CANCER

ZBTB18 restricts chromatin accessibility and prevents transcriptional adaptations that drive metastasis

Ruishan Wang[1], Akshita B. Bhatt[1,2], Benjamin A. Minden-Birkenmaier[1,2], Olivia K. Travis[1,2], Srishti Tiwari[1,2], Hong Jia[1], Wojciech Rosikiewicz[3], Ophelie Martinot[1], Eleanor Childs[1]†, Robin Loesch[1], Guenole Tossou[1], Sophie Jamieson[2], David Finkelstein[4], Beisi Xu[3], Myriam Labelle[1,2]*

Metastases arise from rare cancer cells that successfully adapt to the diverse microenvironments encountered during dissemination through the bloodstream and colonization of distant tissues. How cancer cells acquire the ability to appropriately respond to microenvironmental stimuli remains largely unexplored. Here, we report an epigenetic pliancy mechanism that allows cancer cells to successfully metastasize. We find that a decline in the activity of the transcriptional repressor ZBTB18 defines metastasis-competent cancer cells in mouse models. Restoration of ZBTB18 activity reduces chromatin accessibility at the promoters of genes that drive metastasis, such as *Tgfbr2*, and this prevents TGFβ1 pathway activation and consequently reduces cell migration and invasion. Besides repressing the expression of metastatic genes, ZBTB18 also induces widespread chromatin closing, a global epigenetic adaptation previously linked to reduced phenotypic flexibility. Thus, ZBTB18 is a potent chromatin regulator, and the loss of its activity enhances chromatin accessibility and transcriptional adaptations that promote the phenotypic changes required for metastasis.

INTRODUCTION

Metastasis is the leading cause of cancer-related mortality, but the mechanisms responsible for cancer cell dissemination and disease recurrence are incompletely understood. Metastasis occurs when malignant cancer cells succeed in completing all the steps of the metastatic cascade: invasion of the tissues surrounding the primary tumor, transit through the bloodstream, arrest and survival in distant organs, and metastatic outgrowth (*1*, *2*). Despite the aggressive nature of metastatic cancers, only very few of the cancer cells shed by a primary tumor ultimately form clinically relevant metastases in distant tissues (*1*–*3*). This is due to the inability of the vast majority of cancer cells to survive the immunological challenges and mechanical stresses associated with systemic dissemination through the circulation and the hurdles of invading, surviving, and growing into the parenchyma of distant organs (*1*–*3*).

Because the phenotypic traits that facilitate a specific step of the metastatic cascade (e.g., invasion) differ from those needed for other steps (e.g., proliferation), classical models postulate that overt metastases originate from rare tumor cells that initially had or acquired en route a high degree of phenotypic plasticity (*2*, *4*). In particular, it has been proposed that highly metastatic tumor cells respond to microenvironmental signals [such as WNT (Wingless/Integrated) proteins, transforming growth factor–β (TGFβ) ligands, and inflammatory cytokines] and transiently acquire stem cell–like properties associated with increased phenotypic flexibility (*2*, *4*–*6*). Furthermore, some highly metastatic tumor cells are capable of shuttling between opposite cell differentiation states (e.g., epithelial versus mesenchymal phenotypes), and this adaptability confers them the capacity to proliferate or migrate depending on the metastatic step and associated microenvironment (*2*, *4*, *6*). However, it remains poorly defined how tumor cells achieve the state of pliancy that allows them to respond to diverse microenvironmental cues and to surmount the challenges encountered during metastatic dissemination.

Epigenetic alterations play key roles in enabling cell plasticity during cell differentiation and tumorigenesis. Alterations such as aberrant promoter and enhancer activity, as well as widespread changes in heterochromatin have also been implicated in metastasis of various cancer types (*7*–*14*). These changes are characterized by DNA hypomethylation and by the gain of activating and loss of repressive histone marks, which can result from a reduction in PRC2 (polycomb repressive complex 2)-mediated gene silencing (*9*, *12*). For example, in breast cancer cells, the presence of a poised (bivalent) chromatin conformation at the promoter of the transcription factor ZEB1 (zinc finger E-box binding homeobox 1) enables tumor cells to respond to TGFβ1 and to transition to a stem cell–like phenotype (*15*).

Compared to differentiated cells, undifferentiated stem cells display a more open chromatin state that facilitates transcriptional responses to environmental stimuli and promotes phenotypic flexibility (*16*). Enhanced chromatin accessibility at the promoter of specific transcription factors has also been associated with metastasis (*8*). Moreover, recent studies have indicated that global reorganization of chromatin conformation contributes to small cell lung cancer (SCLC) metastasis (*17*). In particular, *Nfib* copy number amplification drives a permanent widespread increase in chromatin accessibility that facilitates metastasis in a subset of SCLCs (*17*). However, it remains unknown whether similar changes in chromatin accessibility also drive metastasis in other cancer types. Furthermore, the factors responsible for restricting chromatin accessibility

[1]Comprehensive Cancer Center, Solid Tumor Program, Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [2]Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [3]Center for Applied Bioinformatics, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [4]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA.
†Present address: The Francis Crick Institute, London, NW1 1AT, UK.
*Corresponding author. Email: myriam.labelle@stjude.org

and the mechanisms by which they might be overridden during metastasis remain incompletely understood.

By using an orthotopic breast cancer model that recapitulates the entire metastatic cascade in immunocompetent mice, here, we show that loss of activity of the transcriptional repressor ZBTB18 (zinc finger and BTB domain containing 18; RP58/ZFN238) characterizes metastasis-derived tumor cells that retain a higher ability to metastasize. Restoration of ZBTB18 activity via its overexpression leads to a reduction in chromatin accessibility at the promoters of genes associated with metastasis. In particular, we find that ZBTB18 overexpression considerably limits chromatin accessibility at the Tgfb receptor 2 (*Tgfbr2*) promoter, thereby preventing TGFβ1 signaling and consequently inhibiting downstream cell motility and metastasis.

In addition to repressing target genes associated with metastasis, we find that ZBTB18 overexpression also induces widespread chromatin closing. Thus, we propose that loss of ZBTB18 activity enhances global chromatin accessibility in tumor cells and that this facilitates transcriptional adaptations that promote the phenotypic changes required for metastasis.

RESULTS

Breast cancer cells isolated from distinct lung metastases that originated from the same primary tumor display drastically different metastatic competence

Comparison of the gene expression profiles of cancer cells isolated from metastases and from their matched primary tumors is an approach that has been successfully used to identify candidate drivers of metastasis (*18*, *19*). However, only few such pairs of metastases and parental primary tumors have been developed in immunocompetent mouse models, hindering the possibility to identify drivers of metastasis in the physiological context of an intact immune system.

To address this gap and establish new models of tumor-metastasis pairs in immunocompetent mice, we have generated novel breast cancer cell lines from distinct lung metastases that originated from the same primary tumor. To this end, poorly metastatic E0771 cells [luminal B ($ERα^-$, $ERβ^+$, PR^+, and $ErbB2^+$) breast cancer cell line (*20*) derived from a spontaneous mammary adenocarcinoma that arose in a C57BL/6 mouse (*21*)] were stably transduced with copGFP (copepod green florescent protein) (E0771GFP cells) and implanted into the fourth mammary fat pad of a C57BL/6 female. One month later, two distinct spontaneous lung metastases were isolated, expanded in culture, and fluorescence-activated cell sorting (FACS)–sorted based on green fluorescent protein (GFP) expression, yielding the M11GFP and M12GFP cell lines (Fig. 1A). When reimplanted orthotopically into mice, these two cell lines gave rise to tumors that grew significantly faster than parental E0771GFP tumors, as indicated by higher tumor masses 28 days after tumor implantation (Fig. 1B). Furthermore, quantitation of lung metastatic foci revealed that mice bearing M11GFP primary tumors developed significantly more metastases than mice inoculated with either parental E0771GFP or M12GFP cells (Fig. 1C). Thus, although M12GFP cells give rise to bigger tumors than E0771GFP or M11GFP cells, they are not more metastatic than parental E0771GFP cells. In contrast, M11GFP cells are more metastatic than both E0771GFP and M12GFP cells (Fig. 1C), even when the number of metastases in each mouse is normalized by the primary tumor mass (Fig. 1D). The increased ability to seed lung metastases of M11GFP in comparison to E0771GFP or M12GFP cells was confirmed in an experimental metastasis model based on the tail vein injection of cancer cells: M11GFP cells yielded approximately four times more lung metastases than M12GFP or E0771GFP cells (Fig. 1E). Consistent with their enhanced metastatic competence, M11GFP cells migrated more efficiently in vitro than E0771GFP or M12GFP cells (Fig. 1F), while all three cell lines displayed similar proliferation rates (Fig. 1G). Together, via in vivo selection from a parental cell line, we have here generated a novel breast cancer cell line with increased metastatic potential (M11GFP cells) and another with enhanced in vivo tumor growth (M12GFP) but limited metastatic potential. Together with their common parental line, these two cell lines provide an attractive model system to delineate gene expression/phenotypic programs that are differentially modulated in metastatic tumor cells with highversuslow ability to further metastasize.

Target genes of the transcriptional repressor ZBTB18 are up-regulated in breast cancer cells with high metastatic competence

RNA sequencing (RNA-seq) was performed to delineate the gene expression changes that underlie the different metastatic behavior of E0771GFP, M11GFP, and M12GFP cells. Principal components analysis (PCA) indicated that these three related cell lines each display distinct gene expression programs (principal component 1), with principal component 2 demonstrating the difference in gene expression between highly metastatic M11GFP cells and poorly metastatic M12GFP and E0771GFP cells (Fig. 2A and fig. S1A). Furthermore, the expression of 242 genes is significantly modulated [$|\log_2(\text{fold change})| > 0.5$, $P < 0.05$] in M11GFP cells in comparison to both E0771GFP and M12GFP cells (Fig. 2B and fig. S1B), suggesting that these genes might play a role in metastasis.

To pinpoint potential signaling pathways globally driving these gene expression changes and responsible for the enhanced metastatic potential of M11GFP cells, we examined whether genes up-regulated in M11GFP cells share any transcription factor binding sites in their promoter. Gene set enrichment analysis (GSEA) using Molecular Signatures Database (MSigDB) transcription factor target prediction (C3 TFT gene sets) identified 13 transcription factors associated with target genes up-regulated in M11GFP cells when compared to related but less metastatic M12GFP and E0771GFP cells (Fig. 2C and table S1). Because M12GFP have a similar metastatic behavior as E0771GFP cells (Fig. 1), any transcription factor up-regulated in M12GFP versus E0771GFP was excluded from the list of putative transcription factors that drive metastasis (Fig. 2C).

Among differentially regulated candidates, target genes of the transcriptional repressor ZBTB18 (RP58/ZNF238) were the most significantly up-regulated in M11GFP in comparison to the two other related cell lines with lower metastatic potential (Fig. 2D and table S1). ZBTB18 is a transcriptional repressor that plays a crucial role in brain development, neuronal differentiation, and myogenesis (*22*–*25*). In mouse gliomas and colon cancer, loss of *Zbtb18* expression via epigenetic silencing has been implicated in tumor progression, and restoration of its expression has been shown to limit tumor growth (*24*, *26*, *27*). However, whether ZBTB18 plays a functional role in cancer metastasis remains unknown.

On the basis of the observed increase in ZBTB18 target gene expression in highly metastatic cells, we hypothesized that loss of

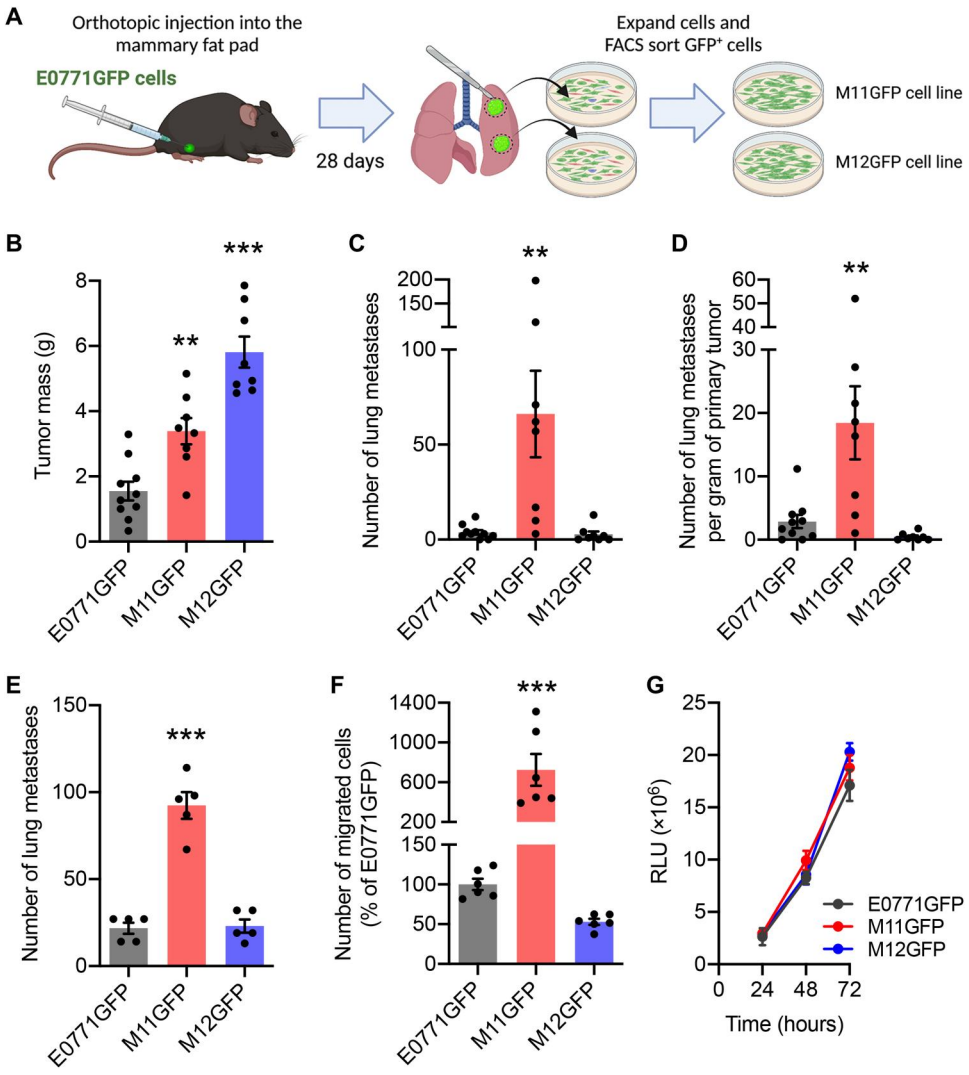

Fig. 1. Breast cancer cells isolated from distinct lung metastases that originated from the same primary tumor display drastically different metastatic competence. (**A**) Scheme of the experimental approach used to generate the M11GFP and M12GFP cell lines. (**B**) Primary tumor mass, 28 days after orthotopic inoculation of 10^6 E0771GFP, M11GFP, and M12GFP cells into the fourth mammary fat pad of C57BL/6 female mice (E0771GFP, $n = 10$ mice; M11GFP and M12GFP, $n = 8$ mice per group). (**C**) Number of spontaneous lung metastases in mice from (B). (**D**) Number of spontaneous lung metastases normalized by tumor mass in mice from (B). (**E**) Number of lung metastases in C57BL/6 female mice 14 days after tail vein injection of 5×10^5 E0771GFP, M11GFP, or M12GFP cells ($n = 5$ mice per group). (**F**) Migration of E0771GFP, M11GFP, and M12GFP cells through Transwell inserts ($n = 6$ biological replicates). (**G**) Proliferation rate of E0771GFP, M11GFP, and M12GFP cells ($n = 3$ independent experiments done in quadruplicate). RLU, relative luminescence units. (B to G) Means ± SEM, one-way ANOVA, followed by Tukey's posttest. $^{**}P < 0.01$ and $^{***}P < 0.001$.

ZBTB18 repressor activity enhances metastasis. ZBTB18 inhibition could occur via a reduction in ZBTB18 expression levels and/or activity. However, no changes were detected in ZBTB18 expression at both the mRNA and protein levels in M11GFP versus E0771GFP cells (Fig. 2, E and F). Furthermore, there are no mutations in the protein coding sequence of *Zbtb18* in these cell lines (table S2). Together, these results thus suggest that ZBTB18 activity (rather than its expression) is repressed in M11GFP cells. Similarly, no difference in ZBTB18 expression was found when comparing the highly metastatic 4T1 breast cancer cell line to the poorly metastatic 67NR cell line (Fig. 2, E and F), two well-studied cell lines isolated from the same spontaneous breast tumor that developed in a BALB/c mouse (*28*). However, increased expression of some of the most strongly enriched ZBTB18 target genes (*Tgfbr2*, *Ntn1*, *Nuak1*, and *Ltbp1*) was seen when comparing M11GFP to E0771GFP and 4T1 to 67NR cells (Fig. 2, G and H). Together, these data suggest that the activity of the transcriptional repressor ZBTB18 is inhibited in M11GFP and 4T1 cells, thereby allowing expression of its target genes in these highly metastatic cells.

