LMO2 is critical for early metastatic events in breast cancer

Authors:

Shaheen Sikandar[1,†,*], Jane Antony[1,†], Gunsagar S. Gulati[1,†], Angera H. Kuo[1], William Hai Dang Ho[1], Soumyashree Das[2], Chloé B. Steen[3], Thiago Almeida Pereira[1], Dalong Qian[1], Philip A. Beachy[1], Fredrick Dirbas[4], Kristy Red-Horse[1,2], Terence H. Rabbitts[5], Jean Paul Thiery[6], Aaron M. Newman[1,7,‡], and Michael F. Clarke[1,8,‡,*]

Affiliations:

[1]Institute for Stem Cell Biology and Regenerative Medicine, 265 Campus Drive, School of Medicine, Stanford, CA-94305.

[2]Department of Biology, Stanford University, Stanford, CA 94305, USA.

[3]Division of Oncology, Department of Medicine, Stanford Cancer Institute, Stanford University, Stanford, CA, USA.

[4]Department of Surgery, Stanford University School of Medicine, Stanford Cancer Institute, 875 Blake Wilbur Drive, Rm CC2235, Stanford, CA, 94305, USA.

[5]Institute of Cancer Research, Division of Cancer Therapeutics, London, SM2 5NG, UK.

[6]Bioland Laboratory, Guangzhou Regenerative Medicine and Health, Guangzhou, China.

[7]Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA.

[8]Department of Medicine, Stanford University, Stanford, CA 94305, USA.

[†]These authors contributed equally.

[‡]Co-senior authors

[*]Correspondence to: mfclarke@stanford.edu, ssikanda@ucsc.edu

One sentence summary:

LMO2 modulates STAT3 signaling in breast cancer metastasis.

SUMMARY

Metastasis is responsible for the majority of breast cancer-related deaths, however identifying the cellular determinants of metastasis has remained challenging. Here, we identified a minority population of immature *THY1*$^{+}$/*VEGFA*$^{+}$ tumor epithelial cells in human breast tumor biopsies that display angiogenic features and are marked by the expression of the oncogene, *LMO2*. Higher abundance of *LMO2*$^{+}$ basal cells correlated with tumor endothelial content and predicted poor distant recurrence-free survival in patients. Using *MMTV-PyMT/Lmo2*^{CreERT2} mice, we demonstrated that *Lmo2* lineage-traced cells have a higher propensity to metastasize. LMO2 knockdown in human breast tumors reduced lung metastasis by impairing intravasation, leading to a reduced frequency of circulating tumor cells. Mechanistically, we find that LMO2 binds to STAT3 and is required for STAT3 activation by TNFα and IL6. Collectively, our study identifies a population of metastasis-initiating cells with angiogenic features and establishes the LMO2-STAT3 signaling axis as a therapeutic target in breast cancer metastasis.

INTRODUCTION

While significant progress has been made to treat early-stage breast cancer, treatment options and outcomes for metastatic breast cancer have been largely unchanged in a decade (Esposito et al., 2021; Siegel et al., 2011; Siegel et al., 2021). In order to improve outcomes for breast cancer patients, it is critical to identify and elucidate signaling pathways active in metastatic cells. However, it has been difficult to pinpoint cancer cell populations involved in metastasis as they represent a transient state (Lu and Kang, 2019). Previous studies employing lineage tracing and cell surface marker profiling have implicated distinct subsets of tumor epithelial cells in breast cancer metastasis, primarily using lineage markers such as E-cadherin (Beerling et al., 2016, Padmanaban et al., 2019), N-cadherin (Li et al., 2020) and S100a4 (Fischer et al., 2015). Recent studies have also suggested that metastatic cells display hybrid features of both epithelial and mesenchymal lineages (Kröger et al., 2019; Pastushenko et al.,

2021). This has led to a debate in the field about the precise molecular identity of metastasis-initiating cells (Lu and Kang, 2019; Shen and Kang, 2019; Ye et al., 2017).

Our previous work has demonstrated that in breast cancer, minority populations of phenotypically immature cells in the tumor are enriched in tumor-initiating potential and metastasis (Al-Hajj et al., 2003; Liu et al., 2010; Sikandar et al., 2017). Recent advances in single-cell technologies have revealed complex transcriptional landscapes in human tumors and enabled precise molecular characterization of these minority cell populations (Lawson et al., 2018). However, the functional and clinical significance of these populations remains to be elucidated (Lawson et al., 2018; Tanay and Regev, 2017). To understand the transcriptional heterogeneity in breast cancer, we performed single-cell RNA sequencing (scRNA-seq) in primary patient samples and developed a novel computational method that can predict immature cell populations *in silico* (Gulati et al., 2020). Using our scRNA-seq data, bulk tumor expression deconvolution, lineage tracing, and functional assays, we have now identified a clinically relevant population of metastasis-initiating cells that express the hematopoietic transcription factor and T-cell oncogene, LMO2. Here, we mechanistically define the role of LMO2 in breast cancer metastasis by its association with tumor vasculature and identify LMO2 as a previously unknown regulator of STAT3 signaling in breast cancer.

