules formed in left lung lobes as generated in **h** at the experimental endpoint ($n = 4$ mice per group). All data are represented as mean ± SEM. *P* values were determined by two-tailed Student's t-test (**a**, **e-f**, **j**), two-way ANOVA with correction using Geisser-Greenhouse method (**c-d**, **g-h**). Source data are provided.

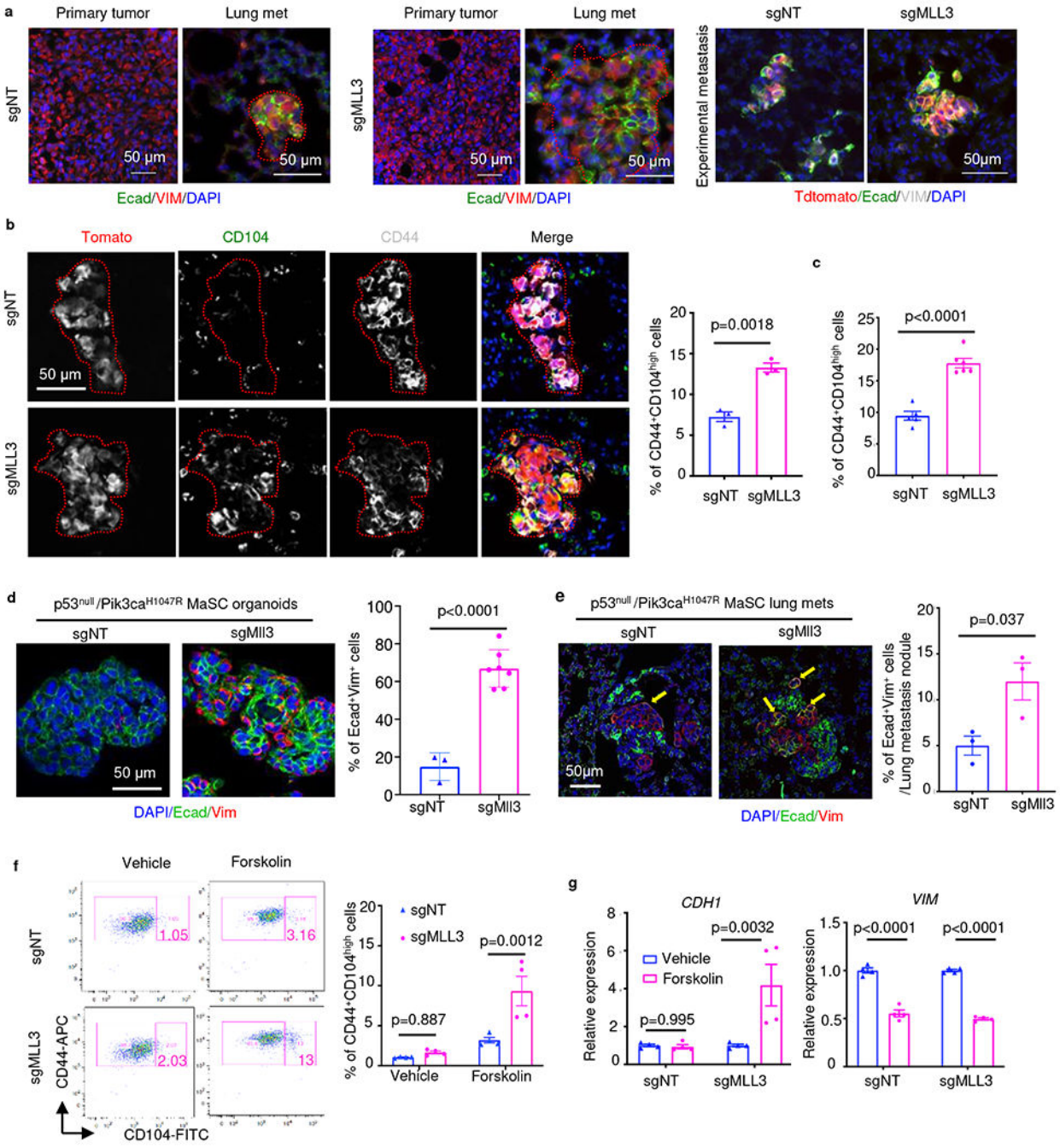


Fig 3. *MLL3* loss facilitates the acquisition of a hybrid EMT state.

a, Left and Middle: Immunofluorescence of vimentin (VIM) and E-cadherin (Ecad) in primary tumors and lung metastases formed by sgNT or sgMLL3 MDA-MB-231 cells. Right: IF of Ecad and Vim in experimental lung metastases formed by sgNT and sgMLL3 MDA-MB-231 cells (Tomato^{+}) at day 14 post tail vein injection. $n = 3$ primary tumors or lungs per group. b, Representative IF images and quantification of CD44^{+}CD104^{high} cells in lung metastases formed by sgNT or sgMLL3 MDA-MB-231 cells at day 14 post tail vein injection ($n = 3$ mice per group). **c**, Flow cytometry measuring the frequency of