A well-known mechanism that modulates the activity of certain transcription factors consists in their cytoplasmic-to-nuclear shuttling, which results in their inactivation when retained in the cytoplasm and their activation upon translocation to the nucleus (*29*). On this basis, we next tested whether cytoplasmic-to-nuclear shuttling differentially regulates ZBTB18 activity in cancer cells with distinct metastatic potential. Western blot analysis of nuclear fractions and whole-cell lysates (Fig. 2I) and immunofluorescence staining (Fig. 2J and fig. S2A) revealed that a lower proportion of ZBTB18

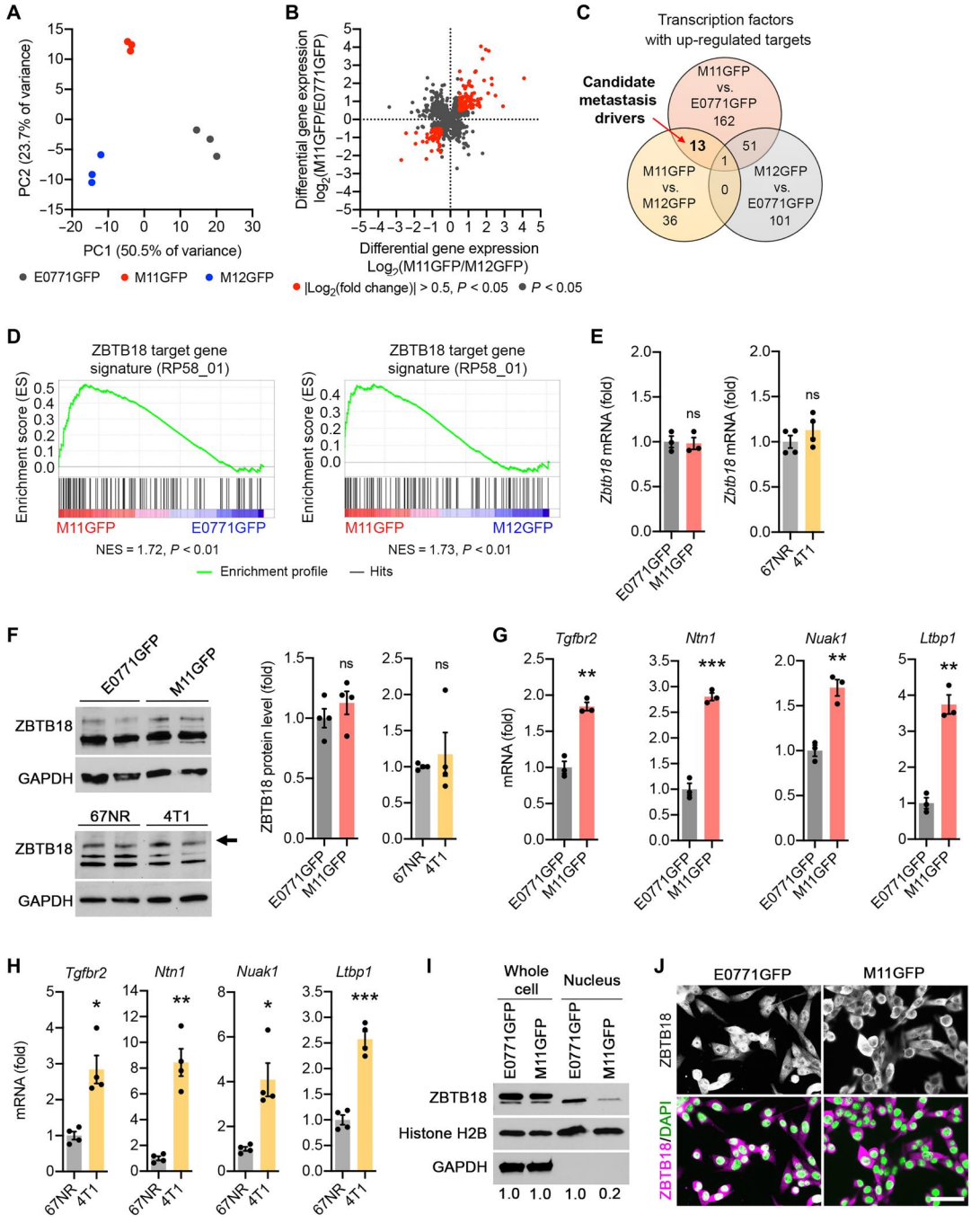

Fig. 2. Target genes of the transcriptional repressor ZBTB18 are up-regulated in breast cancer cells with high metastatic competence. (**A**) PCA of RNA-seq data indicates that E0771GFP, M11GFP, and M12GFP have distinct gene expression patterns ($n = 3$ biological replicates per cell line). (**B**) Differential gene expression for M11GFP versus M12GFP and M11GFP versus E0771GFP. Only genes with significant differential expression ($P < 0.05$) for both comparisons are shown. (**C**) Numbers of transcription factors with increased target gene expression ($P < 0.05$) when comparing RNA-seq data obtained from E0771GFP, M11GFP, and M12GFP cell lines, as determined using GSEA analysis with the MSigDB C3 TFT (transcription factor target) collection. Genes with an FPKM of >1 were included in this analysis. (**D**) GSEA enrichment plot for ZBTB18 target gene expression in M11GFP cells compared with E0771GFP cells (left) and M12GFP cells (right). NES, normalized enrichment score. (**E**) Relative *Zbtb18* mRNA levels in E0771GFP versus M11GFP (left; $n = 3$ biological replicates) and 67NR versus 4T1 (right; $n = 4$ biological replicates) cells, as determined by RT-qPCR. (**F**) ZBTB18 protein levels, detected by immunoblotting. GAPDH (glyceraldehyde-3-phosphate dehydrogenase) was used as loading control. Two biological replicates are shown. Right: Relative ZBTB18 expression levels calculated after normalization with GAPDH ($n = 4$ biological replicates). (**G** and **H**) Relative expression levels of selected ZBTB18 target genes in E0771GFP and M11GFP (G) ($n = 3$ biological replicates) or 67NR and 4T1 (H) ($n = 4$ biological replicates) cells, as determined by RT-qPCR. (**I**) ZBTB18 protein expression levels, in whole-cell lysates and nuclear fractions, detected by immunoblotting. Nuclear histone H2B (H2B) and cytoplasmic GAPDH were used as controls. Relative ZBTB18 expression levels calculated after normalization with H2B are shown. (**J**) Immunofluorescence staining for ZBTB18 in E0771GFP and M11GFP cells. Cell nuclei are stained with DAPI (4′,6-diamidino-2-phenylindole). Scale bar, 50 μm. (E to H) Means ± SEM, unpaired two-sided *t* test. $^{ns}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

localizes to the nucleus of highly metastatic M11GFP cancer cells in comparison to poorly metastatic E0771GFP cells. Thus, exclusion from nuclei is a mechanism that impedes ZBTB18 interaction with DNA, and this process may underlie the increased metastatic behavior of cancer cells.

As M11GFP cells are derived from E0771GFP cells, we next sought to determine the level of heterogeneity of ZBTB18 subcellular localization in the E0771GFP cell population. To this end, we derived clonal populations from 55 FACS-sorted E0771GFP single cells. Immunofluorescence staining of these clonal populations revealed that ZBTB18 is predominantly nuclear in most (52 of 55, ~95%) E0771GFP clones, whereas rare clones (3 of 55, ~5%) show enrichment of ZBTB18 in their cytoplasm (fig. S2B). Furthermore, all cells derived from a given

single-cell display the same intracellular localization of ZBTB18 (fig. S2B). Thus, although these results do not exclude the possibility that changes in ZBTB18 subcellular localization might be triggered during cancer progression and completion of the metastasis cascade in vivo, they raise the possibility that M11GFP cells originate from a rare E0771GFP cell with predominantly cytoplasmic ZBTB18. Together, these data indicate that a reduction in nuclear ZBTB18 levels that occurs independently from mutations or gene expression changes characterizes highly metastatic M11GFP cells.

Lower levels of nuclear ZBTB18 associate with more aggressive human breast cancers

We next sought to determine whether ZBTB18 protein levels and subcellular localization are altered in tumors from patients with breast cancer. Consistent with our data with murine models, bioinformatics analyses of publicly available datasets revealed that there is no significant difference in *ZBTB18* mRNA levels between breast cancers and normal breast tissues (Fig. 3A). Furthermore, no difference in *ZBTB18* mRNA levels in primary breast tumors were seen when comparing patients stratified according to distant metastasis status (Fig. 3B), lymph node metastasis status (Fig. 3C), or tumor stage (Fig. 3D). Comparison of samples obtained from distant metastases versus primary tumors also did not reveal any major differences in *ZBTB18* expression levels (Fig. 3E). Furthermore, mutations in the protein coding sequence of *ZBTB18* were very rare among patients with breast cancer. Mutations in ZBTB18 were found in only 3 of 1084 (0.28%) of the breast tumors analyzed (tables S3 and S4), indicating that loss-of-function mutations or reductions in *ZBTB18* expression levels are not likely to drive disease progression in patients with breast cancer. Mutations in ZBTB18 were also very rare in other cancer types, with only 108 of 10,967 (0.98%) of the tumors analyzed harboring a mutation in ZBTB18 (tables S3 and S4). In addition, except for glioblastoma, where *ZBTB18* expression is known to be repressed by promoter hypermethylation (*26*), no major differences in *ZBTB18* expression were noted when comparing tumors to normal tissues (fig. S3A) or when comparing metastatic to nonmetastatic cancers (fig. S3B). Thus, together, these data indicate that, in breast cancer and in many other cancer types, *ZBTB18* gene expression is not significantly altered during cancer progression.

To determine the subcellular localization of ZBTB18 in samples from patients with breast cancer, immunofluorescence staining was performed on tissue arrays, which included normal breast tissues (from patients with breast cancer), primary breast tumors from patients with different lymph node metastasis statuses (N0 to N3 according to the TNM Staging System), and samples from lymph node metastases. No significant differences in average ZBTB18 protein levels were noted across the different groups (Fig. 3, F and G). However, while ZBTB18 is generally expressed at similar levels in both the cytoplasm and nucleus in normal breast tissues and primary tumors from patients with no evidence of lymph node metastases (N0), it is more frequently enriched in the cytoplasm versus the nucleus in primary tumors from patients with metastases to multiple lymph nodes (N3) and in samples from lymph node metastases (Fig. 3, F and H). Similar results were also obtained with a second independent breast cancer tissue array where a depletion of nuclear ZBTB18 and concomitant enrichment in the cytoplasm was more frequently seen in invasive ductal carcinoma and lymph node metastases than in ductal carcinoma in situ and normal breast tissues (from patients with cancer; fig. S3, C to E). Overall, these data suggest that exclusion of ZBTB18 from the nucleus and its translocation to the cytoplasm associate with more aggressive and metastatic breast cancers in the clinic. As exclusion from nuclei is a mechanism that impedes ZBTB18 interaction with DNA, this process may underlie the increased metastatic behavior of cancer cells.

ZBTB18 inhibits breast cancer cell migration and metastasis

To investigate whether ZBTB18 is an inhibitor of breast cancer metastasis, we established M11GFP and E0771GFP cell lines with stable overexpression of ZBTB18 (overexpressed ZBTB18 localizes to both the cytoplasm and nucleus; fig. S4, A to C) and inoculated them orthotopically into the fourth mammary fat pad of C57BL/6 females. These experiments revealed that ZBTB18 overexpression in M11GFP cells does not affect breast tumor growth (Fig. 4A) but inhibits spontaneous lung metastasis (Fig. 4B). Overexpression of ZBTB18 in poorly metastatic E0771GFP cells further inhibited metastasis (Fig. 4, C and D), indicating that although metastasis is more limited in this model, it can also be blocked by ZBTB18 overexpression. Furthermore, following tail vein injection into syngeneic mice, E0771GFP and M11GFP tumor cells that stably overexpress ZBTB18 yielded fewer lung metastasis compared to control cells expressing an empty vector (EV) (Fig. 4E). A similar inhibition of metastasis was also observed upon overexpression of ZBTB18 in 4T1, 67NR, Ep5, and Ep5ExTu breast cancer cells and in MC38GFP colon carcinoma cells, further confirming the antimetastatic capacity of ZBTB18 across many cancer models (Fig. 4E and fig. S4, D to H).

Consistent with these observations, overexpression of ZBTB18 inhibited in vitro cell migration (Fig. 4F and fig. S4I) and invasion through collagen I (Fig. 4G) in cancer cell lines, whereas cell proliferation was not affected (Fig. 4H and fig. S4J). Furthermore, decreased collective invasion of tumor cells was observed when ZBTB18-overexpressing cells were grown as spheroids in a type I collagen:Matrigel three-dimensional matrix (fig. S4K). Together, these results demonstrate that ZBTB18 overexpression inhibits tumor cell motility and metastasis.

The predominantly cytoplasmic localization of ZBTB18 in lymph node metastases from patients with breast cancer (Fig. 3F) and in the highly metastatic M11GFP cells (Fig. 2, I and J) suggests that ZBTB18 localization to the nucleus might be necessary for its antimetastatic function. To test this hypothesis, we generated M11GFP cells with stable overexpression of ZBTB18 constructs with either nuclear localization (ZBTB18-Nucl) or nuclear exclusion (ZBTB18-Cyto) signals (fig. S4L). Overexpression of ZBTB18-Nucl in M11GFP cells inhibited lung metastasis in vivo (Fig. 4I) and tumor cell migration in vitro (Fig. 4J), while cell proliferation remained unaffected (Fig. 4K). In contrast, ZBTB18-Cyto overexpression affected neither metastasis, migration, nor proliferation (Fig. 4, I to K). Together, these results indicate that localization of ZBTB18 to the nucleus is necessary for its antimetastatic function. Conversely, they also support the hypothesis that inhibition of ZBTB18 activity would promote metastasis.