RESULTS

LMO2 is expressed in a minority population of immature *THY1*$^+$/*VEGFA*$^+$ human breast cancer cells.

To understand the substructure of the epithelial populations in breast cancer, we started by analyzing scRNA-seq profiles (Gulati et al., 2020) of human breast tumor epithelial cells from patients with triple-negative (n = 5) or estrogen receptor positive (ER$^+$) breast cancer (n = 13). We identified a minority population of *THY1*$^+$ cells that were largely restricted to the basal compartment, comprising 11% of all basal cells (**Fig. S1A, Table S1**). Moreover, within this subset, 33% of cells expressed *VEGFA* (**Fig. S1A**). We were struck by this combination since THY1$^+$ cells are enriched in reconstitution potential in

the normal mammary gland (Lobo et al., 2018) and tumorigenic potential in mouse tumors (Cho et al., 2008) and VEGFA is a pro-angiogenic factor linked to tumor growth and distant metastasis (Mercurio et al., 2005; Zhao et al., 2015). To determine whether *THY1*$^{+}$/*VEGFA*$^{+}$ cells represent a potential immature cell population, we applied CytoTRACE, a computational framework for predicting cellular differentiation status on the basis of single-cell transcriptional diversity (Gulati et al., 2020). We found that relative to other basal cells, *THY1*$^{+}$/*VEGFA*$^{+}$ cells are predicted to be significantly less differentiated, suggesting a role for this population in tumor growth or metastasis (**Fig. 1A**).

To identify potential molecular regulators within this population, we next searched for genes with expression patterns that overlap *THY1* and *VEGFA* expression in our dataset. Intriguingly, we found that *LMO2*, a hematopoietic stem cell regulator (Yamada et al., 1998) and T-cell oncogene (Larson, 1995), was among the top five hits (**Fig. 1B, Table S2**). *LMO2* also marked *THY1*$^{+}$/*VEGFA*$^{+}$ cells in an independent scRNA-seq atlas of triple-negative human breast tumors (Kim et al., 2018), corroborating this result (**Fig. 1C**). Analysis of the *LMO2*$^{+}$ basal epithelial subset showed that these cells not only express *THY1* and epithelial cytokeratins (**Fig. 1D**), but also display a coherent gene expression program significantly enriched in angiogenesis genes, including *VEGFA* and *S100A4* (**Fig. 1E, Table S3**).

We next measured the relative abundance of distinct endothelial, immune, stromal, and epithelial populations in human breast tumors with respect to *LMO2*$^{+}$ basal cells. As *LMO2* is expressed in myriad cell types, including immune, stromal, and endothelial cells, the expression of the gene is insufficient to distinguish cell types. Therefore, we defined unique transcriptional signatures for various niche and breast epithelial cells from our scRNA-seq data and utilized CIBERSORTx, a deconvolution approach, to calculate the cellular composition of bulk RNA admixtures from breast cancer clinical cohorts (Newman et al., 2019) (**Methods**). In line with our previous results, we observed a striking correlation between the abundance of *LMO2*$^{+}$ basal cells and

endothelial cell content imputed in 508 breast tumors (Esserman et al., 2012) (r = 0.45; $P < 2 \times 10^{-16}$; **Fig. 1F**).

Human *LMO2*$^{+}$ basal cells are associated with poor outcomes in breast cancer patients.

Deconvolution of an additional 3,024 human breast tumors from three clinical cohorts (Curtis et al., 2012; TCGA, 2012) revealed that basal *LMO2*$^{+}$ cells are more abundant in 'Basal' breast cancer subtypes which correlate with more aggressive breast cancers as compared to other PAM50 classes (Perou et al., 2000) (**Fig. S1B**). We also found a significant increase in basal *LMO2*$^{+}$ cells with worsening clinical grade and stage of the tumor (**Fig. S1C, D**), suggesting that *LMO2*$^{+}$ cells increase with tumor progression. Importantly, higher levels of *LMO2*$^{+}$ basal cells were significantly associated with inferior distant recurrence-free survival (**Fig. 1G**), independent of estrogen receptor status. These data link the abundance of *LMO2*$^{+}$ basal epithelial cells with more aggressive breast tumors and distant metastasis.

***Lmo2* lineage-traced cells have a higher propensity to metastasize.**

To experimentally verify our *in silico* findings, we began by employing the CreERT2 system (Rios et al., 2014; van Amerongen et al., 2012; Van Keymeulen et al., 2011) to delineate the fate of epithelial cells that have expressed *LMO2*$^{+}$ in breast tumors. We obtained *Lmo2*^{CreERT2} mice (Forster, Drynan, Pannell, Rabbitts *in preparation*) and crossed them to *Rosa26*^{mTmG} reporter and *MMTV-PyMT* tumor mice to generate triple-transgenic *Lmo2*^{CreERT2}/*Rosa26*^{mTmG}/*MMTV-PyMT* mice, which we termed *Lmo2-PyMT* (**Fig. 2A**). *MMTV-PyMT* tumors are an aggressive luminal subtype of breast cancer (Herschkowitz et al., 2007) that metastasize to the lungs (Guy et al., 1992) and have been extensively used to explore the cellular underpinnings of breast cancer metastasis (Beerling et al., 2016; Fischer et al., 2015; Padmanaban et al., 2019; Pastushenko et al., 2018). As *Lmo2* is expressed in other cells such as stromal and endothelial cells (Gratzinger et al., 2009), we orthotopically transplanted lineage depleted (CD45^{-}/CD31^{-}/Ter119^{-}) tumor cells from TdTomato-fluorescent *Lmo2-PyMT* into non-fluorescent BL6 mice to clearly assess the contribution of *Lmo2* lineage-traced breast cancer cells from