$CD44^{+}CD104^{high}$ tumor cells in the lung 14 days after tail vein injection with sgNT ($n = 5$ mice) or sgMLL3 ($n = 6$ mice) MDA-MB-231 cells. **d**, Representative immunofluorescence and quantification of $Ecad^{+}VIM^{+}$ cells in sgNT ($n = 3$) and sgMll3 ($n = 7$) $p53^{null}$/$Pik3ca^{H1047R}$ MaSC organoids. **e**, Representative immunofluorescence and quantification of $Ecad^{+}VIM^{+}$ cells in experimental lung metastases derived from sgNT ($n = 3$ mice) and sgMll3 ($n = 3$ mice) $p53^{nul1}$/$Pik3ca^{H1047R}$ cells at week 8 post tail vein injection. **f**, The induction of $CD44^{+}CD104^{high}$ cells by the PKA activator forskolin in sgNT and sgMLL3 MDA-MB-231 cells. Cells were treated with10 μM forskolin for 14 days and then analyzed by flow cytometry ($n = 4$ biological samples examined over 2 independent experiments for each group). **g**, Expression levels of *CDH1* and *VIM* in the sgNT or sgMLL3 MDA-MB-231 cells. Cells were treated with 10 μM forskolin for 14 days and then analyzed by qRT-PCR ($n = 4$ biological samples examined over 2 independent experiments for each group). All data are represented as mean ± SEM. *P* values were determined by two-tailed Student's t-test (**b-e**), two-way ANOVA with Šídák's multiple comparisons test (**f-g**). Source data are provided.

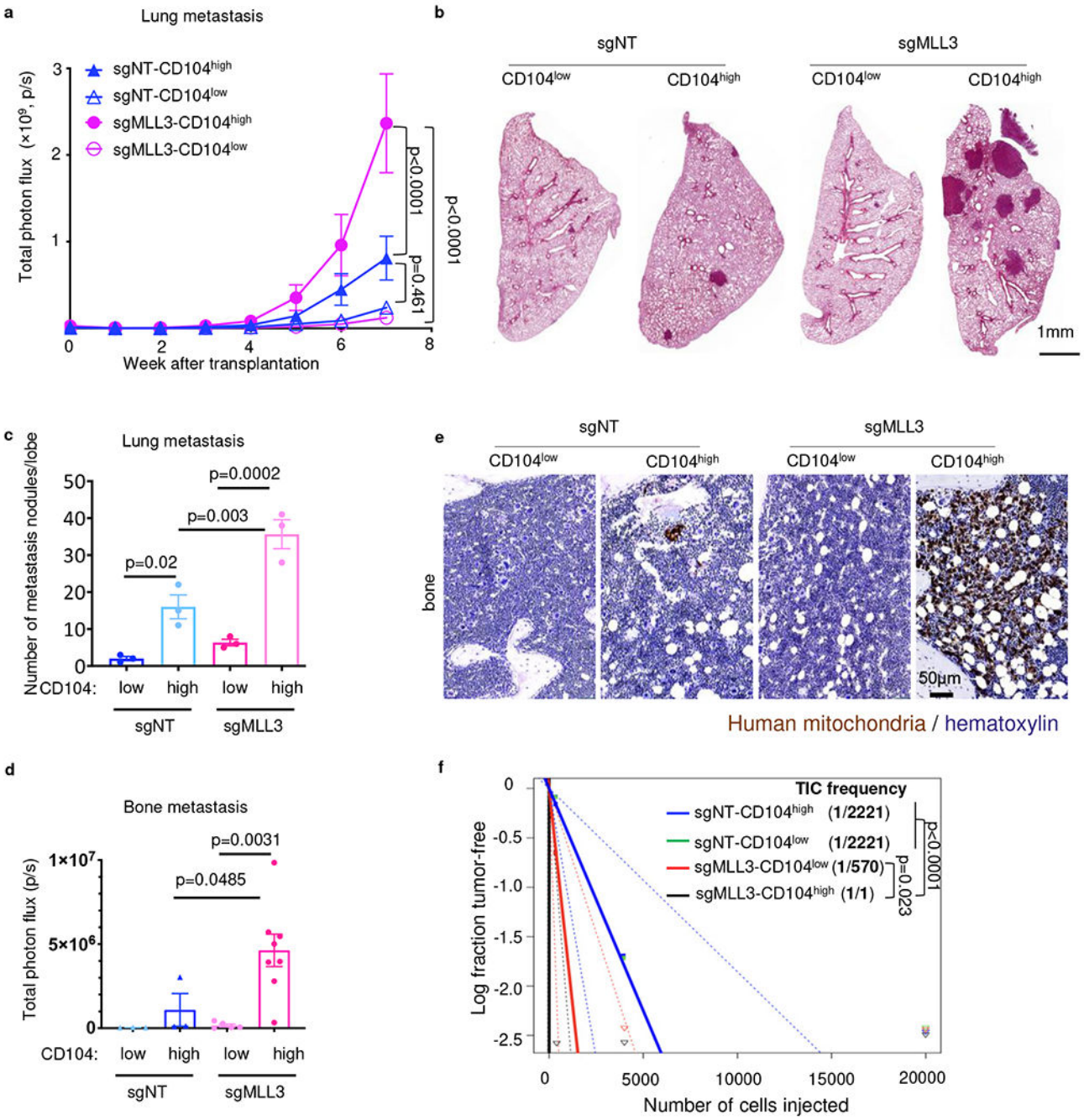


Fig 4. Loss of *MLL3* potentiates the multi-organ metastasis-initiating ability of hybrid E/M cells.