To test directly whether inhibition of ZBTB18 function would be sufficient to enhance metastasis, we next performed loss-of-function experiments by knocking out *Zbtb18* in the poorly metastatic E0771GFP cell line via the CRISPR-Cas9 technology (fig. S4M). Knockout of *Zbtb18* resulted in increased lung metastasis

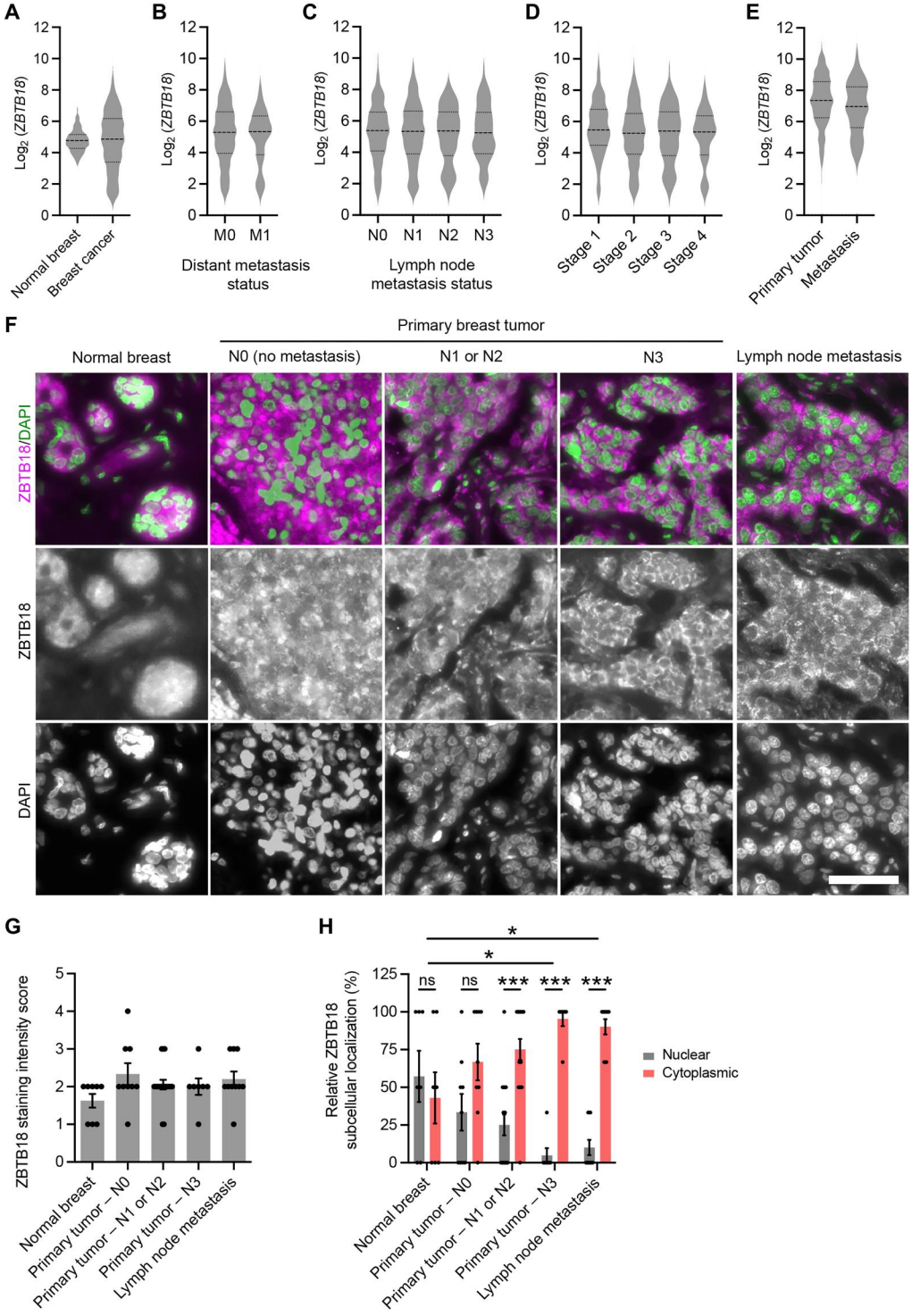

Fig. 3. Lower levels of nuclear ZBTB18 associate with more aggressive human breast cancers. (**A**) *ZBTB18* gene expression levels in tumors and adjacent normal tissues from patients with breast invasive carcinoma (normal breast, $n = 113$; breast cancer, $n = 1119$). (**B**) *ZBTB18* gene expression levels in primary tumors from patients with breast invasive carcinoma, stratified according to the presence (M1) or absence (M0) of metastases to distant organs (M0, $n = 1017$; M1, $n = 24$). (**C**) *ZBTB18* gene expression levels in primary tumors from patients with breast invasive carcinoma, stratified according to their lymph node metastasis status (TNM Staging System; N0, $n = 536$; N1, $n = 417$; N2, $n = 132$; N3, $n = 83$). (**D**) *ZBTB18* gene expression levels in primary tumors from patients with breast invasive carcinoma, stratified according to tumor stages (stage 1, $n = 203$; stage 2, $n = 693$; stage 3, $n = 277$; stage 4, $n = 22$). (**E**) *ZBTB18* gene expression levels in primary tumors and metastases from patients with breast carcinoma (primary tumor, $n = 868$; metastasis, $n = 140$). (**F**) Representative immunostaining for ZBTB18 in samples from patients with breast cancer, stratified according to lymph node metastasis status. DAPI staining delineates cell nuclei. Scale bar, 50 μm. (**G**) Average ZBTB18 staining intensity in tissue array samples from (F) (normal breast, $n = 7$; N0, $n = 9$; N1 or N2, $n = 18$; N3, $n = 7$; lymph node metastasis, $n = 10$). $P > 0.05$ for all comparisons. (**H**) Relative ZBTB18 subcellular localization in tissue array samples from (F) (normal breast, $n = 7$; N0, $n = 9$; N1 or N2, $n = 18$; N3, $n = 7$; lymph node metastasis, $n = 10$). (A to E) Median and upper and lower quartiles are indicated by dashed lines. (G and H) Means ± SEM. (A, B, and E) Mann-Whitney test (C, D, G, and H) and Kruskal-Wallis test, followed by Dunn's posttest. No significant difference in *ZBTB18* expression [$|\log_2(\text{fold change})| > 1$, $P < 0.05$] was found for any of the comparisons of the groups in (A) to (E). $^{ns}P > 0.05$, $^{*}P < 0.05$, and $^{***}P < 0.001$.

(Fig. 4L) and cell invasion (Fig. 4M) but did not affect cell proliferation (Fig. 4N). Thus, knockout of ZBTB18 promotes metastasis, whereas overexpression of nuclear ZBTB18 inhibits this process.

ZBTB18 induces a global decrease in chromatin accessibility and represses prometastatic gene expression programs

To decipher the molecular mechanism by which ZBTB18 suppresses breast cancer metastasis, we next performed RNA-seq of E0771GFP cells that stably overexpress ZBTB18 and compared them to control E0771GFP cells that express an EV (Fig. 5A).

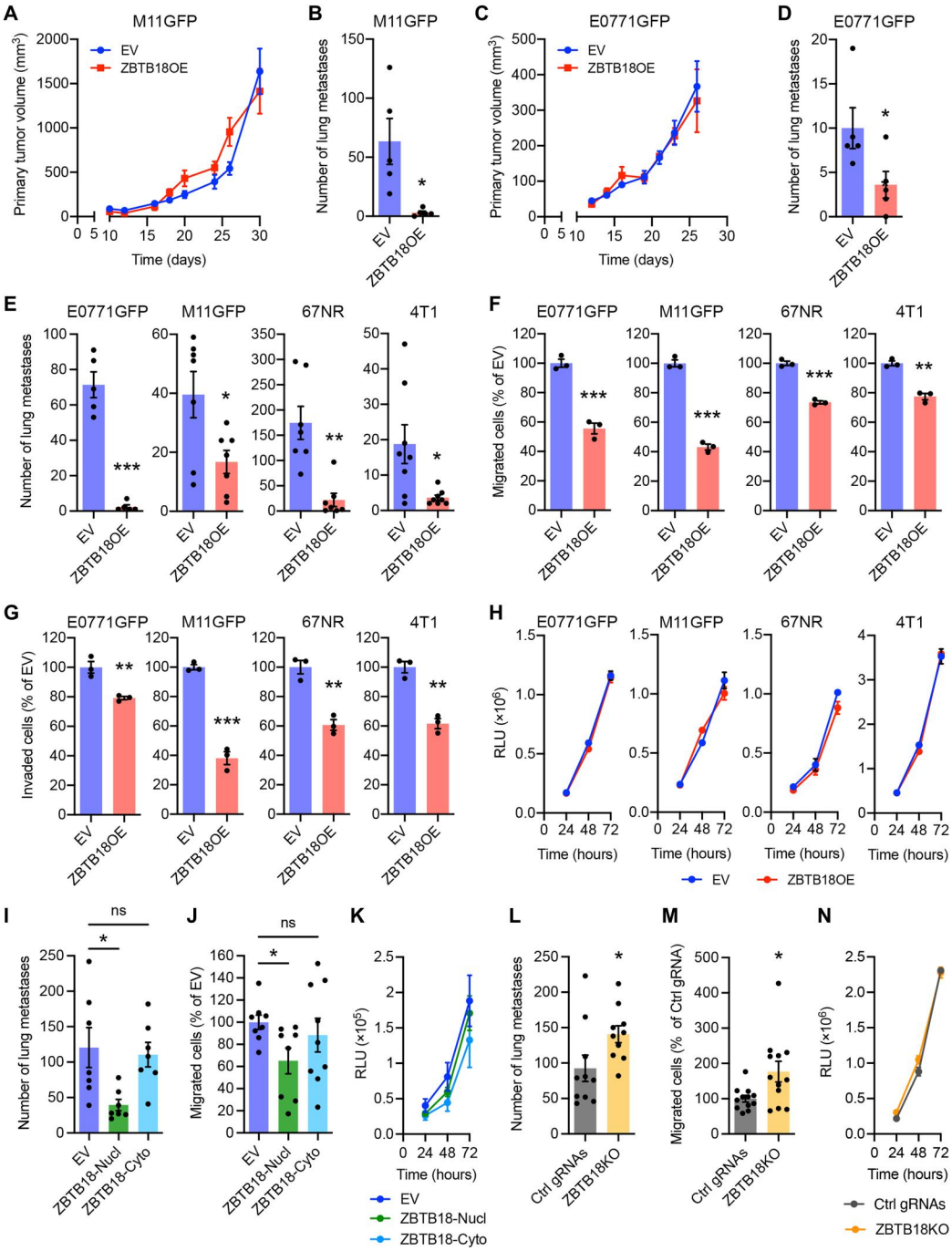

Fig. 4. ZBTB18 inhibits metastasis and suppresses tumor cell migration and invasion. (**A** and **C**) Tumor growth following orthotopic injection of 10^6 cells [M11GFP in (A) and E0771GFP in (C)] overexpressing ZBTB18 (ZBTB18OE) or an EV into the fourth mammary fat pad of C57BL/6 mice ($n = 5$ per group). (**B** and **D**) Number of spontaneous lung metastases in mice from (A) and (C), respectively, 30 or 28 days after tumor cell inoculation. (**E**) Number of lung metastases 14 days after tail vein injection of E0771GFP (10^6 per C57BL/6 mouse, $n = 5$ mice/group), M11GFP (2.5×10^5 per C57BL/6 mouse, $n = 7$ per group), 67NR (10^6 per BALB/c mouse, $n = 7$ per group), or 4T1 (2.5×10^5 per BALB/c mouse, $n = 8$ per group) cells overexpressing ZBTB18 (ZBTB18OE) or an EV. (**F**) Migration of E0771GFP, M11GFP, 67NR, and 4T1 cells overexpressing ZBTB18 (ZBTB18OE) or an EV through Transwell inserts ($n = 3$ biological replicates). (**G**) Invasion of cells from (F) through collagen I–coated Transwell inserts ($n = 3$ biological replicates). (**H**) Proliferation of cells from (F) ($n = 3$ independent experiments done in quadruplicate). (**I**) Number of lung metastases 14 days after tail vein injection of 5×10^5 M11GFP cells overexpressing ZBTB18 constructs with either nuclear localization (ZBTB18-Nucl) or nuclear exclusion (ZBTB18-Cyto) signals or an EV into C57BL/6 females ($n = 7$ mice/group). (**J**) Migration of cells from (I) ($n = 8$ to 9 from three independent experiments). (**K**) Proliferation of cells from (I) ($n = 3$ independent experiments done in quadruplicate). (**L**) Number of lung metastases 10 days after tail vein injection of 2×10^5 E0771GFP cells overexpressing Cas9 and either control nontargeting gRNAs (Ctrl gRNAs) or *Zbtb18*-targeting gRNAs (ZBTB18KO) into NSG females ($n = 10$ mice per group). (**M**) Migration of cells from (L) ($n = 4$ independent experiments done in triplicate). (**N**) Proliferation of cells from (L) ($n = 3$ independent experiments done in quadruplicate). Means ± SEM; (A to H, L, and M) unpaired two-sided *t* test; (I to K) one-way ANOVA and Tukey's posttest. $^{ns}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

This revealed that ZBTB18 overexpression induces the up-regulation of 252 genes [$\log_2$(fold change) > 1, false discovery rate (FDR) < 0.05], whereas 450 genes are significantly down-regulated [$\log_2$(fold change) < −1, FDR < 0.05; Fig. 5A]. This enrichment for down-regulated gene expression is consistent with the transcriptional repressor role of ZBTB18. Moreover, ATAC-seq (assay for transposase-accessible chromatin using sequencing) analyses of the same cell lines revealed that ZBTB18 overexpression induces changes in global DNA accessibility of 18,131 regions, with 6,349 regions showing increased accessibility and 11,782 showing decreased accessibility [|$\log_2$(fold change)| > 1, FDR < 0.05; Fig. 5, B and C]. These widespread changes are highly reproducible and biased toward an increase in closed chromatin conformation, with 12.7% of all regions adopting a newly closed conformation [$\log_2$(fold change) < −1, FDR < 0.05] and 6.8% adopting a more open conformation [$\log_2$(fold change) > 1, FDR < 0.05; Fig. 5, B and C

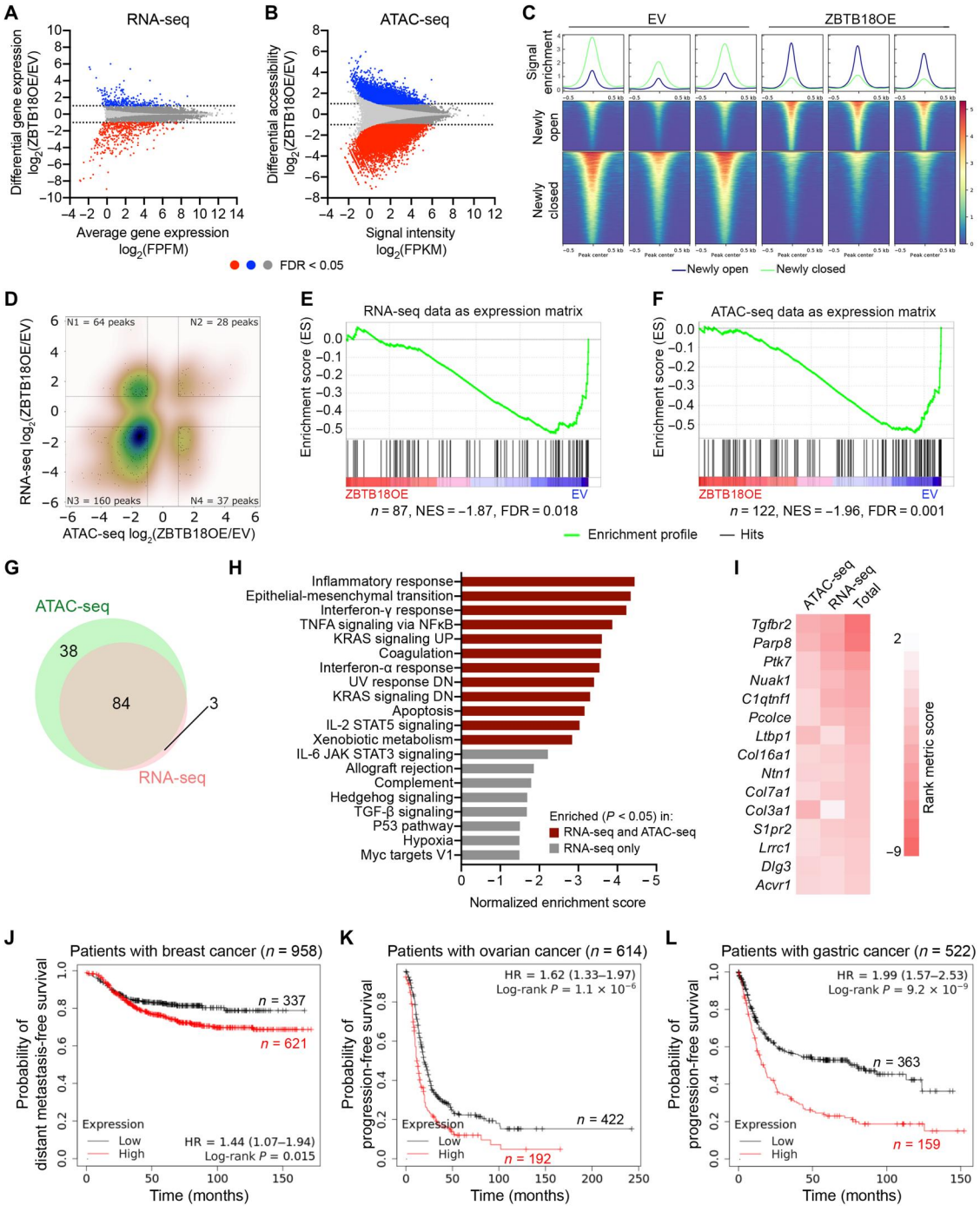

Fig. 5. ZBTB18 induces a global decrease in chromatin accessibility and represses prometastatic gene expression programs. (**A**) RNA-seq of E0771GFP cells stably overexpressing ZBTB18 (ZBTB18OE) compared to E0771GFP cells stably expressing an EV (n = 3 biological replicates). (**B**) Chromatin accessibility in cell lines from (A), determined by ATAC-seq (n = 3 biological replicates). (**C**) Heatmap centered at ATAC-seq nucleosome-free peak summits for differentially accessible regions in data from (B). Plots are from three biological replicates. (**D**) Correlation plot of genes and peaks modulated by >2-fold in both the RNA-seq and ATAC-seq datasets [from (A) and (B)]. (**E** and **F**) GSEA enrichment plot for ZBTB18 target gene expression in the RNA-seq data from (A) and the ATAC-seq data from (B). (**G**) Venn diagram for ZBTB18 target genes identified by RNA-seq (E) and ATAC-seq (F). (**H**) Pathways and biological processes (from the GSEA MSigDB Hallmark collection) that are enriched in genes down-regulated by >2-fold (P < 0.05) in RNA-seq and ATAC-seq data from (A) and (B). (**I**) ZBTB18 target gene signature. List of the 15 ZBTB18 target genes most significantly down-regulated in both the RNA-seq and ATAC-seq datasets. Total rank metric score is the sum of the rank metric scores from the GSEA analyses shown in (E) and (F). (**J**) Probability of distant metastasis-free survival for patients with breast cancer, stratified according to the mean expression levels of genes in the ZBTB18 target gene signature. (**K** and **L**) Probability of progression-free survival for patients with gastric (K) or ovarian (L) cancer, stratified according to the mean expression levels of genes in the ZBTB18 target gene signature. (J to L) Cutoff values were determined using the auto select best cutoff function in Kaplan-Meier plotter [(J), cutoff = 1449, range = 339 to 3979; (K), cutoff = 1360, range = 218 to 3475; (L), cutoff = 1651, range = 333 to 3527]. HR, hazard ratio and 95% confidence intervals.

and fig. S5, A to D]. As a comparison, global changes of an equivalent magnitude (although not biased toward chromatin closing) were reported when comparing basal and luminal A breast cancer subtypes (fig. S5E) (*30*, *31*). The changes in chromatin accessibility induced by ZBTB18 are also greater than those seen when comparing CD4 and CD8 T lymphocytes but smaller than those found when comparing B and T lymphocytes (fig. S5E) (*32*). Moreover, while ZBTB18 modulates accessibility at 19.5% of all chromatin regions (12.7% newly closed + 6.8% newly open), acute depletion of CTCF (CCCTC-binding factor), a master regulator of global three-dimensional chromatin architecture, affects 12.8% of all regions (6.7% newly closed + 6.1% newly open) in an acute lymphoblastic leukemia cell line (fig. S5E) (*33*).