the tumor. After tumors were formed, we pulsed the mice with tamoxifen to induce expression of GFP in *Lmo2*-expressing cells (**Fig. 2B**). At 48h post-pulse, we verified that expression of *Lmo2* was enriched in the transplanted GFP$^+$ cancer cells (**Figs. 2C**, **S2,** and **S3A**). FACS quantification demonstrated that GFP$^+$ cells represented a minor fraction of all tumor cells and expressed the epithelial marker, EpCAM (**Fig. 2C**).

To assess the population dynamics of *Lmo2* lineage-traced cells, we plated TdTomato$^+$ tumor cells from *Lmo2-PyMT* mice in 3D organoid assays and pulsed the organoids with 4-hydroxytamoxifen. Consistent with the *in vivo* model, lineage-traced GFP$^+$ cells comprised a minority of tumor organoids (~2%) 7 days post-pulse. This percentage was unchanged even after 4 weeks in culture, suggesting similar proliferative capacity between GFP$^+$ and TdTomato$^+$ cells (**Fig. S3B**). We confirmed this by plating sorted GFP$^+$ and TdTomato$^+$ cells in 3D organoid cultures and showing that both populations formed organoids at similar frequencies (**Fig. S3C**).

To determine whether *Lmo2*$^+$ cells co-associate with endothelial cells, as predicted *in silico* (**Fig. 1F**), we stained vasculature with endomucin and visualized their co-localization with 3D imaging. We found that *Lmo2* lineage-traced cells not only resided near tumor blood vessels (**Fig. 2E**) but surprisingly ~20% showed co-localization with tumor vasculature and appeared to be incorporated into the tumor vasculature (**Fig. 2E** and **S3D**).

Given that abundance of *LMO2*$^+$ cells in patients predicts distant recurrence-free survival (**Fig. 1G**) and *Lmo2* lineage-traced cells reside closer to tumor vasculature, we next tested whether *Lmo2*$^+$ cells have metastatic capabilities. As dissemination of metastatic cells occurs continuously during tumor growth, to lineage-trace tumor cells expressing *Lmo2*, we pulsed *Lmo2-PyMT* mice with tamoxifen 2-3 times per week once the tumors were palpable and continued until tumor endpoint **(see Methods**; **Fig. 3F**). At the end of the experiment, we found that in the primary tumor only 10-15% of tumor cells were GFP$^+$ (**Fig. 2G**). Surprisingly, even though the tumor was majority TdTomato$^+$, the lungs had a disproportionately higher number of GFP$^+$ metastases,

several of which were also larger than the TdTomato$^+$ metastases (P = 0.034, Wilcoxon signed-rank unpaired test) (**Fig. 2H**). These data suggest that *Lmo2* lineage-traced cells have a higher propensity to form metastases in the *PyMT* mice and is consistent with our findings in human breast cancer patients (**Fig. 1G**). Furthermore, a subset of GFP tumor cells did not remain Lmo2 positive (**Fig. S3E**), suggesting that expression of Lmo2 in some cells represents a transient state, in agreement with previous studies linking transient cell states to metastases (Pastushenko et al., 2018).

***LMO2* knockdown abrogates lung metastasis in human breast cancer models.**

To understand the functional role of LMO2 in human breast cancer, we knocked down *LMO2* expression in MDA-MB-468 cells using two independent shRNA vectors tagged with a GFP reporter (**Fig. S4A-C**). We then implanted the cells orthotopically in immunodeficient mice (**Fig. 3A** and **S4D**). In contrast to a previous report (Liu et al., 2016), knockdown of *LMO2* did not affect primary tumor growth (**Fig. 3B**) or proliferation *in vitro* (**Fig. S5A**). Nevertheless, *LMO2*-knockdown tumors had significantly fewer lung metastases relative to control (P = 0.003, ANOVA; **Fig. 2C**). Moreover, *LMO2*-knockdown in tumor-bearing mice led to a significantly reduced number of circulating tumor cells compared to control mice (P < 0.0001, ANOVA; **Fig. 2D**), implicating LMO2 in tumor cell shedding, a key step in metastasis initiation. To extend our findings to more clinically relevant models, we used patient-derived xenograft (PDX) models previously generated in our lab (Sikandar et al., 2017). Consistent with our MDA-MB-468 studies, knockdown of LMO2 dramatically decreased metastasis to the lung in three different PDX models of breast cancer (**Fig. 3E, F**), but did not significantly impact tumor growth (**Fig. S5B-D**).