a, Lung metastasis-initiating ability of $CD44^{+}CD104^{low}$ or $CD44^{+}CD104^{high}$ cells sorted from 14-day forskolin-treated sgNT or sgMLL3 MDA-MB-231 cells. 2×10^4 cells were injected via tail vein and monitored for lung colonization by bioluminescence imaging (sgNT $CD44^{+}CD104^{low}$, $n = 3$; sgNT $CD44^{+}CD104^{high}$, $n = 3$; sgMLL3 $CD44^{+}CD104^{low}$, $n = 5$; sgMLL3 $CD44^{+}CD104^{high}$, $n = 8$ mice). **b**, Representative H&E staining of lung sections of mice at the experimental endpoint as shown in **a. c**, Number of metastatic nodules formed in left lung lobes as generated in **a** ($n = 3$ mice per group). **d**, *Ex vivo*

bioluminescence signals in bones (hind limbs) at the experimental endpoint as shown in **a** (sgNT $CD44^{+}CD104^{low}$, $n = 3$; sgNT $CD44^{+}CD104^{high}$, $n = 3$; sgMLL3 $CD44^{+}CD104^{low}$, $n = 5$; sgMLL3 $CD44^{+}CD104^{high}$, $n = 8$ mice). **e**, Human mitochondria IHC detecting MDA-MB-231 metastases in bones at the endpoint as shown in **a** (sgNT $CD44^{+}CD104^{low}$, $n = 3$; sgNT $CD44^{+}CD104^{high}$, $n = 3$; sgMLL3 $CD44^{+}CD104^{low}$, $n = 5$; sgMLL3 $CD44^{+}CD104^{high}$, $n = 8$ mice). **f**, Limiting dilution transplantation assay measuring tumor-initiating cell (TIC) frequency of $CD104^{high}/CD44^{+}$ and $CD104^{low}/CD44^{+}$ populations of WT (sgNT) or *MLL3*-mutant (sgMLL3) MDA-MB-231 cells. TIC frequencies were calculated with the ELDA program (Supplementary Table 1). All data are represented as mean ± SEM. *P* values were determined by two-way ANOVA with Šídák's multiple comparisons test (**a, c**), two-tailed Student's t-test (**d**) or chi-square likelihood ratio test (**f**). Source data are provided.