ZBTB18 overexpression preferentially modulates chromatin accessibility at gene-distal regions while also affecting regions proximal to transcription start sites (TSSs; fig. S5F). Consistent with these observations, motif enrichment analysis revealed a significant enrichment for ZBTB18 binding motifs in regions with reduced accessibility in cells overexpressing ZBTB18 (table S5). Moreover, 257 transcription factor binding motifs are enriched [log2(fold change) > 0.5, FDR < 0.001] in regions with decreased accessibility, whereas only 56 motifs are enriched in regions with increased accessibility (tables S5 and S6), further indicating that ZBTB18-induced chromatin closing prevents access to many transcription factor binding sites. Thus, these data suggest that ZBTB18 inhibits gene expression by restricting DNA accessibility on a global scale, including at the enhancers and promoters of specific target genes. Examination of genes modulated by more than twofold in both RNA-seq and ATAC-seq reveals that 81% of the down-regulated genes (160 of 197) also adopt a less accessible DNA conformation (Fig. 5D and fig. S5G). As expected, GSEA analyses (with MSigDB C3 TFT gene sets) confirmed that ZBTB18 target genes are significantly down-regulated and less accessible in cells that overexpress ZBTB18, based on RNA-seq (Fig. 5E and table S7) and ATAC-seq datasets (Fig. 5F and table S8). The majority of genes predicted to have a ZBTB18 binding site in their promoter are transcriptionally down-regulated upon ZBTB18 overexpression (Fig. 5E) and display decreased chromatin accessibility in their promoter region (Fig. 5F). Furthermore, ZBTB18-regulated genes found with RNA-seq and ATAC-seq analyses significantly overlap (Fig. 5G), further suggesting that impairment of chromatin accessibility is a key mechanism by which ZBTB18 inhibits expression of its target genes.

Many of the genes suppressed by ZBTB18 overexpression are functionally clustered in biological processes and signaling pathways previously reported to play prominent roles in metastasis, such as inflammation, epithelial-to-mesenchymal transition, and nuclear factor κB signaling (Fig. 5H) (*34*–*36*). Furthermore, high expression of a ZBTB18 target gene signature consisting of the 15 genes most significantly repressed by ZBTB18 (Fig. 5I) correlates with shorter distant metastasis-free survival for patients with breast cancer (Fig. 5J and fig. S6, A to D). In addition, high expression of this ZBTB18 target gene signature also correlates with faster disease progression for patients with ovarian (Fig. 5K) or gastric (Fig. 5L) cancers. Correlations were significant when using either the mean, the geometric mean, or the sum of *z* scores of the expression levels of genes in the ZBTB18 target gene signature to stratify patients (fig. S6, E to J).

Our data indicate that ZBTB18 mRNA levels remain unchanged when comparing poorly and highly metastatic cell lines (Fig. 2E) and during tumor progression in humans (based on group averages; Fig. 3, A to E). However, given the wide distribution of *ZBTB18* gene expression levels among patients with cancer (Fig. 3, A to E) and the antimetastatic role played by ZBTB18, individual patients with lower expression levels of *ZBTB18* in their tumors might have a poorer prognosis. To test this hypothesis, we examined the correlation of *ZBTB18* expression levels with progression-free survival. These analyses revealed that high expression of *ZBTB18* correlates with slower disease progression for patients with breast cancer or gastric cancer (fig. S7). However, the magnitude of the absolute hazard ratios (|HR|) obtained when using *ZBTB18* mRNA as biomarker is smaller than those obtained when using the ZBTB18 target gene signature (Fig. 5, J to L, and fig. S7). Furthermore, no correlation is seen between *ZBTB18* mRNA levels and progression-free survival for ovarian cancer patients (fig. S7, B and E). Thus, consistent with a posttranscriptional regulation of ZBTB18 activity in aggressive cancers, the ZBTB18 target gene signature performs better than *ZBTB18* expression levels for stratifying patients according to the probability of progression-free survival. Together, these data suggest that ZBTB18-induced chromatin closing prevents the expression of genes associated with metastasis and that loss of ZBTB18 activity may underlie high metastatic capacity also in humans.

ZBTB18 directly binds to the promoter region of *Tgfbr2* and inhibits TGFβ1-induced signaling and cell motility

We next explored the functional involvement of ZBTB18 target genes in metastasis. Closer analysis of the ZBTB18 target genes identified by GSEA revealed that *Tgfbr2*, a gene encoding a receptor responsible TGFβ signal transduction, is one of the genes most significantly down-regulated by ZBTB18 overexpression (Fig. 5I) and that it harbors a profound loss of chromatin accessibility at its promoter (Fig. 6A). Consistent with these data, decreased TGFBR2 mRNA and protein levels were observed in E0771GFP, M11GFP, 67NR, and 4T1 cells that overexpress ZBTB18, in comparison to control cells transduced with an EV (Fig. 6, B and C).

Furthermore, chromatin immunoprecipitation (ChIP)–quantitative polymerase chain reaction (qPCR) analyses indicated that, in E0771GFP and M11GFP cells, ZBTB18 directly binds to two putative binding sites present in the *Tgfbr2* promoter region (Fig. 6, D and E). These analyses also revealed that ZBTB18 is more markedly associated to the *Tgfbr2* promoter in poorly metastatic E0771GFP cells than in highly metastatic M11GFP cells (Fig. 6E), consistent with diminished ZBTB18 repressor activity in aggressive tumor cells. Furthermore, transient transfection of ZBTB18 represses the activity of the *Tgfbr2* promoter, as determined by luciferase reporter assays (Fig. 6F). However, when ZBTB18 binding sites in the *Tgfbr2* promoter are mutated, *Tgfbr2* promoter activity remains unaffected by ZBTB18 overexpression (Fig. 6F). Repression of *Tgfbr2* promoter activity is dependent on ZBTB18 localization to the nucleus, as transfection of ZBTB18 flanked by nuclear localization signals (ZBTB18-Nucl) inhibits *Tgfbr2* promoter activity, whereas ZBTB18 harboring a nuclear exclusion signal (ZBTB18-Cyto) does not (Fig. 6G and fig. S8A). Thus, ZBTB18 directly interacts with the promoter of *Tgfbr2* to control its expression. Given that TGFβ signaling via TGFBR2 is required for efficient migration of E0771GFP, M11GFP, 67NR, and 4T1 tumor cells (fig. S8, B and C), our findings suggest that ZBTB18 might modulate the ability of tumor cells to respond to TGFβ signaling and to metastasize.

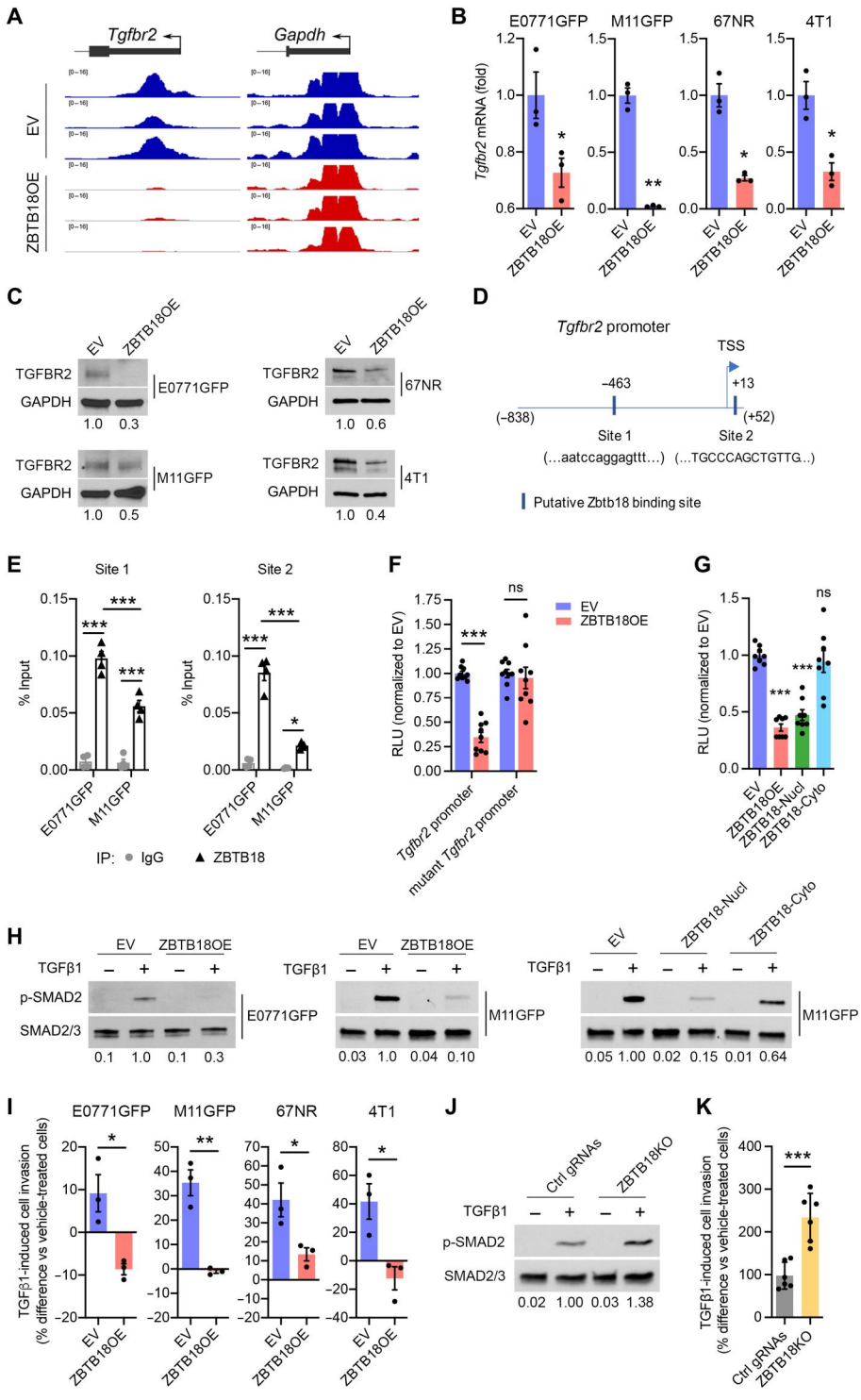

Fig. 6. ZBTB18 directly binds to the *Tgfbr2* promoter region and inhibits TGFβ1-induced signaling and cell motility. (**A**) Chromatin accessibility peaks in the *Tgfbr2* and *Gapdh* (control) promoters ($n = 3$ biological replicates). (**B**) *Tgfbr2* mRNA levels in E0771GFP, M11GFP, 67NR, and 4T1 cells overexpressing ZBTB18 (ZBTB18OE) or an EV ($n = 3$ biological replicates). (**C**) TGFBR2 protein levels in cells from (B), normalized to GAPDH. (**D**) Putative ZBTB18 binding sites in the *Tgfbr2* promoter, predicted using JASPAR2016 with 80% threshold; (**E**) ChIP-qPCR with anti-ZBTB18 antibodies and IgG controls. ZBTB18 binding to the *Tgfbr2* promoter in E0771GFP versus M11GFP cells is presented as % enrichment of input ($n = 4$ independent experiments). (**F**) Luciferase activity of wild-type and mutant *Tgfbr2* promoters (lacking ZBTB18 binding sites) in ZBTB18-overexpressing (ZBTB18OE) and control (EV) cells ($n = 9$, from three independent experiments). (**G**) Luciferase activity of wild-type *Tgfbr2* promoter in EV controls and in cells overexpressing wild-type (ZBTB18OE), nuclear-localized (ZBTB18-Nucl), and nuclear-excluded (ZBTB18-Cyto) ZBTB18 ($n = 8$, from four independent experiments). (**H**) Cells overexpressing ZBTB18 (ZBTB18OE), ZBTB18-Nucl, ZBTB18-Cyto, or EV controls treated with TGFβ1 (2 ng/ml) for 1 hour. Phosphorylated SMAD2 (p-SMAD2) is normalized to total SMAD2/3. (**I**) TGFβ1-induced invasion of ZBTB18OE and EV cells treated with TGFβ1 (2 ng/ml) for 16 hours compared to vehicle-treated controls ($n = 3$ biological replicates). (**J**) E0771GFP cells with *Zbtb18* knockout (ZBTB18KO; Cas9 + ZBTB18 gRNAs) and controls (Cas9 + Ctrl gRNAs) treated with TGFβ1 (2 ng/ml) for 1 hour; p-SMAD2 is normalized to total SMAD2/3. (**K**) TGFβ1-induced invasion of E0771GFP cells with *Zbtb18* knockout (ZBTB18KO) and controls treated with TGFβ1 (2 ng/ml) for 16 hours versus vehicle controls ($n = 6$, from two independent experiments). (B, E to G, I, and K) Means ± SEM. (B, I, and K) unpaired two-sided *t* test; (E and F) two-way ANOVA with Tukey's posttest; (G) one-way ANOVA with Tukey's posttest; $^{ns}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

In accordance with this hypothesis, treatment with TGFβ1 led to a stronger induction of SMAD2 phosphorylation in M11GFP cells than in E0771GFP cells (fig. S8D). In addition, *Tgfbr2* repression by ZBTB18 led to a reduction in TGFβ1-induced signaling, as indicated by the marked decrease in SMAD2 phosphorylation in tumor cells overexpressing ZBTB18 or ZBTB18-Nucl (Fig. 6H). Furthermore, whereas treatment of control E0771GFP, M11GFP, 67NR, and 4T1 tumor cells (expressing an EV) with TGFβ1 enhanced their migration and invasion through collagen lattices, these effects were blocked in ZBTB18 overexpressing cells (Fig. 6I and fig. S8E). Conversely, CRISPR-Cas9–mediated *Zbtb18* knockout enhanced TGFβ1-mediated SMAD2/3 phosphorylation and TGFβ1-induced tumor cell invasion (Fig. 6, J and K). Together, these

results indicate that ZBTB18 inhibits TGFβ1-dependent tumor cell invasion and migration.

Forced expression of TGFBR2 in tumor cells expressing ZBTB18 enhances cell motility, promotes metastasis, and reenables prometastatic TGFβ1 signaling

To further delineate whether down-regulation of *Tgfbr2* contributes to the effect of ZBTB18 on metastasis, rescue experiments were done. For these studies, a doxycycline-inducible *Tgfbr2* construct was introduced into cell lines that stably express ZBTB18 (fig. S9). Reintroduction of TGFBR2 in ZBTB18-expressing cells rescued SMAD2 phosphorylation in response to TGFβ1 treatment (Fig. 7A), increased in vitro tumor cell migration and invasion (Fig. 7, B and C), and promoted in vivo metastasis to the lungs (Fig. 7D). Together, these results indicate that *Tgfbr2* down-regulation is a key effector of the antimetastatic function of ZBTB18. Furthermore, consistent with these data, a significant inverse correlation was found between TGFβ1- and ZBTB18-induced gene expression signatures in E0771GFP cells (Fig. 7E). Specifically, 69.9% of the genes significantly modulated by both TGFβ1 treatment and ZBTB18 overexpression displayed an inverse correlation, thus further demonstrating the antagonistic effect of ZBTB18 on TGFβ1 signaling.