To better understand how LMO2 affects metastasis, we rigorously studied the effects of LMO2 knockdown *in vitro* in MDA-MB-468 cells. Knockdown of LMO2 showed significant impairment in the ability of cancer cells to migrate across transwells and invade through a 3D hydrogel matrix (**Fig. S6A**, **B**). Importantly, since *LMO2*$^+$ epithelial cells associated with endothelial cells in patient samples, we tested whether knockdown of *LMO2* decreased this association in co-culture assays. We found that in 3D co-

culture assays with human vascular endothelial cells (HUVECs), LMO2 knockdown significantly impacted incorporation of cancer cells into HUVEC tubes (**Fig. S6C**). To confirm that the effects of knockdown were specific to LMO2, we overexpressed LMO2 in cells with shRNA targeting the 3'UTR. We found that all phenotypes of migration (**Fig. 3G**), invasion (**Fig. 3H**), and incorporation into the vasculature *in vitro* (**Fig. 3I**) could be rescued by overexpression of LMO2 in LMO2-deficient cells. Lastly, to test whether LMO2 is required after metastatic cells enter circulation, we injected control and LMO2 knockdown cells into the tail vein. We found that LMO2 knockdown did not significantly impact the formation of lung metastases when cells were directly injected in the tail vein, suggesting that LMO2 is critical for the initial dissemination of cancer cells from the tumor, but not extravasation and formation of metastatic foci (**Fig. S6D**).

RNA sequencing identifies LMO2 as a regulator of IL6-JAK-STAT3 signaling.

To elucidate the molecular function of LMO2 in breast cancer cells, we performed bulk RNA sequencing of MDA-MB-468 cells after transfection with control and LMO2 shRNA vectors (**Fig. 4A**). Among the top 50 genes downregulated after LMO2 knockdown were genes previously implicated in metastasis, such as *BMP2* (Bach et al., 2018; Huang et al., 2017; Wang et al., 2017), *LGR6* (Leushacke and Barker, 2012; Ruan et al., 2019), *EGR4* (Matsuo et al., 2014), *TDO2* (D'Amato et al., 2015) and *S100A4* (Boye and Maelandsmo, 2010; Garrett et al., 2006; Helfman et al., 2005) (**Fig. 4A, Table S4**). Using Gene Set Enrichment Analysis (GSEA) (Mootha et al., 2003; Subramanian et al., 2005), we found that inflammatory pathways, such as TNFα via NF-kB signaling, IL6-JAK-STAT3 signaling, and IFNγ response, were significantly downregulated in LMO2 knockdown as compared to control conditions (**Fig. 4B**). To confirm our findings in primary patient samples we performed single-sample GSEA in our scRNA-seq data set as well as a larger published dataset of primary human breast cancer cells (Kim et al., 2018). We found that IL6-JAK-STAT3 signaling was significantly enriched in *LMO2*$^{+}$ versus *LMO2*$^{-}$ single cells (**Fig. 4C**) compared to other pathways (**Fig. S7**). In the hematopoietic system, LMO2 is an adaptor protein that facilitates formation of functional protein complexes which then activate transcription of downstream targets (Chambers and Rabbitts, 2015). Hence, we asked whether LMO2 may similarly behave as a

bridging molecule to drive downstream signaling in breast epithelial cells. Using proximity ligation assays, we found that LMO2 had a significantly high binding affinity to STAT3, but not to NF-kB, further confirming our pathway analysis (**Fig. 4D**).

LMO2 is required for STAT3 activation by IL6 and TNFα.

To demonstrate specificity and functional significance of the LMO2-STAT3 interaction, we first showed that LMO2 knockdown significantly reduced LMO2-STAT3 binding ($P < 0.0001$, ANOVA; **Fig. 5A**). We also confirmed the LMO2-STAT3 interaction using co-immunoprecipitation assays (Co-IP) of LMO2 with STAT3 (**Fig. 5B**) and, a reverse Co-IP of STAT3 with LMO2 (**Fig. 5C**). In breast cancer, STAT3 is activated by cytokines, such as IL6 (Zhong et al., 1994), TNFα (De Simone et al., 2015), IFNα (Beadling et al., 1994; Cho et al., 1996; Darnell et al., 1994) and IFNγ (Darnell et al., 1994; Will et al., 1996), as well as receptor tyrosine kinases such as EGFR (Kim et al., 2012; Zhao et al., 2020), leading to phosphorylation of STAT3. Dimerization of pSTAT3 and translocation to the nucleus activates transcription of downstream target genes involved in several processes, including metastasis. To understand whether the STAT3-LMO2 interaction has an effect on downstream STAT3 signaling, we used a STAT3-luciferase reporter assay. We stimulated control or LMO2 knockdown cells with IL6, TNFα, IFNγ, IFNα, and EGF. We found that cells with knockdown of LMO2 were unable to induce transcription of the STAT3-luciferase reporter when treated with IL6 and TNFα as compared to control (**Fig. 5D**), but STAT3-luciferase was activated by IFNγ, IFNα, and EGFR treatment. This suggests that LMO2 function in breast cancer cells is specific to activation of STAT3 signaling through IL6 and TNFα. On a molecular level, we found that knockdown of LMO2 significantly reduced STAT3 phosphorylation at Tyr705, which is required for its dimerization and transcriptional activity (**Fig. 5E and Fig. S8**). To understand how LMO2 regulates phosphorylation of STAT3, we examined the interaction of STAT3 with its upstream activator JAK2 and its cytoplasmic inhibitor PIAS3. Knockdown of LMO2 decreased the interaction of STAT3 with JAK2 (**Fig. 5F**) and allowed for increased interaction with its inhibitor, PIAS3 (**Fig. 5G**). This suggests that LMO2 works as an adaptor protein in the cytoplasm to stabilize the STAT3-JAK2 interaction, thereby allowing efficient phosphorylation and activation of STAT3 while