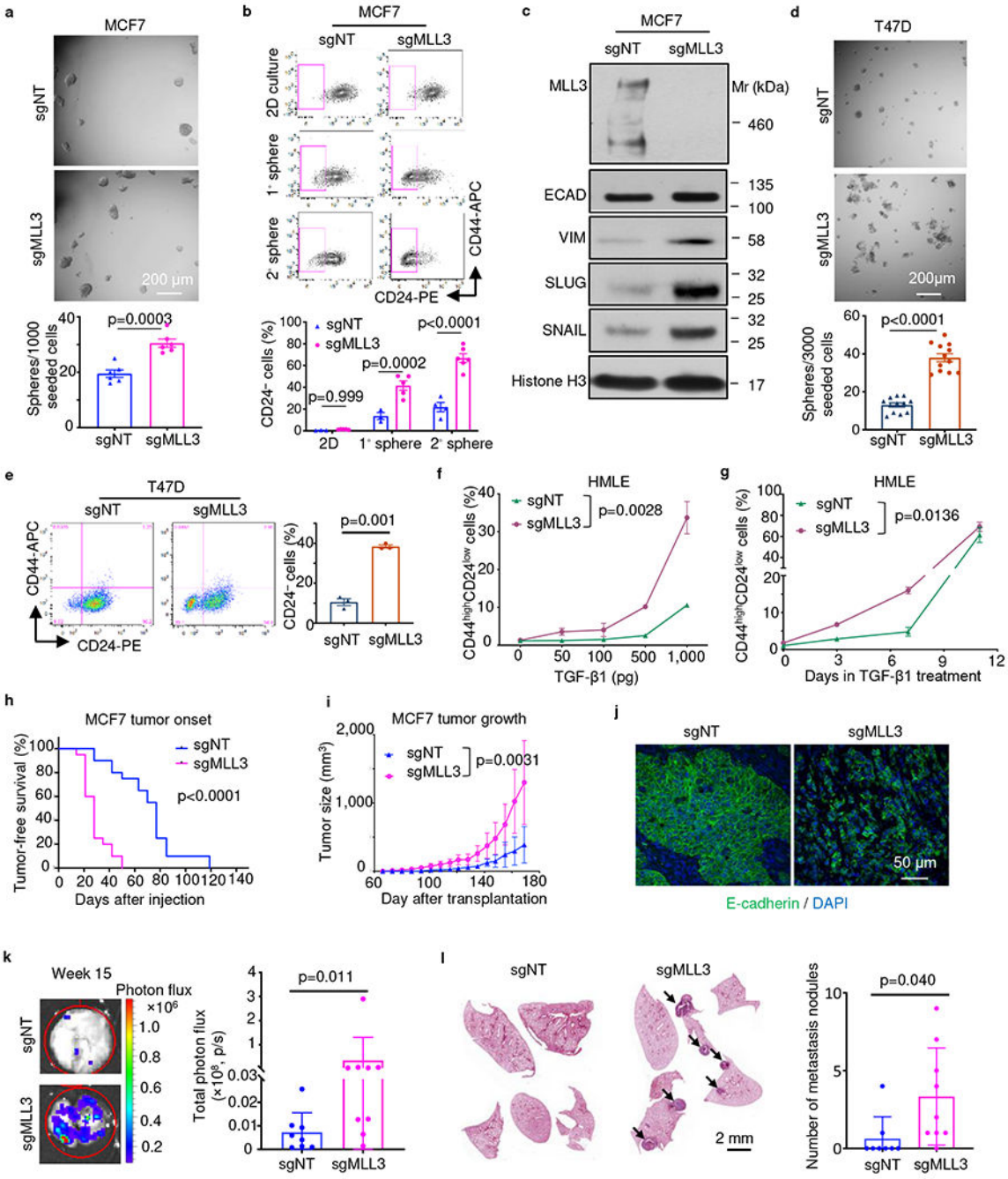


Fig 5. *MLL3* loss enables epithelial cells to gain mesenchymal features.

a, Sphere-forming efficiency of sgNT or sgMLL3 MCF7 cells ($n = 3$ biological samples in each experiment, two experiments were performed). **b**, Flow cytometry for CD24 in sgNT or sgMLL3 MCF7 spheres (sgNT: 2D, $n = 3$; 1° sphere, $n = 3$; 2° sphere, $n = 4$; sgMLL3: 2D, $n = 5$; 1° sphere, $n = 5$; 2° sphere, $n = 6$ biological replicates). **c**, Western blot for MLL3, E-cadherin (ECAD), Vimentin (VIM), SLUG and SNAIL in sgNT or sgMLL3 MCF7 spheres (Similar results was repeated with 2 biological samples). **d**, Sphere-forming efficiency of sgNT or sgMLL3 T47D cells ($n = 2$ biological samples with six experimental replicates). **e**,

Flow cytometry for CD24 in sgNT or sgMLL3 T47D spheres (sgNT, $n = 3$; sgMLL3, $n = 3$ biological samples). **f**, Induction of $CD44^{high}CD24^{low}$ cells in sgNT or sgMLL3 HMLE cells treated with different doses of TGF-β1 for 11 days (sgNT, $n = 5$; sgMLL3, $n = 5$ biological samples). **g**, Kinetics of $CD44^{high}CD24^{low}$ cell induction by TGF-β1 (5 ng/ml) in sgNT or sgMLL3 HMLE cells (sgNT, $n = 3$; sgMLL3, $n = 3$ biological samples). **h**, Kaplan-Meier survival analysis of tumor onset in mice orthotopically injected with sgNT or sgMLL3 MCF7 cells ($n = 20$ xenografts per group). **i**, Growth rates of sgNT ($n = 5$) or sgMLL3 ($n = 8$) MCF7 xenografts. **j**, Representative images of E-cadherin expression in sgNT or sgMLL3 MCF7 tumors ($n = 10$ tumors per group). **k**, Lung colonization of sgNT and sgMLL3 MCF7 measured by the experimental metastasis assay. 200,000 Luc2-Tomato-labeled cells were injected via the tail vein. Lung metastatic burdens were measured by at week 15 (sgNT, $n = 8$; sgMLL3, $n = 9$ mice). **I**, Representative H&E and number of lung metastatic nodules per mouse at the experimental endpoint as shown in **k**. All data are represented as mean ± SEM. *P* values were determined by two-tailed Student's t-test (**a**, **d-e**, **k-l**), two-way ANOVA with Šídák's multiple comparisons test (**b, f-g, i**), or log-rank test (**h**). Source data are provided.