Next, we investigated whether ZBTB18 exerts its influence on TGFβ1-regulated genes mainly via its repression of *Tgfbr2* expression or, alternatively, via its capacity to alter chromatin accessibility at the promoter of these genes. For these studies, we analyzed the effect of ZBTB18 overexpression on chromatin accessibility at the promoter of TGFβ1-regulated genes [$|\log_2(\text{fold change})| > 1$]. This analysis revealed that, although 27 of 89 (30%) of TGFβ1-modulated genes display changes in chromatin accessibility upon ZBTB18 overexpression, the magnitude of changes in TGFβ1-induced gene expression and ZBTB18-induced chromatin accessibility do not significantly correlate (Fig. 7F). Together, these findings indicate that ZBTB18 regulates TGFβ1-induced gene expression primarily via its modulation of *Tgfbr2* expression rather than through changes in the accessibility of TGFβ1 target gene promoters. Nonetheless, increased chromatin accessibility due to loss of ZBTB18 activity likely also contributes to other prometastatic gene expression changes that might synergize with increased TGFBR2-mediated signaling.

In summary, this study reveals that inhibition of ZBTB18 activity in tumor cells promotes metastasis by increasing chromatin accessibility and by enabling prometastatic TGFβ1-TGFBR2–driven changes. This gene expression program and global chromatin remodeling confer cancer cells the phenotypic adaptations needed to complete the metastatic cascade.

DISCUSSION

Metastasis is an inefficient and challenging process with a very high rate of failure. To successfully metastasize, tumor cells must be able to adapt to the specific challenges and demands of each step of the metastatic cascade. However, the molecular mechanisms that enable the remarkable cellular plasticity needed for efficient metastasis remain incompletely understood.

By using a model that recapitulates the entire metastatic cascade in an intact immune environment, we find that loss of activity of the transcriptional repressor ZBTB18 confers cancer cells the capacity to complete the metastatic cascade and form overt lung metastases. While total ZBTB18 protein levels remain unchanged, a reduction in nuclear ZBTB18 levels, accompanied by up-regulation of its target genes, is seen in highly metastatic breast cancer cells in comparison to less aggressive cells. Furthermore, we show that restoration of ZBTB18 activity via its overexpression inhibits tumor cell invasion and metastasis, without affecting tumor growth. Metastasis and cell motility are inhibited when ZBTB18 overexpression is localized to the cell nucleus but not when it is largely restricted to the cytoplasm, suggesting that nuclear localization of ZBTB18 is necessary for its antimetastatic effect. Conversely, we find that knockout of ZBTB18 enhances metastasis and motility of poorly metastatic cancer cells, further supporting the notion that loss of ZBTB18 activity facilitates metastasis. Furthermore, as the antimetastatic effect of ZBTB18 overexpression is observed in both highly and poorly metastatic breast cancer models, our data suggest that ZBTB18 activity might be partially repressed also in poorly metastatic tumor cell populations. In this scenario, interventions aiming at enhancing the activity of ZBTB18 might limit the progression not only of highly metastatic cancers but also of less aggressive (but still metastatic) cancers. Lower ZBTB18 activity triggers expression programs associated with inflammation, epithelial-to-mesenchymal transition, and tumor cell motility and invasion. These transcriptional changes facilitate the completion of many steps of the metastatic cascade such as invasion of surrounding tissues, intravasation, and extravasation into the parenchyma of distant tissues (*1*, *37*, *38*). In accordance with these observations, low levels of nuclear ZBTB18 accompanied by its higher cytoplasmic localization are more frequently seen in aggressive primary tumors and lymph node metastasis samples from patients with breast cancer than in earlier-stage tumors or normal tissues. Furthermore, high expression of a ZBTB18 target gene signature correlates with shorter distant metastasis-free or progression-free survival for patients with breast, ovarian, or gastric cancers, suggesting that inhibition of ZBTB18 activity may underlie high metastatic capacity also in humans.

Consistent with its transcriptional repressor function, we find that the effect of ZBTB18 on gene expression is strongly biased toward inhibition and that the majority of genes inhibited by ZBTB18 also display a loss of chromatin accessibility at their promoter. Our ATAC-seq analyses also revealed that ZBTB18 overexpression induces a widespread decrease in chromatin accessibility, as a substantial proportion [12.7% by $\log_2(\text{fold change}) < -1$, FDR < 0.05] of all regions accessible in control cells are repressed in ZBTB18-overexpressing cells. In contrast, a lower fraction [6.8% by $\log_2(\text{fold change}) > 1$, FDR < 0.05] of chromatin regions show increased accessibility, indicating that the effect of ZBTB18 is biased toward the induction of chromatin closing. To put these findings in perspective, global changes of a similar magnitude (although not biased toward chromatin closing) were reported when comparing basal and luminal A breast cancer subtypes (*30*, *31*). The changes in chromatin accessibility induced by ZBTB18 are also greater than those seen when comparing CD4 and CD8 T lymphocytes but smaller than those found when comparing B and T lymphocytes (*32*). Moreover, while we find that ZBTB18 modulates accessibility at 19.5% of all chromatin regions (12.7% newly closed + 6.8% newly open), acute depletion of CTCF, a master regulator of global three-dimensional chromatin architecture, was previously reported to affect 12.8% of all regions (6.7% newly closed +

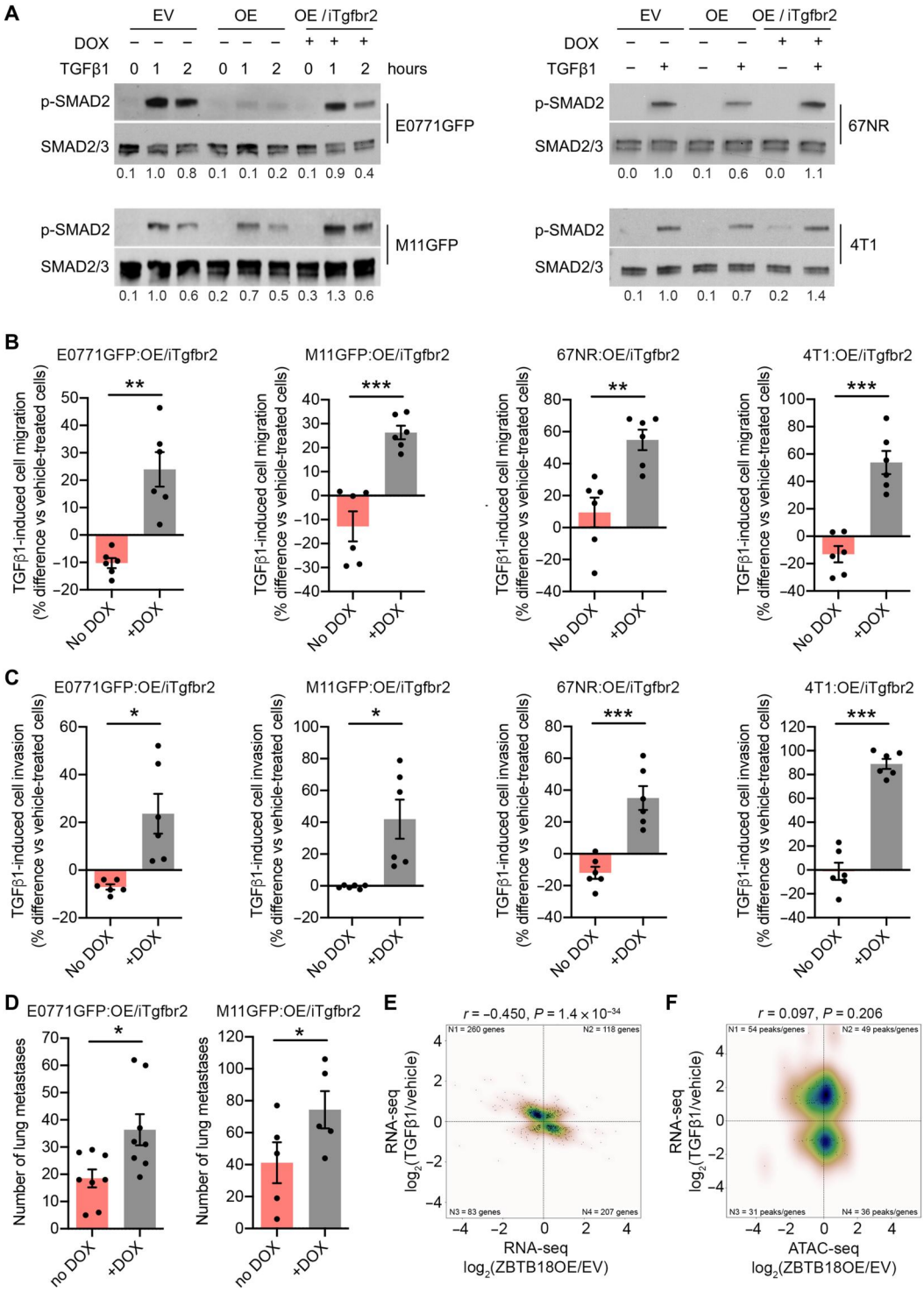

Fig. 7. TGFBR2 overexpression rescues ZBTB18-mediated inhibition of cell migration and metastasis. (**A**) p-SMAD2 and total SMAD2/3 levels detected by immunoblotting in cell lines stably overexpressing ZBTB18 (OE), ZBTB18 + doxycycline-inducible TGFBR2 (OE/iTgfbr2) or an EV. TGFBR2 was induced with doxycycline (1 μg/ml), and cells were treated with TGFβ1 (2 ng/ml) or vehicle. p-SMAD2 expression levels relative to EV + TGFβ1 (1 hour) (calculated after normalization with total SMAD2/3) are shown. (**B** and **C**) TGFβ1-induced migration (B) and invasion (C) through Transwell inserts of cell lines overexpressing ZBTB18 and doxycycline-inducible TGFBR2 (OE/iTgfbr2). TGFBR2 expression was induced with doxycycline (1 μg/ml; DOX+; see fig. S9), and cells were treated with TGFβ1 (2 ng/ml) or vehicle. Data are shown relative to vehicle-treated control groups ($n = 3$ biological replicates). (**D**) Number of lung metastases in C57BL/6 mice 14 days after tail vein injection of 10^6 E0771GFP or M11GFP cells stably overexpressing ZBTB18 + inducible TGFBR2 (OE/iTgfbr2). TGFBR2 expression was induced via daily intraperitoneal injection of doxycycline (5 mg/kg body weight) starting 24 hours after tumor cell inoculation (E0771GFP, $n = 8$ mice per group; M11GFP, $n = 5$ mice per group). (**E**) Correlation plot of transcripts significantly modulated (FDR < 0.05) in E0771GFP cells upon ZBTB18 overexpression (*y* axis) and upon reexpression of TGFBR2 and treatment with TGFβ1 (*x* axis) ($n = 3$ biological replicates). Pearson correlation = −0.450, $P < 10^{-34}$. (**F**) Correlation plot of ZBTB18-induced modulation of chromatin accessibility (ATAC-seq data; *x* axis) and of TGFβ1-modulated transcripts with $|\log_2(\text{fold change})| > 1$ (TGFβ1-induced gene expression signature; *y* axis) ($n = 3$ biological replicates). Pearson correlation = −0.097, $P = 0.206$. (B to D) Means ± SEM, unpaired two-sided *t* test. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

6.1% newly open) in an acute lymphoblastic leukemia cell line (*33*). We thus propose that ZBTB18 overexpression triggers a substantial decrease in chromatin accessibility that is similar in magnitude to the chromatin remodeling that occurs during some cell differentiation programs. Furthermore, consistent with a role for ZBTB18 in reducing chromatin accessibility, it has been previously reported that, in neuronal cells, almost all ZBTB18 is localized to condensed chromatin regions (*39*) where it actively represses stem/progenitor cell programs during brain development, thereby allowing neuronal differentiation (*40*).

Although the association of specific chromatin accessibility states with metastatic competence remains largely unexplored, it has been previously reported that cells isolated from liver metastases of an SCLC mouse model harbor a more open chromatin state than

cells from primary tumors (*17*). Together with our data, this suggests that a more open chromatin state promotes metastasis. In the SCLC model, increased chromatin accessibility and enhanced metastasis are induced by *Nfib* overexpression, which results from copy number amplification (*17*). Because *Nfib* overexpression stems from a genetic alteration, this implies that the change in chromatin accessibility that it induces is predetermined and irreversible. In contrast, our data indicate that reduction of ZBTB18 nuclear localization and activity, but not in its expression, characterize highly metastatic breast cancer cells. This raises the possibility that ZBTB18-dependent chromatin remodeling might occur in a dynamic and reversible manner and thereby contribute to the acquisition of transient states associated with more invasive phenotypes (*4*, *41*). However, whether ZBTB18 subcellular localization is irreversible in a given tumor cell or whether it evolves during cancer metastasis remains to be explored.

Our data also pinpoint *Tgfbr2* as a key gene that is up-regulated in highly metastatic cell lines and directly inhibited by ZBTB18-induced chromatin closing. In addition, we find that ZBTB18 directly binds to the *Tgfbr2* promoter in poorly metastatic cells but less so in highly metastatic cells, indicating that *Tgfbr2* expression is directly regulated by ZBTB18. TGFBR2 is the cell surface receptor responsible for mediating TGFβ signaling, which has been shown to play a dichotomous and context-dependent role in cancer progression. While TGFβ triggers cytostatic signals that can limit early tumor growth, it also promotes cancer metastasis via other processes (*42*, *43*). These mechanisms include the enhancement of the migratory and invasive properties of tumor cells, the induction of epithelial-mesenchymal transition, the gain of stem-like properties, and the promotion of immune evasion (*42*, *43*). In accordance with a prometastatic role of TGFβ signaling, treatment with TGFβ induced invasion and migration of all cell lines used in this study. However, ZBTB18 overexpression and consequent loss of TGFBR2 led to the inhibition of TGFβ/SMAD signaling and prevented the induction of tumor cell migration and invasion by TGFβ1. Forced expression of TGFBR2 in tumor cells expressing ZBTB18 reenables TGFβ1-induced cell motility and promotes metastasis. Thus, inhibition of the TGFβ1 pathway is one of the key mechanisms by which ZBTB18 exerts its antimetastatic effect. Furthermore, this suggests that ZBTB18 loss and consequent up-regulation of TGFBR2 enable cells to readily respond to proinvasive TGFβ1 present in primary tumors or provided at the site of extravasation, as a result of platelet-tumor cell interactions (*36*). Along these lines, in patient-derived xenograft models of breast cancer, *Tgfbr2* was identified as one of the genes most highly up-regulated in early metastatic lesions in comparison to primary tumors or overt metastases (*5*), suggesting that it may play a key role during the establishment of metastases and that its expression levels are subject to modulation during metastatic progression.

ZBTB18 widely affects chromatin accessibility and therefore likely impedes metastasis by repressing many prometastatic genes. On this basis, future experiments should identify additional prometastatic genes that are repressed by ZBTB18 and test their possible interactions with TGFβ signaling. Together, these studies would provide further insight into the molecular mechanisms of ZBTB18-mediated repression of cancer metastasis.

In conclusion, our study uncovers ZBTB18 as a factor that induces widespread chromatin closing and that suppresses breast cancer invasion and metastasis at least in part via the repression of TGFβ1-driven prometastatic gene expression changes. We thus propose a model whereby the loss of ZBTB18 activity empowers tumor cells to respond to specific prometastatic environmental cues.

MATERIALS AND METHODS

Experimental design

The objectives of this study were to generate new models of spontaneous breast cancer metastasis that are syngeneic to immunocompetent C57BL/6 mice and to use these models to delineate gene expression and phenotypic programs that are differentially modulated in tumor cells with high versus low metastatic ability. RNA-seq and bioinformatics analyses of two new metastasis-derived cell lines and of their parental cell line indicated that the loss of activity of the transcriptional repressor ZBTB18 characterizes tumor cells that have a higher ability to metastasize. The mechanism of action of ZBTB18 and its impact on breast cancer metastasis was then assessed in six breast cancer models [E0771GFP, M11GFP (developed in this study), 67NR, 4T1, Ep5, and Ep5ExTu cell lines] and one colon cancer mode (MC38GFP) using syngeneic mouse models of spontaneous and experimental metastasis; RNA-seq; ATAC-seq; reverse transcription qPCR (RT-qPCR); immunoblotting; ChIP-qPCR; promoter activity reporter assays; and cell proliferation, migration, and invasion assays. We also examined the expression levels and subcellular localization of ZBTB18 in primary tumors and lymph node metastasis samples from patients with breast cancer and determined the correlation of a ZBTB18 target gene signature with distant metastasis-free survival for patients with breast, ovarian, or gastric cancers.