simultaneously preventing its negative regulation by PIAS3 (**Fig. 5H**). This LMO2-mediated control of a core inflammatory response pathway could enable cancer cells to rapidly transition between cellular phenotypes required for metastasis and represents a therapeutic vulnerability that could be targeted.

DISCUSSION

Efficient metastasis of tumor cells requires transition from a proliferative to an invasive state and back to a proliferative state at a distant site (Beerling et al., 2016). Previous studies using mouse tumor models have demonstrated the requirement of a basal epithelial program in metastasis (Cheung et al., 2013; Padmanaban et al., 2019) and showed that hybrid epithelial-mesenchymal states (Beerling et al., 2016; Kröger et al., 2019; Nieto et al., 2016) in metastasis express angiogenic factors (Pastushenko et al., 2018). Here, we have identified a population of *THY1*$^{+}$/*VEGFA*$^{+}$ human basal epithelial cells with higher transcriptional diversity that is marked by transient expression of *LMO2*. Moreover, we demonstrate that *Lmo2* lineage-traced epithelial cells have a higher propensity to form lung metastases. Moreover, knockdown of LMO2 decreases lung metastasis in multiple tumor models of human breast cancer by affecting multiple steps during intravasation. It is important to note that only a subset of *Lmo2* lineage-traced cells show vascular phenotypes, suggesting specific epigenetic regulation that is activated in the presence of TNFα and IL6 from the microenvironment. Our observations highlight a heterogenous, cancer-cell-intrinsic response to the microenvironment while previous studies have demonstrated that there is a reciprocal effect of cancer cells on the tumor microenvironment with recruitment of macrophages and cross-talk with tumor endothelial cells during metastasis (Borriello et al., 2020).

LMO2 has been extensively studied in hematological malignancies and is well-established as a transcriptional adaptor protein (Chambers and Rabbitts, 2015). Recent studies have attempted to understand the role of LMO2 in breast cancer (Hu et al., 2021; Liu et al., 2016; Liu et al., 2017) but have suffered from contradictory results, were limited to cell lines, and did not attribute LMO2 to any particular tumor cell population. We demonstrate that LMO2 is a previously unidentified binding partner of

STAT3 in breast cancer cells and modulates STAT3 signaling in response to IL6 and TNFα. We speculate that the expression of LMO2 provides the necessary threshold to stabilize STAT3 signaling, which in turn enables the tumor cells to enter a transient metastatic state (Wendt et al., 2014) and escape the primary tumor. STAT3 signaling is involved in a number of processes and its targets may be defined in unison with other contextual signals such as inflammation. Several studies have linked low chronic inflammation in cancer to metastasis (Joyce and Pollard, 2009; Liu et al., 2015). We speculate that LMO2 is a critical molecular link between these processes and define a novel function for LMO2 in breast cancer metastasis. The development of new methods targeting adaptor proteins (Wang et al., 2020) and small molecules that disrupt the LMO2-STAT3 axis (Milton-Harris et al., 2020) could provide novel therapeutic strategies to modulate STAT3 signaling and inhibit metastatic colonization in breast cancer.

REFERENCES AND NOTES

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A *100*, 3983-3988.

Bach, D.H., Park, H.J., and Lee, S.K. (2018). The Dual Role of Bone Morphogenetic Proteins in Cancer. Mol Ther Oncolytics *8*, 1-13.

Beadling, C., Guschin, D., Witthuhn, B.A., Ziemiecki, A., Ihle, J.N., Kerr, I.M., and Cantrell, D.A. (1994). Activation of JAK kinases and STAT proteins by interleukin-2 and interferon alpha, but not the T cell antigen receptor, in human T lymphocytes. EMBO J *13*, 5605-5615.

Beerling, E., Seinstra, D., de Wit, E., Kester, L., van der Velden, D., Maynard, C., Schäfer, R., van Diest, P., Voest, E., van Oudenaarden, A.*, et al.* (2016). Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity. Cell Reports *14*, 2281-2288.

Borriello, L., Karagiannis, G.S., Duran, C.L., Coste, A., Oktay, M.H., Entenberg, D., and Condeelis, J.S. (2020). The role of the tumor microenvironment in tumor cell intravasation and dissemination. European Journal of Cell Biology *99*, 151098.