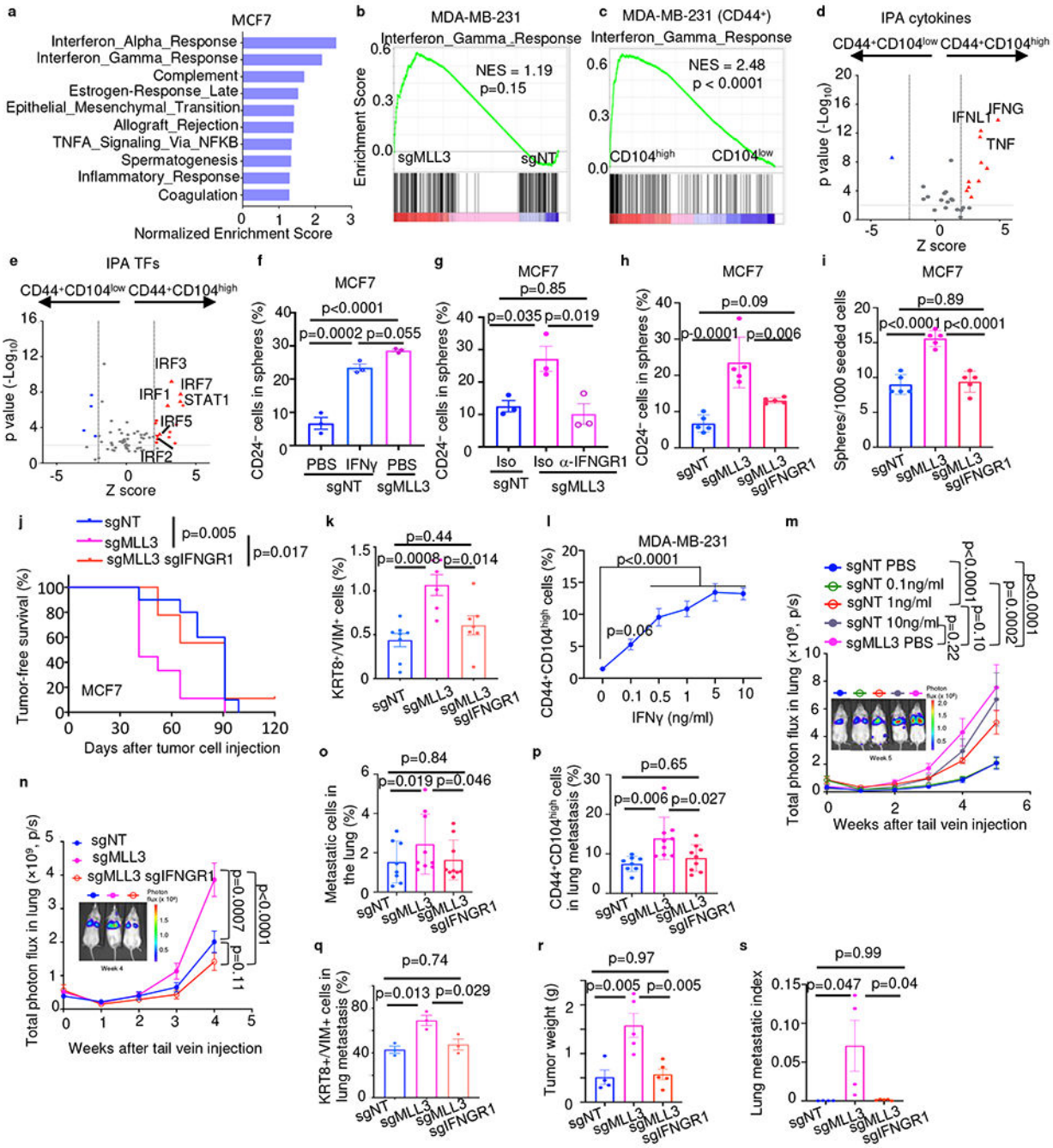


Fig 6. *MLL3* loss enhances tumor progression and metastasis through activation of the IFNγ pathway.

a, Top 10 enriched Hallmark signatures in sgMLL3 MCF7 cells. **b** and **c**, GSEA comparing IFNγ signature in the indicated cells. **d** and **e**, Upstream cytokines (**d**) and transcription factors (**e**) enriched in the $CD44^{+}CD104^{high}$ MDA-MB-231 cells. **f**, Increase of $CD24^{-}$ cells by IFNγ (10ng/ml, 3 days) in MCF7 spheres ($n = 3$ biological samples). **g**, Reduction of $CD24^{-}$ cells by anti-IFNGR1 (1μg/ml, 6 days) ($n = 3$ biological samples). **h**, Frequency of $CD24^{-}$ cells in the indicated MCF7 spheres ($n = 5$ technique replicates over 3 independent