Mice

C57BL/6J (RRID:IMSR_JAX:000664) female mice were purchased from the Jackson Laboratory, and BALB/cAnNTac (RRID:IMSR_TAC:balb) female mice were purchased from Taconic Biosciences. NSG (NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ; RRID:IMSR_JAX:005557) mice were from the Jackson Laboratory and maintained at St. Jude Children's Research Hospital. Experiments were started when the mice were 7 to 8 weeks old. Mice were housed and handled in accordance with approved St. Jude Children's Research Hospital Institutional Animal Care and Use Committee protocols.

Cell culture

E0771 breast cancer cells (CH3 BioSystems; RRID:CVCL_GR23) and its derivative cell lines (E0771GFP, M11GFP, and M12GFP) were maintained in RPMI 1640 (Thermo Fisher Scientific) supplemented with 10% calf serum with iron (HyClone) and 1% penicillin-streptomycin. 67NR (from F. Miller, Karmanos Cancer Institute; RRID:CVCL_9723) and 4T1 [American Type Culture Collection (ATCC), catalog no. CRL-2539; RRID:CVCL_0125] breast cancer cells were grown in RPMI 1640 supplemented with 10% fetal bovine serum (HyClone) and 1% penicillin-streptomycin at 37°C and 5% CO_2. 293FT cells (ATCC, catalog no.PTA-5077; RRID:CVCL_6911), Ep5, and Ep5ExTu breast cancer cells (*44*, *45*), and MC38GFP colon cancer cells (*46*, *47*) were cultivated in Dulbecco's modified Eagle's medium, 10% fetal bovine serum, and 1% penicillin-streptomycin. Cells were used at passage 4 to 15 after thawing and were mycoplasma negative, as tested with the MycoAlert Mycoplasma Detection Kit (Lonza).

For some experiments, cells were treated with doxycycline (1 μg/ml; Sigma-Aldrich, #9891), recombinant TGFβ1 (2 ng/ml; R&D Systems, #240-B), anti-TGFβ1 blocking antibody (6 μg/ml; R&D Systems, #MAB240; RRID:AB_358119), mouse immunoglobulin G1 (IgG1) isotype control (6 μg/ml; R&D Systems, #MAB002; RRID:AB_357344), or 10 μM TGFBR1 inhibitor SB431542 (Cayman Chemical, #13031), as indicated in the figure legends.

Establishment of E0771GFP, M11GFP, and M12GFP cell lines

To establish the E0771GFP cell line, E0771 cells were stably transduced with a GFP-expressing lentiviral vector (pCDH-EF1-MCS-T2A-copGFP vector; System Biosciences) as described previously (*48*). GFP^+ cells were then FACS-sorted, expanded in tissue culture, and inoculated (10^6 cells) into the fourth mammary fat pad of a C57BL/6J female. After 28 days, the mouse was euthanized, and two distinct GFP^+ lung metastatic foci were dissected under a fluorescence stereomicroscope (Zeiss). Recovered tissue was then dissociated and digested with collagenase I at 37°C for 45 min. The cell suspensions were then filtered through a 70-μm cell strainer and plated in 10-cm dishes containing RPMI 1640 supplemented with 10% calf serum with iron and 1% penicillin-streptomycin. Tumor cells were then expanded in tissue culture, FACS-sorted on the basis of GFP expression, and maintained as the M11GFP and M12GFP cell lines.

Establishment of cell lines with stable expression of ZBTB18 constructs and inducible TGFBR2 or with CRISPR-Cas9–mediated knockout of *Zbtb18*

A *Zbtb18* expression construct (ZBTB18OE) was produced by subcloning *Zbtb18* cDNA flanked by Xba I and Bam HI restriction sites (gBlocks; IDT) into the pCDH-CMV-MCS-EF1-Puro lentiviral vector (EV; System Biosciences). To generate a ZBTB18 construct that localizes preferentially to the nucleus (ZBTB18-Nucl), a c-Myc nuclear localization signal [PAAKRVKLD (*49*)] was added at the 5′ end and an SV40 large T antigen nuclear localization signal [PKKKRKV (*50*)] was added at the 3′ end of the ZBTB18 protein coding sequence by PCR cloning. Alternatively, a nuclear export sequence [SELQNKLEELDLDSYK (*51*)] was added at the 3′ end of ZBTB18 to generate a construct that localizes preferentially to the cytoplasm (ZBTB18-Cyto; see table S9 for primers). ZBTB18 constructs were then subcloned into the pCDH-CMV-MSC-EF1-Puro lentiviral vector as above.

For CRISPR-Cas9–mediated *Zbtb18* knockout, E0771GFP cells were first transduced with the LentiCas9-Blast vector (gift from F. Zhang; Addgene, plasmid no. 52962) and selected with blasticidin to establish the E0771GFP-Cas9 stable cell line. E0771GFP-Cas9 cells were then further transduced with a pool of three control guide RNA (gRNA) or a pool of five *Zbtb18*-targeting gRNA viral vectors with puromycin resistance gene expression (lentiGuide-Puro vector; a gift from F. Zhang; Addgene, plasmid no. 52963; see table S10 for gRNA sequences). This approach, also used for pooled screens (*52*), yields high cleavage efficiency and thus does not require clonal selection of targeted cells.

For the inducible *Tgfbr2* construct, *Tgfbr2* was cut out from the pCMV-mTgfbr2-GFPSpark plasmid (Sino Biological) with the Nhe I and Bam HI restriction enzymes and then subcloned into the Xba I and Bam HI restriction sites of the pCW-Cas9-Blast vector (a gift from M. Babu; Addgene, plasmid no. 83481).

Packaging of the vectors described above was obtained by co-transfecting 293FT cells with 1.3 μg of transfer vector and 0.67 μg of pCAG-kGP1-1R, 0.22 μg of pCAG-VSV-G, and 0.22 μg of pCAG4-RTR2 or 1 μg of pCMV-dR8.9 and 1 μg of pCAG-VSV-G helper plasmids using TransIT-293 reagent (Mirus) in Opti-MEM I (Thermo Fisher Scientific). Forty-eight hours after transfection, conditioned medium from 293FT cells was collected, filtered through a 0.45-μm filter, and applied to tumor cells with Polybrene (4 μg/ml; Sigma-Aldrich). The tumor cells were then selected with puromycin or blasticidin (5 μg/ml; Gibco) beginning 24 hours after transduction.

Mouse models of metastasis

To analyze tumor progression and spontaneous metastasis, breast cancer cells resuspended in 200 μl of Hanks′ balanced salt solution (HBSS) were inoculated orthotopically into the fourth (inguinal) mammary fat pad of syngeneic female mice (see the figure legends for the number of cells inoculated). Tumor growth in the orthotopic site was monitored by measuring individual tumors every 2 to 3 days with calipers and calculating the tumor volume [tumor volume (mm^3) = (length × $width^2$)/2]. Twenty-eight or thirty days after inoculation, mice were euthanized with CO_2. Primary breast tumors were then resected and weighed. Lungs were collected and fixed in HBSS and 4% formaldehyde for 30 min. For tumor cell lines expressing GFP, metastatic foci at the surface of lungs were counted under a fluorescence EVOS microscope (Thermo Fisher Scientific). For tumor cell lines without GFP expression, lungs were fixed in 4% formaldehyde for 24 hours, sectioned, stained with hematoxylin and eosin (H&E), and imaged. The number of metastases per animal was determined by counting metastatic foci in five H&E-stained lung sections per mouse.

For the experimental metastasis models, tumor cells resuspended in 100 μl of HBSS were inoculated into syngeneic female mice via tail-vein injection. After 14 days, mice were euthanized, lungs were resected, and metastases were quantified as described above for spontaneous metastasis. For experiments with tumor cells stably overexpressing ZBTB18 and doxycycline-inducible TGFBR2 (OE/iTgfbr2), TGFBR2 expression was induced via daily intraperitoneal injection of doxycycline (5 mg/kg body weight) starting 24 hours after tumor cell inoculation.

Proliferation assay

The number of viable cells was determined using the CellTiter-Glo Luminescent Cell Viability Assay Kit (Promega) following the manufacturer′s protocol. Luminescence was detected with an Infinite M200Pro plate reader and the i-control 1.10 software (Tecan).

Migration and invasion assay

Transwell migration was performed using the HTS FluoroBlok 96-well Multiwell System plates (Corning). A total of 1.25×10^4 cells per well in serum-free medium were seeded in the upper insert with 8-μm pores. Medium plus 5% serum was used as chemoattractant in the lower chamber. Twenty-four hours later, cells that had migrated to the bottom chamber were stained with Calcein AM (4 μg/ml in HBSS) for 1 hour and imaged using an Olympus IX70 fluorescence inverted microscope and the DP Manager software (version 33.1.222, Olympus). Migrated cells in each image were analyzed with the Analyze Particle tool in the Fiji software.

Transwell invasion was performed as for the migration assay using FluoroBlok Transwell inserts precoated with collagen type I (PureCol, Advanced BioMatrix) at 37°C for 3 hours before cell seeding.

Spheroid invasion assay

Spheroid invasion was performed as described (*53*). Briefly, 40 drops of single-cell suspension (20 μl per drop containing 750 cells) were seeded onto the lid of a 10-cm dish and incubated inverted over a culture dish containing 5 ml of phosphate-buffered saline (PBS) for 72 hours to generate spheroids. The spheroids were then collected and resuspended in 200 μl of 1:1, serum-free medium:Matrigel (Corning) and seeded into a precoated 24-well plate in quadruplicate. The culture was polymerized at 37°C for 30 min; conditioned medium (1 ml) was added, and spheroids were incubated overnight. The invasion of spheroids was imaged with an EVOS microscope and analyzed as the total area of invasion outside the initial spheroid area using the Fiji software.

RNA-seq and data analysis

Total RNA was isolated from cell lysates using the RNeasy Plus Mini Kit (QIAGEN), and RNA-seq analysis was performed as previously described (*48*). Alternatively, raw sequencing reads in fastq format were quality-filtered with Trim Galore tool [F. Krueger (2012); www.bioinformatics.babraham.ac.uk/projects/trim_galore/] and mapped to the mouse genome reference (GRCm38) with STAR (*54*). Reads per gene were quantified with the RSEM (RNA-seq by Expectation Maximization) software (*55*), and the limma-voom approach (*56*, *57*) was used to calculate differential gene expression statistics, using only level 1 and 2 protein-coding genes. Moreover, only the genes with more than 10 read counts per mean library size (in million) in at least the minimum group sample size were kept. PCA mapping was performed using $\log_2$(FPKM) (fragments per kilo base per million mapped reads) values in Partek Genomic Suite 6.6 (www.partek.com/partek-genomics-suite/).

ATAC-seq and data analysis

ATAC-seq libraries were prepared as previously described (*58*, *59*). Briefly, transposition was performed using the Nextera DNA Library Prep Kit (Illumina, #20034197) with nuclei from 5×10^4 cells. After purification with a DNA Clean & Concentrator-5 kit (Zymo Research, #D4013), the libraries were amplified by PCR using the NEBNext HiFi 2× PCR Master Mix (NEB, #M0541S). PCR fragments were purified with the QIAGEN MinElute PCR Purification Kit, and the library concentration and quality were determined with an Agilent High Sensitivity DNA Kit and Bioanalyzer, respectively. ATAC-seq libraries were sequenced with an Illumina HiSeq Platform with 100–base pair (bp) paired-end reads in triplicate. Next, sequenced reads were processed with Trim Galore tool (v0.4.4); potential adapters were removed, and the 3′ ends of reads were quality-trimmed with Cutadapt (DOI:10.14806/ej.17.1.200) using the quality cutoff of Q20. The first 15 bp of each read were clipped to avoid GC (guanine-cytosine) bias. Reads were then mapped to the mouse reference genome (GRCm38) with bwa sampe v0.7.17-r1198 (*60*); duplicated reads were identified with the bamsormadup tool from biobambam2 (v2.0.87, DOI: 10.1186/1751-0473-9-13), and properly paired and uniquely mapped reads were extracted in BAM format with samtools v1.2 (*61*). Then, using bedtools v2.24.0 (*62*), nucleosome-free fragments (defined as fragments <100 bp) were extracted, and MACS2 tool v2.1.1.20160309 (*63*) was used to call peaks in narrow format with --extsize 200 ---nomodel -q 0.05 flags (high-confidence peaks). Separately, ATAC-seq peaks were also called using more relaxed criteria, setting the -q flag to 0.5, which are here referred to as low-confidence peaks. Last, the reproducible high-confidence peaks between the biological replicates from the same condition were identified as those that, for the same genomic region, either had high- or low-confidence peaks called in all other biological replicates. Reproducible peaks from each condition were then merged into one collection of reproducible peaks.

To identify differentially accessible regions between experimental groups, the nucleosome-free fragments for each reproducible peak were calculated with intersect command from pybedtools v0.8.1 (*62*, *64*). Next, the number of raw reads mapping per peak was converted to FPKM unit and trimmed mean of M-values from edgeR (*65*). The limma-voom approach (*56*, *57*) was used to assess the significance of differential peak accessibility.

The same approach was used to reanalyze publicly available ATAC-seq datasets from The Cancer Genome Atlas (TCGA) collection of primary human tumors (basal versus luminal A breast cancer) (*31*), GSE74912 (CD4 versus CD8 T lymphocytes and B versus T lymphocytes) (*32*), and GSE153237 (control versus CTCF depletion) (*33*). Given the high number of basal and luminal A breast cancer samples in the TCGA collection and the higher variability within these two groups, a more relaxed criteria were used to determine reproducible peaks: Reproducible high-confidence peaks between the biological replicates from the same condition were identified as those that, for the same genomic region, either had high- or low-confidence peaks called in at least 50% of the remaining samples.

Correlation between RNA-seq and ATAC-seq data

To determine the correlation between RNA-seq and ATAC-seq data, differentially accessible ATAC-seq peaks were first annotated with genes if located within 2 kb from the TSS. Therefore, if a gene was annotated with multiple peaks, it might be displayed multiple times. Next, to generate Fig. 5D, the data points were filtered, and only points with $|\log_2(\text{fold change})| > 1$ in both ATAC-seq and RNA-seq datasets were retained. Kernel density estimation (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gaussian_kde.html) was applied to visually emphasize the areas with the highest concentration of the data points. For Fig. 7F, the filtering of points by the threshold of $|\log_2(\text{fold change})| > 1$ was only conducted at the level of the RNA-seq data. For Fig. 7E, differential gene expression data from the two RNA-seq experiments were matched by gene name. Only genes with statistically significant differential gene expression (FDR < 0.05) were included in subsequent analyses. A scatter plot displaying the $\log_2$(fold change) values of those genes and the corresponding kernel density estimation was generated. For all analyses, the Pearson correlation coefficient was calculated.

Gene set enrichment analysis

The enrichment of transcription factor target genes was determined using the GSEA software (*66*, *67*) and the MSigDB C3 TFT molecular signature database collection, which contains targets of regulation by specific transcription factors (www.gsea-msigdb.org/gsea/msigdb/genesets.jsp?collection=TFT). In addition, the MSigDB H

(Hallmark; www.gsea-msigdb.org/gsea/msigdb/genesets.jsp?collection=H) database collection was interrogated to determine the pathways and biological processes enriched in E0771GFP, M11GFP, M12GFP, or ZBTB18-overexpressing cells. GSEA with RNA-seq data was based on FPKM values from each sample. For GSEA with ATAC-seq data, peaks were associated with a gene if located within 2 kb of its TSS. If a gene was associated with more than one ATAC-seq peak, the average FPKM value from all the peaks was used to represent the chromatin state of the gene. Gene set collections, which represent human genes, were converted to mouse orthologs on the basis of the reference human-mouse homology from the Mouse Genome Database (www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt) (*68*) using in-house scripts. GSEA was conducted with parameters set to -permute gene_set -nperm 10000 -set_min 10 -set_max 10000 -metric Signal2Noise -rnd_seed 149 -collapse No_collapse -scoring_scheme weighted -norm meandiv.