Boye, K., and Maelandsmo, G.M. (2010). S100A4 and metastasis: a small actor playing many roles. Am J Pathol *176*, 528-535.

Chambers, J., and Rabbitts, T.H. (2015). LMO2 at 25 years: a paradigm of chromosomal translocation proteins. Open Biol *5*, 150062-150062.

Cheung, K.J., Gabrielson, E., Werb, Z., and Ewald, A.J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. Cell *155*, 1639-1651.

Cho, R.W., Wang, X., Diehn, M., Shedden, K., Chen, G.Y., Sherlock, G., Gurney, A., Lewicki, J., and Clarke, M.F. (2008). Isolation and Molecular Characterization of Cancer Stem Cells in MMTV-Wnt-1 Murine Breast Tumors. STEM CELLS *26*, 364-371.

Cho, S.S., Bacon, C.M., Sudarshan, C., Rees, R.C., Finbloom, D., Pine, R., and O'Shea, J.J. (1996). Activation of STAT4 by IL-12 and IFN-alpha: evidence for the involvement of ligand-induced tyrosine and serine phosphorylation. J Immunol *157*, 4781-4789.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y.*, et al.* (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346-352.

D'Amato, N.C., Rogers, T.J., Gordon, M.A., Greene, L.I., Cochrane, D.R., Spoelstra, N.S., Nemkov, T.G., D'Alessandro, A., Hansen, K.C., and Richer, J.K. (2015). A TDO2-AhR signaling axis facilitates anoikis resistance and metastasis in triple-negative breast cancer. Cancer Res *75*, 4651-4664.

Darnell, J.E., Jr., Kerr, I.M., and Stark, G.R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science *264*, 1415-1421.

De Simone, V., Franze, E., Ronchetti, G., Colantoni, A., Fantini, M.C., Di Fusco, D., Sica, G.S., Sileri, P., MacDonald, T.T., Pallone, F.*, et al.* (2015). Th17-type cytokines, IL-6 and TNF-alpha synergistically activate STAT3 and NF-kB to promote colorectal cancer cell growth. Oncogene *34*, 3493-3503.

Esposito, M., Ganesan, S., and Kang, Y. (2021). Emerging strategies for treating metastasis. Nature Cancer *2*, 258-270.

Esserman, L.J., Berry, D.A., Cheang, M.C., Yau, C., Perou, C.M., Carey, L., DeMichele, A., Gray, J.W., Conway-Dorsey, K., Lenburg, M.E.*, et al.* (2012). Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). Breast Cancer Res Treat *132*, 1049-1062.

Fischer, K.R., Durrans, A., Lee, S., Sheng, J., Li, F., and Wong, S.T. (2015). Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature *527*.

Garrett, S.C., Varney, K.M., Weber, D.J., and Bresnick, A.R. (2006). S100A4, a mediator of metastasis. J Biol Chem *281*, 677-680.

Gratzinger, D., Zhao, S., West, R., Rouse, R.V., Vogel, H., Gil, E.C., Levy, R., Lossos, I.S., and Natkunam, Y. (2009). The transcription factor LMO2 is a robust marker of vascular endothelium and vascular neoplasms and selected other entities. Am J Clin Pathol *131*, 264-278.

Gulati, G.S., Sikandar, S.S., Wesche, D.J., Manjunath, A., Bharadwaj, A., Berger, M.J., Ilagan, F., Kuo, Angera H., Hsieh, R.W., Cai, S.*, et al.* (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. Science *367*, 405-411.

Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Molecular and Cellular Biology *12*, 954-961.

Helfman, D.M., Kim, E.J., Lukanidin, E., and Grigorian, M. (2005). The metastasis associated protein S100A4: role in tumour progression and metastasis. Br J Cancer *92*, 1955-1958.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S.*, et al.* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biology *8*, R76.

Hu, A., Hong, F., Li, D., Xie, Q., Chen, K., Zhu, L., and He, H. (2021). KDM3B-ETF1 fusion gene downregulates LMO2 via the WNT/β-catenin signaling pathway, promoting metastasis of invasive ductal carcinoma. Cancer Gene Therapy.

Huang, P., Chen, A., He, W., Li, Z., Zhang, G., Liu, Z., Liu, G., Liu, X., He, S., Xiao, G.*, et al.* (2017). BMP-2 induces EMT and breast cancer stemness through Rb and CD44. Cell Death Discov *3*, 17039.

Joyce, J.A., and Pollard, J.W. (2009). Microenvironmental regulation of metastasis. Nature reviews Cancer *9*, 239-252.

Kim, C., Gao, R., Sei, E., Brandt, R., Hartman, J., Hatschek, T., Crosetto, N., Foukakis, T., and Navin, N.E. (2018). Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. Cell *173*, 879-893.e813.

Kim, S.M., Kwon, O.J., Hong, Y.K., Kim, J.H., Solca, F., Ha, S.J., Soo, R.A., Christensen, J.G., Lee, J.H., and Cho, B.C. (2012). Activation of IL-6R/JAK1/STAT3 signaling induces de novo resistance to irreversible EGFR inhibitors in non-small cell lung cancer with T790M resistance mutation. Mol Cancer Ther *11*, 2254-2264.