experiments). **i**, Sphere-forming ability of the indicated MCF7 cells ($n = 5$ technique replicates, similar results were repeated three times). **j**, Tumor onset in mice injected with sgNT ($n = 10$), sgMLL3 ($n = 9$) and sgMLL3/sgIFNGR1 MCF7 cells ($n = 9$ xenografts). **k**, Quantification of $KRT8^+VIM^+$ cells in sgNT ($n = 8$), sgMLL3 ($n = 6$) and sgMLL3/sgIFNGR1 MCF7 tumors ($n = 7$ tumors). **l**, $CD44^+CD104^{high}$ cell frequency in cells treated as indicated for 4 days ($n = 6$ technique replicates over 2 independent experiments). **m**, Lung colonization of MDA-MB-231 cells pretreated with IFNγ as indicated for 4 days in culture, then injected via tail vein into recipient mice (sgNT-PBS, $n = 7$; sgNT-0.1ng/ml, $n = 4$; sgNT-1ng/ml, $n = 3$; sgNT-10ng/ml, $n = 5$; sgMLL3-PBS, $n = 6$ mice). **n**, Lung colonization by sgNT ($n = 6$), sgMLL3 ($n = 6$) or sgMLL3/sgIFNGR1 ($n = 7$ mice) MDA-MB-231 cells. **o** and **p**, Frequency of $Tomato^+$ tumor cells (**o**) and $CD44^+CD104^{high}$ cells in tumor cells (**p**) in the lungs 14 days after tail-vein injection with sgNT ($n = 8$), sgMLL3 ($n = 9$) or sgMLL3/sgIFNGR1 MDA-MB-231 cells ($n = 9$ mice). **q**, Frequency of $KRT8^+/VIM^+$ cells in tumor cells in the lung 14 days after tail-vein injection with sgNT, sgMLL3 or sgMLL3/sgIFNGR1 MDA-MB-231 cells ($n = 3$ per group). **r**, Weight of sgNT ($n = 4$), sgMLL3 ($n = 4$) or sgMLL3/sgIFNGR1 MDA-MB-231 primary tumors ($n = 5$) at week 13 after injection. **s**, Lung metastasis in sgNT ($n = 4$), sgMLL3 ($n = 4$) or sgMLL3/sgIFNGR1 MDA-MB-231 tumor-bearing mice ($n = 5$) as generated in **r**. Metastatic index = lung photon flux/primary tumor photon flux. All data are represented as mean ± SEM. *P* values were determined by permutation-based test using GSEA (**b-c**), Fisher's right tailed exact test (**d-e**), one-way ANOVA with Turkey's multiple comparisons test (**f-i, k, o-s**) or Dunnett's multiple comparisons test (**l**), log-rank test (**j**), two-way ANOVA with Turkey's multiple comparisons test (**m-n**). Source data are provided.