Motif enrichment analysis

Motif enrichment analysis was conducted using Homer (*69*), with parameters "-bits -local 2 -size given" and a collection of default Homer motifs, enlarged by the transcription factor binding motifs download from the TRANSFAC database (*70*). The analysis was first run for the regions with significantly increased chromatin accessibility upon ZBTB18 overexpression [log2(fold change) > 1, FDR < 0.05], using the regions with decreased accessibility [log2(fold change) < −1, FDR < 0.05] as the background (-bg flag) for motif enrichment search. The same analysis was repeated in the opposite direction to determine the motifs enriched among the regions showing decreased accessibility, compared to the regions showing increased accessibility.

Reverse transcription quantitative polymerase chain reaction

Total RNA was isolated using the RNeasy Mini Kit (QIAGEN) and reverse-transcribed with the iScript cDNA Synthesis Kit (Bio-Rad) following the manufacturer's procedures. RT-qPCR was performed using the iQ SYBR Green Supermix (Bio-Rad) with the primers reported in table S11. Data were normalized to 18*S* ribosomal RNA or *Gapdh* expression. Relative mRNA levels were calculated using the CFX Manager Software version 3.1 (Bio-Rad) and the comparative CT method.

Immunoblotting

Cell lysates were prepared in radioimmunoprecipitation assay buffer (25 mM tris-HCl, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 0.1% SDS) supplemented with cOmplete Protease Inhibitor Cocktail (Roche) and protease/phosphatase inhibitors (Cell Signaling Technology). Nuclear extracts were prepared as described (*71*). Briefly, 4×10^6 cells were lysed in 450 μl of LB1 buffer [50 mM Hepes-KOH (pH 7.5), 140 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol, 0.5% (v/v) IGEPAL CA-630, 0.25% (v/v) Triton X-100, and protease inhibitors] at 4°C for 10 min. One hundred microliters of the lysate was removed as "whole-cell lysate." The remaining lysate was centrifuged at 2000*g* for 5 min at 4°C, and the supernatant was aliquoted as "cytosolic fraction." The pellet was washed once with 900 μl of LB2 buffer [10 mM tris-HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and protease inhibitors]. The pellet was resuspended in 250 μl of LB3 buffer [10 mM tris-HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% (w/v) sodium deoxycholate, 0.5% (v/v) *N*-lauroylsarcosine, and protease inhibitors] as "nuclear extract." To shear the chromatin DNA, nuclear extracts and whole-cell lysates were sonicated with a Sonic Dismembrator Model 500 (Fisher Scientific) at 15% output for 15 s and four pulses in an ice water bath. Immunoblot was performed as described previously (*48*). The primary antibodies used were anti-ZBTB18 (ZNF238; Proteintech, #12714-1-AP; RRID:AB_2218388; Atlas Antibodies, HPA074019), anti–glyceraldehyde-3-phosphate dehydrogenase (GAPDH; Millipore, #MAB374; RRID:AB_2107445), anti–histone H2B (H2B; Cell Signaling Technology, #12364;; RRID:AB_2714167), anti–phospho-Smad2 (Cell Signaling Technology, #3108, RRID:AB_490941), anti-Smad2/3 (Cell Signaling Technology, #8685; RRID:AB_10889933), and anti-TGFBR2 (Cell Signaling Technology, #79424; RRID:AB_2799933). The secondary antibodies were horseradish peroxidase–conjugated anti-rabbit (#7074; RRID:AB_2099233) or anti-mouse (#7076; RRID:AB_330924) IgG (Cell Signaling Technology).

Immunofluorescence staining

Paraffin-embedded breast cancer tissue arrays (Cooperative Human Tissue Network, CHTN_BrCaProg2 and Novus Biologicals, #NBP2-30212) were deparaffinized in xylene and rehydrated in 100, 95, and 70% ethanol. Antigen retrieval was carried out in ammonium tris-EDTA buffer (pH 9) at 95°C for 20 min. The slides were then incubated in PBS for 5 min, blocked in blocking buffer (5% bovine serum albumin, 0.2% Triton X-100, and 15% donkey serum in HBSS) for 30 min, and incubated with anti-ZBTB18 (1:50; Proteintech, #12714-1-AP; RRID:AB_2218388) overnight at 4°C. The slides were then washed 3 × 5 min in wash buffer (HBSS and 0.3% Triton X-100) and incubated with goat anti-rabbit IgG (H+L)–Alexa Fluor 568 secondary antibody (Thermo Fisher Scientific) for 30 min. The slides were washed three times in wash buffer, followed by a brief wash in water, dried, and mounted using VECTASHIELD Antifade Mounting Medium with 4′,6-diamidino-2-phenylindole (DAPI; Vector Laboratories). The slides were scanned at 20× with an Axioscan Z.1 slide scanner (Zeiss). ZBTB18 intensity scores (0, no staining; 1, medium staining intensity; and 2, high staining intensity) were independently assigned to ZBTB18 expression levels in the nucleus and in the cytoplasm by an investigator blinded to experimental conditions. A total intensity score was obtained by adding the nuclear and cytoplasmic scores. The percentage of nuclear expression was calculated using (nuclear intensity score/total intensity score) × 100. The percentage of cytoplasmic expression was calculated using (cytoplasmic intensity score/total intensity score) × 100.

For immunofluorescence of clonal populations derived from the E0771GFP cell line, single cells were first sorted at one cell per well in 96-well plates using a BD FACSAria fusion flow cytometer. The presence of a single cell/well was confirmed visually, and cells were expanded to ~70% confluency. Cells were then trypsinized, plated in eight-well μ-slides (ibidi), and grown in regular culture medium for 2 days. Immunofluorescence staining and imaging of ZBTB18 were then performed as described above.

ChIP assay

Tumor cells grown to ~85% confluence in 15-cm dishes were fixed for 10 min with 1% formaldehyde, and chromatin was prepared by sonication-mediated shearing. The ChIP assay was performed using the ChIP-IT Express Kit (Active Motif) with anti-ZBTB18 (ZNF238; Proteintech, #12714-1-AP; RRID:AB_2218388) and rabbit IgG (PeproTech, #500-P00; RRID:AB_2722620) following the manufacturer's instructions. To determine the enrichment of ZBTB18 at the *Tgfbr2* promoter, RT-qPCR was performed using the iQ SYBR Green Supermix (Bio-Rad) with primers for the two putative ZBTB18 binding sites, predicted using JASPAR2016 (http://jaspar2016.genereg.net) with 80% threshold (Fig. 6D and table S11). A standard curve was generated with 10-fold dilutions of immunoprecipitated chromatin to test the efficiency of the primers. The enrichment of ZBTB18 and negative control IgG at the *Tgfbr2* promoter is reported as a percentage of the input.

Luciferase reporter assay

The mouse *Tgfbr2* promoter (702 bp, from −650 to +52 relative to the transcription start site) was amplified from genomic DNA of E0771GFP cells by PCR using Q5 Hot Start High-Fidelity 2X Master Mix (NEB). The *Tgfbr2* promoter fragment was cloned into the Kpn I and Hind III restriction sites of the pGL3-Basic firefly luciferase reporter vector (Promega). A mutant *Tgfbr2* promoter where putative ZBTB18 binding sites (5′-[AC]ACATCTG[GT][AC]-3′) are substituted with two Sma I binding sites (5′-CCC GGGCCCGGG-3′) was generated by cloning a gBlock into the Kpn I and Hind III restriction sites of the pGL3-Basic firefly luciferase reporter vector. 293FT cells were plated in 96-well plates at 10^4 cells per well. After 24 hours, cells were cotransfected with 75 ng of luciferase reporter plasmid (pGL3-Basic, pGL3-*Tgfbr2* promoter, or pGL3-mutant *Tgfbr2* promoter); 75 ng of expression plasmid ZBTB18OE, ZBTB18-Nucl, ZBTB18-Cyto, or empty expression vector (EV); and 50 ng of pRL-TK *Renilla* luciferase vector (to normalize for transfection efficiency; Promega) using TransIT-293 reagent in Opti-MEM I. After 24 hours, luciferase activity was measured using the Dual-Glo Luciferase Assay System (Promega). Firefly luciferase values were normalized to *Renilla* luciferase values and protein content and reported relative to the EV condition.

Mining of publicly available datasets

We derived a ZBTB18 target gene signature consisting of the 15 genes most significantly repressed by ZBTB18, based on the sum of the rank metric scores from GSEA of RNA-seq and ATAC-seq data for E0771GFP-ZBTB18OE versus E0771GFP-EV (see Fig. 5I for the list of genes). Kaplan-Meier plotter (www.kmplot.com/analysis/) (*72*–*75*) was then used to determine the correlation of the ZBTB18 target gene signature expression levels with distant metastasis-free survival of patients with breast cancer. We interrogated the data from all patients or only the data from patients with basal, $HER2^+$(human epidermal growth factor receptor 2-positive), luminal A, or luminal B breast cancer subtypes (Prediction Analysis of Microarray 50 subtypes). For gastric and ovarian cancers, the correlation of the ZBTB18 target gene signature expression levels with progression-free survival was determined. For analyses with Kaplan-Meier plotter, plots were generated on the basis of the mean expression of the 15 genes included in the ZBTB18 target gene signature. Cutoff values were determined using the "auto select best cutoff" function, which computes all possible cutoff values between the lower and upper quartiles and identifies the cutoff value with the most significant log-rank test (*76*). In addition, the same datasets were reanalyzed with Cutoff Finder (https://molpathoheidelberg.shinyapps.io/CutoffFinder_v1/) (*77*) using the mean, geometric mean, or sum of *z* scores of the 15 genes included in the ZBTB18 target gene signature as biomarkers. Cutoff values were determined by two different methods: The first method (significance of correlation method), which is similar to the auto select best cutoff function described above, generates Kaplan-Meier plots for each possible cutoff value in the dataset. The optimal cutoff is then defined as the value with the most significant (log-rank test) split (*77*). In the second method [receiver operating character (ROC) curve method], the cutoff value is determined on the basis of the minimal Euclidean distance on the ROC curve of the biomarker point to the left top edge of the diagram (0,1) and thereby identifies the cutoff where the sum of sensitivity and specificity is maximized, independently of patient outcome data (*77*). Hazard ratios and 95% confidence intervals were also calculated.

To determine *ZBTB18* gene expression levels in patients with cancer and adjacent normal tissues, we mined the GSE62944 dataset, which includes RNA-seq data for 9,264 tumor samples and 741 normal samples across 24 cancer types from TCGA. Only solid tumor types with TPM (transcripts per million) data available for at least three samples for both the normal and tumor samples were analyzed. Log_2(TPM + 1) was plotted. *ZBTB18* gene expression in breast cancers with different distant metastasis statuses, lymph node metastasis statuses, or tumor grades were determined by mining the publicly available Mixed Breast (2022-v32) RNA-seq dataset with the R2 Genomics Analysis and Visualization Platform (https://hgserver1.amc.nl/cgi-bin/r2/main.cgi). The following datasets were also analyzed with R2: Tumor Breast Metastatic-Sinn-1108-MAS5.0-u133a [GSE124648 (microarray data)] to compare *ZBTB18* expression levels in metastases versus primary tumors and RNA-seq datasets from TCGA to compare *ZBTB18* expression levels in metastatic versus nonmetastatic cancers.

Analysis of genetic alterations in the protein coding sequence of *ZBTB18*

The cBioPortal for Cancer Genomics (http://cbioportal.org) (*78*, *79*) was used to identify mutations in ZBTB18 among 10,967 samples from patients with cancer. Data were from the 32 studies included in the publicly available TCGA Pan-Cancer Atlas. For analysis of ZBTB18 mutations in E0771GFP and M11GFP cells, genomic DNA sequences were visualized using National Center for Biotechnology Information (NCBI) Genome Workbench and aligned to the mouse RefSeq RNA database (NCBI) by BLAST.

Statistical analysis

Statistics were performed with R studio (version 1.1.453) or with GraphPad Prism (version 8.4.3). Data were analyzed by two-sided unpaired Student's *t* test to compare the means of two independent groups and by one-way analysis of variance (ANOVA) and Tukey's posttest to compare the means of more than two independent groups. Two-way ANOVA was used to compare the mean differences between groups split on two independent variables. If data were not normally distributed, then the Mann-Whitney test was used to compare the means of two independent groups and the Kruskal-Wallis test followed by Dunn's posttest was used to compare the

means of more than two independent groups. Differences among treatments or groups were considered significant at $P < 0.05$ ($^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; and not significant, $^{ns}P > 0.05$).

Supplementary Materials

This PDF file includes:
Figs. S1 to S9
Tables S1, S2, S3, S7 to S11

Other Supplementary Material for this manuscript includes the following:
Tables S4 to S6