Kröger, C., Afeyan, A., Mraz, J., Eaton, E.N., Reinhardt, F., Khodor, Y.L., Thiru, P., Bierie, B., Ye, X., Burge, C.B.*, et al.* (2019). Acquisition of a hybrid E/M state is

essential for tumorigenicity of basal breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America *116*, 7353-7362.
Larson, R.C., Osada, H., Larson, T.A., Lavenir, I. & Rabbitts, T.H (1995). The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice. Oncogene *11*, 853-862.
Lawson, D.A., Kessenbrock, K., Davis, R.T., Pervolarakis, N., and Werb, Z. (2018). Tumour heterogeneity and metastasis at single-cell resolution. Nat Cell Biol *20*, 1349-1360.
Leushacke, M., and Barker, N. (2012). Lgr5 and Lgr6 as markers to study adult stem cell roles in self-renewal and cancer. Oncogene *31*, 3009-3022.
Li, Y., Lv, Z., Zhang, S., Wang, Z., He, L., Tang, M., Pu, W., Zhao, H., Zhang, Z., Shi, Q.*, et al.* (2020). Genetic Fate Mapping of Transient Cell Fate Reveals N-Cadherin Activity and Function in Tumor Metastasis. Dev Cell *54*, 593-607.e595.
Liu, H., Patel, M.R., Prescher, J.A., Patsialou, A., Qian, D., Lin, J., Wen, S., Chang, Y.-F., Bachmann, M.H., Shimono, Y.*, et al.* (2010). Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. Proceedings of the National Academy of Sciences *107*, 18115-18120.
Liu, J., Lin, P.C., and Zhou, B.P. (2015). Inflammation fuels tumor progress and metastasis. Curr Pharm Des *21*, 3032-3040.
Liu, Y., Huang, D., Wang, Z., Wu, C., Zhang, Z., Wang, D., Li, Z., Zhu, T., Yang, S., and Sun, W. (2016). LMO2 attenuates tumor growth by targeting the Wnt signaling pathway in breast and colorectal cancer. Sci Rep *6*, 36050-36050.
Liu, Y., Wang, Z., Huang, D., Wu, C., Li, H., Zhang, X., Meng, B., Li, Z., Zhu, T., Yang, S.*, et al.* (2017). LMO2 promotes tumor cell invasion and metastasis in basal-type breast cancer by altering actin cytoskeleton remodeling. Oncotarget *8*, 9513-9524.
Lobo, N.A., Zabala, M., Qian, D., and Clarke, M.F. (2018). Serially transplantable mammary epithelial cells express the Thy-1 antigen. Breast Cancer Res *20*, 121-121.
Lu, W., and Kang, Y. (2019). Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis. Dev Cell *49*, 361-374.
Matsuo, T., Dat le, T., Komatsu, M., Yoshimaru, T., Daizumoto, K., Sone, S., Nishioka, Y., and Katagiri, T. (2014). Early growth response 4 is involved in cell proliferation of small cell lung cancer through transcriptional activation of its downstream genes. PloS one *9*, e113606.
Mercurio, A.M., Lipscomb, E.A., and Bachelder, R.E. (2005). Non-Angiogenic Functions of VEGF in Breast Cancer. Journal of Mammary Gland Biology and Neoplasia *10*, 283-290.
Milton-Harris, L., Jeeves, M., Walker, S.A., Ward, S.E., and Mancini, E.J. (2020). Small molecule inhibits T-cell acute lymphoblastic leukaemia oncogenic interaction through conformational modulation of LMO2. Oncotarget; Vol 11, No 19.
Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267-273.
Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D.*, et al.* (2019). Determining

cell type abundance and expression from bulk tissues with digital cytometry. Nature Biotechnology *37*, 773-782.
Nieto, M.A., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2016). Emt: 2016. Cell *166*, 21-45.
Padmanaban, V., Krol, I., Suhail, Y., Szczerba, B.M., Aceto, N., Bader, J.S., and Ewald, A.J. (2019). E-cadherin is required for metastasis in multiple models of breast cancer. Nature *573*, 439-444.
Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S.*, et al.* (2018). Identification of the tumour transition states occurring during EMT. Nature *556*, 463-468.
Pastushenko, I., Mauri, F., Song, Y., de Cock, F., Meeusen, B., Swedlund, B., Impens, F., Van Haver, D., Opitz, M., Thery, M.*, et al.* (2021). Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis. Nature *589*, 448-455.
Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.
Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322-327.
Ruan, X., Liu, A., Zhong, M., Wei, J., Zhang, W., Rong, Y., Liu, W., Li, M., Qing, X., Chen, G.*, et al.* (2019). Silencing LGR6 Attenuates Stemness and Chemoresistance via Inhibiting Wnt/β-Catenin Signaling in Ovarian Cancer. Molecular therapy oncolytics *14*, 94-106.
Shen, M., and Kang, Y. (2019). Role Reversal: A Pro-metastatic Function of E-Cadherin. Dev Cell *51*, 417-419.
Siegel, R., Ward, E., Brawley, O., and Jemal, A. (2011). Cancer statistics, 2011. CA: A Cancer Journal for Clinicians *61*, 212-236.
Siegel, R.L., Miller, K.D., Fuchs, H.E., and Jemal, A. (2021). Cancer Statistics, 2021. CA: A Cancer Journal for Clinicians *71*, 7-33.
Sikandar, S.S., Kuo, A.H., Kalisky, T., Cai, S., Zabala, M., Hsieh, R.W., Lobo, N.A., Scheeren, F.A., Sim, S., Qian, D.*, et al.* (2017). Role of epithelial to mesenchymal transition associated genes in mammary gland regeneration and breast tumorigenesis. Nature Communications *8*, 1669.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.
Tanay, A., and Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. Nature *541*, 331-338.
TCGA (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61-70.
van Amerongen, R., Bowman, A.N., and Nusse, R. (2012). Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. Cell Stem Cell *11*, 387-400.
Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189-193.