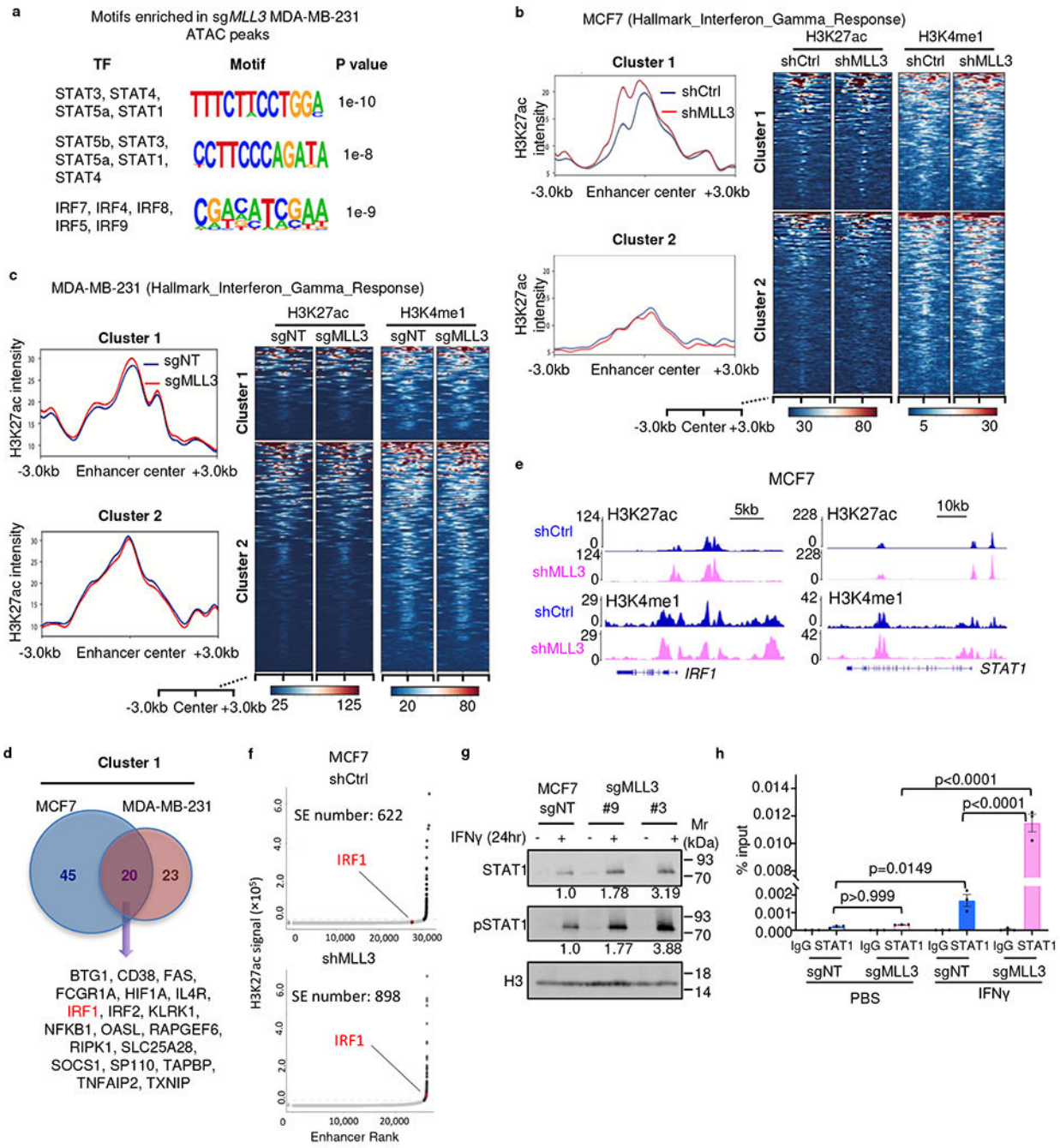


Fig 7. *MLL3* loss increased H3K27ac levels in IFNγ response gene enhancers.

a, Transcription factor motifs enriched in the ATAC peaks of the sgMLL3 MDA-MB-231 cells ($CD104^{high}$ or $CD104^{low}$) compared to the sgNT control counterparts. **b**, H3K27ac signal intensity at enhancer regions of IFNγ response genes in shCtrl and shMLL3 MCF7 cells. The data was generated by reanalyzing the ChIP-seq data from GSE100328. IFNγ response genes are grouped into two clusters according to their H3K27ac intensity at enhancer regions corresponding to individual genes. The regions with increased H3K27ac intensity in sgMLL3 cells were classified as Cluster 1, otherwise as Cluster 2. The gene list

and relevant H3K27ac intensity can be found in Source data. **c**, H3K27ac ChIP-seq signal intensity at enhancer regions of IFNγ response genes in sgNT and sgMLL3 MDA-MB-231 cells. The regions with increased H3K27ac intensity in sgMLL3 cells were classified as Cluster 1, otherwise as Cluster 2. The gene list and relevant H3K27ac intensity can be found in Source data. **d**, Venn diagram showing the overlap of Cluster 1 IFNγ response genes between MCF7 and MDA-MB-231 cells. **e**, H3K27ac and H3K4me1 ChIP-seq peaks at *IRF1* and *STAT1* loci in sgNT and sgMLL3 MCF7 cells. **f**, Hockey-stick plots generated by ROSE showing ranked enhancers based on H3K27ac ChIP-seq in shCtrl and shMLL3 MCF7 cells (SE, super enhancer). The data was generated by reanalyzing the ChIP-seq data from GSE100328. **g**, Western blot analysis of pSTAT1 and STAT1 in sgNT and sgMLL3 MCF7 cells treated with PBS or 100pg/ml IFNγ for 24h ($n = 3$ independent experiments with two different sgRNA). **h**, ChIP-qPCR analysis of STAT1 occupancy at the *STAT1* promoter region in sgNT and sgMLL3 MCF7 cells treated with PBS or 10 ng/ml IFNγ for 4h. ChIP samples were analyzed by qPCR and normalized to input. One of three independent experiments was shown ($n = 3$). Data are represented as mean ± SEM. *P* values for motif enrichment were calculated by HOMER (**a**). *P* values in (**h**) were determined by one-way ANOVA with Turkey's multiple comparisons test. Source data are provided.