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. A. W. Lambert, D. R. Pattabiraman, R. A. Weinberg, Emerging biological principles of metastasis. *Cell* **168**, 670–691 (2017).
2. J. Massagué, K. Ganesh, Metastasis-initiating cells and ecosystems. *Cancer Discov.* **11**, 971–994 (2021).
3. M. Labelle, R. O. Hynes, The initial hours of metastasis: The importance of cooperative host-tumor cell interactions during hematogenous dissemination. *Cancer Discov.* **2**, 1091–1099 (2012).
4. C. L. Chaffer, B. P. San Juan, E. Lim, R. A. Weinberg, EMT, cell plasticity and metastasis. *Cancer Metastasis Rev.* **35**, 645–654 (2016).
5. D. A. Lawson, N. R. Bhakta, K. Kessenbrock, K. D. Prummel, Y. Yu, K. Takai, A. Zhou, H. Eyob, S. Balakrishnan, C.-Y. Wang, P. Yaswen, A. Goga, Z. Werb, Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131–135 (2015).
6. C. Scheel, E. N. Eaton, S.-H. Li, C. L. Chaffer, F. Reinhardt, K.-J. Kah, G. Bell, W. Guo, J. Rubin, A. L. Richardson, R. A. Weinberg, Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926–940 (2011).
7. J.-S. Roe, C.-I. Hwang, T. D. D. Somerville, J. P. Milazzo, E. J. Lee, B. Da Silva, L. Maiorino, H. Tiriac, C. M. Young, K. Miyabayashi, D. Filippini, B. Creighton, R. A. Burkhart, J. M. Buscaglia, E. J. Kim, J. L. Grem, A. J. Lazenby, J. A. Grunkemeyer, M. A. Hollingsworth, P. M. Grandgenett, M. Egeblad, Y. Park, D. A. Tuveson, C. R. Vakoc, Enhancer reprogramming promotes pancreatic cancer metastasis. *Cell* **170**, 875–888.e20 (2017).
8. W. L. Cai, C. B. Greer, J. F. Chen, A. Arnal-Estapé, J. Cao, Q. Yan, D. X. Nguyen, Specific chromatin landscapes and transcription factors couple breast cancer subtype with metastatic relapse to lung or brain. *BMC Med. Genomics* **13**, 33 (2020).
9. Y. Zhang, J. L. Donaher, S. Das, X. Li, F. Reinhardt, J. A. Krall, A. W. Lambert, P. Thiru, H. R. Keys, M. Khan, M. Hofree, M. M. Wilson, O. Yedier-Bayram, N. A. Lack, T. T. Onder, T. Bagci-Onder, M. Tyler, I. Tirosh, A. Regev, J. A. Lees, R. A. Weinberg, 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).
10. J. J. Morrow, I. Bayles, A. P. W. Funnell, T. E. Miller, A. Saiakhova, M. M. Lizardo, C. F. Bartels, M. Y. Kapteijn, S. Hung, A. Mendoza, G. Dhillon, D. R. Chee, J. T. Myers, F. Allen, M. Gambarotti, A. Righi, A. DiFeo, B. P. Rubin, A. Y. Huang, P. S. Meltzer, L. J. Helman, P. Picci, H. H. Versteeg, J. A. Stamatoyannopoulos, C. Khanna, P. C. Scacheri, Positively selected enhancer elements endow osteosarcoma cells with metastatic competence. *Nat. Med.* **24**, 176–185 (2018).
11. P. Rodrigues, S. A. Patel, L. Harewood, I. Olan, E. Vojtasova, S. E. Syafruddin, M. N. Zaini, E. K. Richardson, J. Burge, A. Y. Warren, G. D. Stewart, K. Saeb-Parsy, S. A. Samarajiwa, S. Vanharanta, NF-κB–dependent lymphoid enhancer co-option promotes renal carcinoma metastasis. *Cancer Discov.* **8**, 850–865 (2018).
12. S. Vanharanta, W. Shu, F. Brenet, A. A. Hakimi, A. Heguy, A. Viale, V. E. Reuter, J.-J. Hsieh, J. M. Scandura, J. Massagué, Epigenetic expansion of VHL-HIF signal output drives multiorgan metastasis in renal cancer. *Nat. Med.* **19**, 50–56 (2013).
13. M. Vizoso, H. J. Ferreira, P. Lopez-Serra, F. J. Carmona, A. Martínez-Cardús, M. R. Girotti, A. Villanueva, S. Guil, C. Moutinho, J. Liz, A. Portela, H. Heyn, S. Moran, A. Vidal, M. Martinez-Iniesta, J. L. Manzano, M. T. Fernandez-Figueras, E. Elez, E. Munoz-Couselo, R. Botella-Estrada, A. Berrocal, F. Pontén, J. Van den Oord, W. M. Gallagher, D. T. Frederick, K. T. Flaherty, U. McDermott, P. Lorigan, R. Marais, M. Esteller, Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nat. Med.* **21**, 741–750 (2015).
14. O. G. McDonald, X. Li, T. Saunders, R. Tryggvadottir, S. J. Mentch, M. O. Warmoes, A. E. Word, A. Carrer, T. H. Salz, S. Natsume, K. M. Stauffer, A. Makohon-Moore, Y. Zhong, H. Wu, K. E. Wellen, J. W. Locasale, C. A. Iacobuzio-Donahue, A. P. Feinberg, Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nat. Genet.* **49**, 367–376 (2017).
15. C. L. Chaffer, N. D. Marjanovic, T. Lee, G. Bell, C. G. Kleer, F. Reinhardt, A. C. D'Alessio, R. A. Young, R. A. Weinberg, Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74 (2013).
16. S. Schlesinger, E. Meshorer, Open chromatin, epigenetic plasticity, and nuclear organization in pluripotency. *Dev. Cell* **48**, 135–150 (2019).
17. S. K. Denny, D. Yang, C.-H. Chuang, J. J. Brady, J. S. Lim, B. M. Grüner, S.-H. Chiou, A. N. Schep, J. Baral, C. Hamard, M. Antoine, M. Wislez, C. S. Kong, A. J. Connolly, K.-S. Park, J. Sage, W. J. Greenleaf, M. M. Winslow, Nfib promotes metastasis through a widespread increase in chromatin accessibility. *Cell* **166**, 328–342 (2016).
18. P. Aftimos, M. Oliveira, A. Irrthum, D. Fumagalli, C. Sotiriou, E. N. Gal-Yam, M. E. Robson, J. Ndozeng, A. Di Leo, E. M. Ciruelos, E. de Azambuja, G. Viale, E. D. Scheepers, G. Curigliano, J. M. Bliss, J. S. Reis-Filho, M. Colleoni, M. Balic, F. Cardoso, J. Albanell, C. Duhem, S. Marreaud, D. Romagnoli, B. Rojas, A. Gombos, H. Wildiers, A. Guerrero-Zotano, P. Hall, A. Bonetti, K. F. Larsson, M. Degiorgis, S. Khodaverdi, R. Greil, A. Sverrisdóttir, M. Paoli, E. Seyll, S. Loibl, B. Linderholm, G. Zoppoli, N. E. Davidson, O. T. Johannsson, P. L. Bedard, S. Loi, S. Knox, D. A. Cameron, N. Harbeck, M. L. Montoya, M. Brandao, A. Vingiani, C. Caballero, F. S. Hilbers, L. R. Yates, M. Benelli, D. Venet, M. J. Piccart, Genomic and transcriptomic analyses of breast cancer primaries and matched metastases in AURORA, the Breast International Group (BIG) molecular screening initiative. *Cancer Discov.* **11**, 2796–2811 (2021).
19. M. B. Siegel, X. He, K. A. Hoadley, A. Hoyle, J. B. Pearce, A. L. Garrett, S. Kumar, V. J. Moylan, C. M. Brady, A. E. Van Swearingen, D. Marron, G. P. Gupta, L. B. Thorne, N. Kieran, C. Livasy, E. R. Mardis, J. S. Parker, M. Chen, C. K. Anders, L. A. Carey, C. M. Perou, Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Invest.* **128**, 1371–1383 (2018).
20. A. Le Naour, Y. Koffi, M. Diab, D. Le Guennec, S. Rougé, S. Aldekwer, N. Goncalves-Mendes, J. Talvas, M.-C. Farges, F. Caldefie-Chezet, M.-P. Vasson, A. Rossary, EO771, the first luminal B mammary cancer cell line from C57BL/6 mice. *Cancer Cell Int.* **20**, 328 (2020).
21. K. Sugiura, C. C. Stock, Studies in a tumor spectrum. I. Comparison of the action of methylbis (2-chloroethyl)amine and 3-bis(2-chloroethyl)aminomethyl-4-methoxymethyl -5-hydroxy-6-methylpyridine on the growth of a variety of mouse and rat tumors. *Cancer* **5**, 382–402 (1952).
22. C. Ohtaka-Maruyama, S. Hirai, A. Miwa, A. Takahashi, H. Okado, The 5′-flanking region of the RP58 coding sequence shows prominent promoter activity in multipolar cells in the subventricular zone during corticogenesis. *Neuroscience* **201**, 67–84 (2012).
23. H. Okado, C. Ohtaka-Maruyama, Y. Sugitani, Y. Fukuda, R. Ishida, S. Hirai, A. Miwa, A. Takahashi, K. Aoki, K. Mochida, O. Suzuki, T. Honda, K. Nakajima, M. Ogawa, T. Terashima, J. Matsuda, H. Kawano, M. Kasai, The transcriptional repressor RP58 is crucial for cell-division patterning and neuronal survival in the developing cortex. *Dev. Biol.* **331**, 140–151 (2009).
24. V. M. Tatard, C. Xiang, J. A. Biegel, N. Dahmane, ZNF238 is expressed in postmitotic brain cells and inhibits brain tumor growth. *Cancer Res.* **70**, 1236–1246 (2010).
25. C. Xiang, V. Baubet, S. Pal, L. Holderbaum, V. Tatard, P. Jiang, R. V. Davuluri, N. Dahmane, RP58/ZNF238 directly modulates proneurogenic gene levels and is required for neuronal differentiation and brain expansion. *Cell Death Differ.* **19**, 692–702 (2012).
26. V. Fedele, F. Dai, A. P. Masilamani, D. H. Heiland, E. Kling, A. M. Gätjens-Sanchez, R. Ferrarese, L. Platania, D. Soroush, H. Kim, S. Nelander, A. Weyerbrock, M. Prinz, A. Califano, A. Iavarone, M. Bredel, M. S. Carro, Epigenetic regulation of ZBTB18 promotes glioblastoma progression. *Mol. Cancer Res.* **15**, 998–1011 (2017).
27. S. Bazzocco, H. Dopeso, Á. Martínez-Barriocanal, E. Anguita, R. Nieto, J. Li, E. García-Vidal, V. Maggio, P. Rodrigues, P. G. de Marcondes, S. Schwartz Jr., L. A. Aaltonen, A. Sanchez, J. M. Mariadason, D. Arango, Identification of ZBTB18 as a novel colorectal tumor suppressor gene through genome-wide promoter hypermethylation analysis. *Clin. Epigenetics* **13**, 88 (2021).
28. C. J. Aslakson, F. R. Miller, Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res.* **52**, 1399–1405 (1992).
29. M. Vandromme, C. Gauthier-Rouviere, N. Lamb, A. Fernandez, Regulation of transcription factor localization: Fine-tuning of gene expression. *Trends Biochem. Sci.* **21**, 59–64 (1996).
30. S. E. Pierce, J. M. Granja, M. R. Corces, J. J. Brady, M. K. Tsai, A. B. Pierce, R. Tang, P. Chu, D. M. Feldser, H. Y. Chang, M. C. Bassik, W. J. Greenleaf, M. M. Winslow, LKB1 inactivation modulates chromatin accessibility to drive metastatic progression. *Nat. Cell Biol.* **23**, 915–924 (2021).
31. M. R. Corces, J. M. Granja, S. Shams, B. H. Louie, J. A. Seoane, W. Zhou, T. C. Silva, C. Groeneveld, C. K. Wong, S. W. Cho, A. T. Satpathy, M. R. Mumbach, K. A. Hoadley, A. G. Robertson, N. C. Sheffield, I. Felau, M. A. A. Castro, B. P. Berman, L. M. Staudt, J. C. Zenklusen, P. W. Laird, C. Curtis; Cancer Genome Atlas Analysis Network,

W. J. Greenleaf, H. Y. Chang, The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (2018).
32. M. R. Corces, J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, M. P. Snyder, J. K. Pritchard, A. Kundaje, W. J. Greenleaf, R. Majeti, H. Y. Chang, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **48**, 1193–1203 (2016).
33. B. Xu, H. Wang, S. Wright, J. Hyle, Y. Zhang, Y. Shao, M. Niu, Y. Fan, W. Rosikiewicz, M. N. Djekidel, J. Peng, R. Lu, C. Li, Acute depletion of CTCF rewires genome-wide chromatin accessibility. *Genome Biol.* **22**, 244 (2021).
34. F. R. Greten, S. I. Grivennikov, Inflammation and cancer: Triggers, mechanisms, and consequences. *Immunity* **51**, 27–41 (2019).
35. B. Bakir, A. M. Chiarella, J. R. Pitarresi, A. K. Rustgi, EMT, MET, plasticity, and tumor metastasis. *Trends Cell Biol.* **30**, 764–776 (2020).
36. M. Labelle, S. Begum, R. O. Hynes, Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* **20**, 576–590 (2011).
37. A. Mantovani, P. Allavena, A. Sica, F. Balkwill, Cancer-related inflammation. *Nature* **454**, 436–444 (2008).
38. J. Massague, A. C. Obenauf, Metastatic colonization by circulating tumour cells. *Nature* **529**, 298–306 (2016).
39. K. Aoki, G. Meng, K. Suzuki, T. Takashi, Y. Kameoka, K. Nakahara, R. Ishida, M. Kasai, RP58 associates with condensed chromatin and mediates a sequence-specific transcriptional repression. *J. Biol. Chem.* **273**, 26698–26704 (1998).
40. C. Xiang, K. K. Frietze, Y. Bi, Y. Li, V. Dal Pozzo, S. Pal, N. Alexander, V. Baubet, V. D'Acunto, C. E. Mason, R. V. Davuluri, N. Dahmane, RP58 represses transcriptional programs linked to nonneuronal cell identity and glioblastoma subtypes in developing neurons. *Mol. Cell Biol.* **41**, e0052620 (2021).
41. J. Yang, R. A. Weinberg, Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. *Dev. Cell* **14**, 818–829 (2008).
42. J. Massagué, TGFbeta in cancer. *Cell* **134**, 215–230 (2008).
43. E. Batlle, J. Massagué, Transforming growth factor-β signaling in immunity and cancer. *Immunity* **50**, 924–940 (2019).
44. M. Oft, J. Peli, C. Rudaz, H. Schwarz, H. Beug, E. Reichmann, TGF-β1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. *Genes Dev.* **10**, 2462–2477 (1996).
45. M. Labelle, H. J. Schnittler, D. E. Aust, K. Friedrich, G. Baretton, D. Vestweber, G. Breier, Vascular endothelial cadherin promotes breast cancer progression via transforming growth factor β signaling. *Cancer Res.* **68**, 1388–1397 (2008).
46. L. Borsig, R. Wong, R. O. Hynes, N. M. Varki, A. Varki, Synergistic effects of L- and P-selectin in facilitating tumor metastasis can involve non-mucin ligands and implicate leukocytes as enhancers of metastasis. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 2193–2198 (2002).
47. M. Labelle, S. Begum, R. O. Hynes, Platelets guide the formation of early metastatic niches. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3053–E3061 (2014).
48. H. Jia, J. Janjanam, S. C. Wu, R. Wang, G. Pano, M. Celestine, O. Martinot, H. Breeze-Jones, G. Clayton, C. Garcin, A. Shirinifard, A. M. Zaske, D. Finkelstein, M. Labelle, The tumor cell-secreted matricellular protein WISP1 drives pro-metastatic collagen linearization. *EMBO J.* **38**, e101302 (2019).
49. C. V. Dang, W. M. Lee, Identification of the human c-Myc protein nuclear translocation signal. *Mol. Cell Biol.* **8**, 4048–4054 (1988).
50. D. Kalderon, B. L. Roberts, W. D. Richardson, A. E. Smith, A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499–509 (1984).
51. D. S. Bindels, L. Haarbosch, L. van Weeren, M. Postma, K. E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M. A. Hink, T. W. J. Gadella Jr., mScarlet: A bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods* **14**, 53–56 (2017).
52. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
53. M. Vinci, C. Box, S. A. Eccles, Three-dimensional (3D) tumor spheroid invasion assay. *J. Vi.s Exp.* **e52686**, (2015).
54. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
55. B. Li, C. N. Dewey, RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
56. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
57. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
58. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
59. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21 29 21–21 29 29 (2015).
60. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
61. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
62. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
63. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
64. R. K. Dale, B. S. Pedersen, A. R. Quinlan, Pybedtools: A flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics* **27**, 3423–3424 (2011).
65. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
66. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
67. V. K. Mootha, C. M. Lindgren, K.-F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. N. Hirschhorn, D. Altshuler, L. C. Groop, PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
68. C. J. Bult, J. A. Blake, C. L. Smith, J. A. Kadin, J. E. Richardson; Mouse Genome Database Group, Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* **47**, D801–D806 (2019).
69. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
70. V. Matys, O. V. Kel-Margoulis, E. Fricke, I. Liebich, S. Land, A. Barre-Dirrie, I. Reuter, D. Chekmenev, M. Krull, K. Hornischer, N. Voss, P. Stegmaier, B. Lewicki-Potapov, H. Saxel, A. E. Kel, E. Wingender, TRANSFAC and its module TRANSCompel: Transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**, D108–D 110 (2006).
71. H. Mohammed, C. Taylor, G. D. Brown, E. K. Papachristou, J. S. Carroll, C. S. D'Santos, Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).
72. B. Győrffy, A. Lanczky, A. C. Eklund, C. Denkert, J. Budczies, Q. Li, Z. Szallasi, An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1809 patients. *Breast Cancer Res. Treat* **123**, 725–731 (2010).
73. B. Győrffy, Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput. Struct. Biotechnol. J.* **19**, 4101–4109 (2021).
74. B. Gyorffy, A. Lánczky, Z. Szállási, Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocr. Relat. Cancer* **19**, 197–208 (2012).
75. A. M. Szász, A. Lánczky, A. Nagy, S. Förster, K. Hark, J. E. Green, A. Boussioutas, R. Busuttil, A. Szabó, B. Győrffy, Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1065 patients. *Oncotarget.* **7**, 49322–49333 (2016).
76. A. Lánczky, B. Győrffy, Web-based survival analysis tool tailored for medical research (KMplot): Development and implementation. *J. Med. Internet Res.* **23**, e27633 (2021).
77. J. Budczies, F. Klauschen, B. V. Sinn, B. Győrffy, W. D. Schmitt, S. Darb-Esfahani, C. Denkert, Cutoff finder: A comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLOS ONE* **7**, e51862 (2012).
78. E. Cerami, J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, N. Schultz, The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
79. J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal* **6**, pl1 (2013).

Acknowledgments: We thank K. Lowe for assistance with FACS of single cells, the Hartwell Center for DNA sequencing, Z. Cheng and G. Wu from the St. Jude Center for Applied

Bioinformatics for assistance with bioinformatics analyses, the St. Jude Cell and Tissue Imaging Facility for scanning tissue array slides, S. Miller and the St. Jude Center for Advanced Genome Engineering for gRNA vector design and production, and the St. Jude Vector Core for assistance with lentiviral vector production. **Funding:** This work was supported by NIH grants R01CA245301 (to M.L.) and P30CA021765 and ALSAC, the fundraising and awareness organization for St. Jude Children's Research Hospital. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. **Author contributions:** Conceptualization: R.W. and M.L. Methodology: R.W., A.B.B., B.A.M.-B., W.R., B.X., and M.L. Investigation: R.W., A.B.B., B.A.M.-B., O.K.T., S.T., G.T., S.J., H.J., W.R., O.M., E.C., and R.L. Formal analysis: R.W., A.B.B., W.R., D.F., and M.L. Visualization: R.W., W.R., and M.L. Supervision: B.X. and M.L. Writing (original draft): R.W. Writing (review and editing): R.W., A.B.B., B.A.M.-B., H.J., W.R., O.M., E.C., R.L., and M.L. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. RNA-seq and ATAC-seq data are available at GEO with accession number GSE195944. Publicly available ATAC-seq datasets used in this study are from the TCGA collection of primary human tumors (https://gdc.cancer.gov/about-data/publications/ATACseq-AWG and https://portal.gdc.cancer.gov/), GSE74912 and GSE153237. The research materials in this article can be provided by M.L. pending scientific review and a completed material transfer agreement. Requests for the research materials should be submitted to myriam.labelle@stjude.org.

Submitted 7 April 2022
Accepted 9 December 2022
Published 6 January 2023
10.1126/sciadv.abq3951