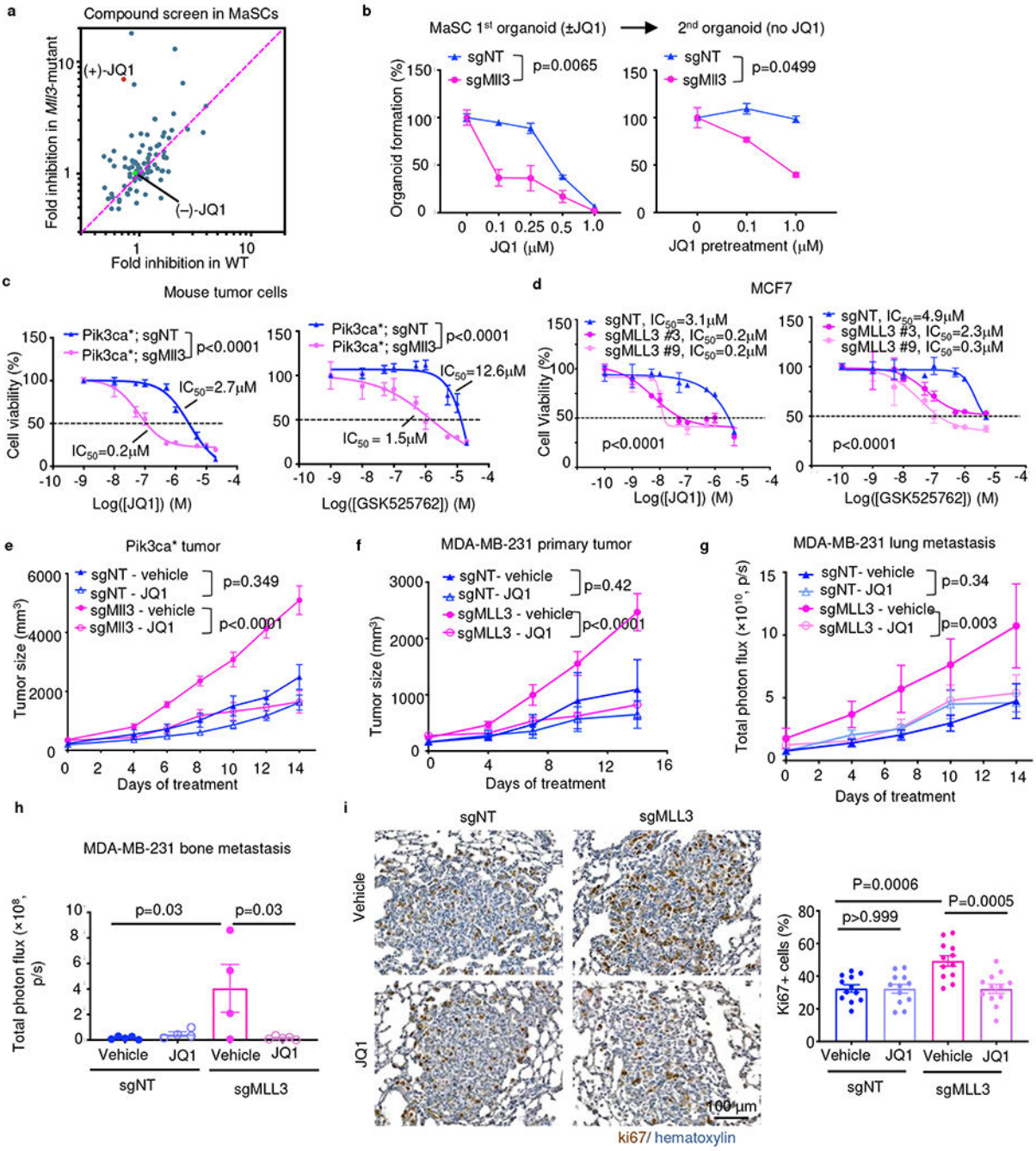


Fig 8. *MLL3*-mutant cells are sensitive to BET inhibition.

a, Chemical screen identifying epigenetic compounds selectively targeting *Mll3*-mutant MaSCs. **b**, Effect of JQ1 on organoid replating efficiency of sgNT or sgMll3 MaSCs. Primary organoids were treated with different doses of JQ1 and then assayed for secondary organoid formation without the drug treatment ($n = 3$ technique replicates for each experiment, a similar result was repeated with another biological sample). **c**, Dose-response of sgNT or sgMll3 primary Pik3ca* tumor cells to BET inhibitors ($n = 6$ technique replicates for each experiment, a similar result was repeated with another independent experiment).

d, Dose-response of sgNT or sgMLL3 MCF7 cells to BET inhibitors ($n = 3$ technique replicates for each experiment, a similar result was repeated with another independent experiment). **e**, Effect of JQ1 on sgNT or sgMll3 Pik3ca* tumor growth (sgNT Vehicle, $n = 7$; sgNT JQ1, $n = 7$; sgMll3 Vehicle: $n = 5$; sgMll3 JQ1, $n = 6$ tumors). **f**, Effect of JQ1 on sgNT or sgMLL3 MDA-MB-231 primary tumor growth (sgNT Vehicle, $n = 5$; sgNT JQ1, $n = 7$; sgMLL3 Vehicle, $n = 5$; sgMLL3 JQ1, $n = 5$ tumor). **g**, Effect of JQ1 on sgNT or sgMLL3 MDA-MB-231 lung metastasis growth, as measured by longitudinal luminescence imaging (sgNT Vehicle: $n = 5$; sgNT JQ1: $n = 4$; sgMLL3 Vehicle: $n = 4$; sgMLL3 JQ1: $n = 6$ mice). **h**, Effect of JQ1 on sgNT or sgMLL3 MDA-MB-231 bone metastasis, as measured by *ex vivo* bioluminescence (sgNT Vehicle: $n = 5$; sgNT JQ1: $n = 4$; sgMLL3 Vehicle: $n = 4$; sgMLL3 JQ1: $n = 5$ mice). **i**, Frequency of $Ki67^+$ cells in sgNT or sgMLL3 MDA-MB-231 lung metastases in animals treated with JQ1 or vehicle for 2 weeks ($n = 12$ each group). All data are represented as mean ± SEM. *P* values were determined by two-way RM ANOVA with correction using Geisser-Greenhouse method (**b-d**), Ordinary two-way ANOVA (**e-g**) or one-way ANOVA with Turkey's multiple comparisons test (**h-i**). Source data are provided.