476 Wang, M.H., Zhou, X.M., Zhang, M.Y., Shi, L., Xiao, R.W., Zeng, L.S., Yang, X.Z.,
477 Zheng, X.F.S., Wang, H.Y., and Mai, S.J. (2017). BMP2 promotes proliferation and
478 invasion of nasopharyngeal carcinoma cells via mTORC1 pathway. Aging (Albany NY)
479 *9*, 1326-1340.
480 Wang, Y., Jiang, X., Feng, F., Liu, W., and Sun, H. (2020). Degradation of proteins by
481 PROTACs and other strategies. Acta Pharm Sin B *10*, 207-238.
482 Wendt, M.K., Balanis, N., Carlin, C.R., and Schiemann, W.P. (2014). STAT3 and
483 epithelial-mesenchymal transitions in carcinomas. JAKSTAT *3*, e28975.
484 Will, A., Hemmann, U., Horn, F., Rollinghoff, M., and Gessner, A. (1996). Intracellular
485 murine IFN-gamma mediates virus resistance, expression of oligoadenylate synthetase,
486 and activation of STAT transcription factors. J Immunol *157*, 4576-4583.
487 Yamada, Y., Warren, A.J., Dobson, C., Forster, A., Pannell, R., and Rabbitts, T.H.
488 (1998). The T cell leukemia LIM protein Lmo2 is necessary for adult mouse
489 hematopoiesis. Proceedings of the National Academy of Sciences of the United States
490 of America *95*, 3890-3895.
491 Ye, X., Brabletz, T., Kang, Y., Longmore, G.D., Nieto, M.A., Stanger, B.Z., Yang, J., and
492 Weinberg, R.A. (2017). Upholding a role for EMT in breast cancer metastasis. Nature
493 *547*, E1-E3.
494 Zhao, C., Yang, L., Zhou, F., Yu, Y., Du, X., Xiang, Y., Li, C., Huang, X., Xie, C., Liu, Z.*,*
495 *et al.* (2020). Feedback activation of EGFR is the main cause for STAT3 inhibition-
496 irresponsiveness in pancreatic cancer cells. Oncogene *39*, 3997-4013.
497 Zhao, D., Pan, C., Sun, J., Gilbert, C., Drews-Elger, K., Azzam, D.J., Picon-Ruiz, M.,
498 Kim, M., Ullmer, W., El-Ashry, D.*, et al.* (2015). VEGF drives cancer-initiating stem cells
499 through VEGFR-2/Stat3 signaling to upregulate Myc and Sox2. Oncogene *34*, 3107-
500 3119.
501 Zhong, Z., Wen, Z., and Darnell, J.E., Jr. (1994). Stat3: a STAT family member
502 activated by tyrosine phosphorylation in response to epidermal growth factor and
503 interleukin-6. Science *264*, 95-98.

ACKNOWLEDGMENTS

We thank Patricia Lovelace, Catherine Carswell Crumpton, and other flow cytometry staff for their help and animal facility core members. The Wolverine Aria instrument was funded by NHI grant S10-1S10RR02933801. We thank Diane Heiser for technical assistance with the tail vein injections. We thank Margaret Cuadro for administrative assistance. We thank the Stanford Neuroscience Microscopy Service, supported by NIH grant NS069375.

Funding: This work was supported by NIH/NCI (U01CA154209-01 and P01 CA139490-05), the Breast Cancer Research Foundation (to M.F.C.), the U.S. Department of Defense (W81XWH-11-1-0287 and W81XWH-13-1-0281 to M.F.C.; S.S.S., W81XWH-12-1-0020), the National Cancer Institute (A.M.N., R00CA187192-03; M.F.C., 5R01CA100225-09; G.S.G., PHS grant no. CA09302), the Stanford Bio-X Interdisciplinary Initiatives Seed Grants Program (IIP) (A.M.N., M.F.C.), the Virginia and D.K. Ludwig Fund for Cancer Research (A.M.N., M.F.C), the Stinehart-Reed Foundation (A.M.N.), the Stanford School of Medicine Dean's Fellowship (J.A.), Stanford Bio-X Bowes Graduate Student Fellowship (G.S.G.), and the Stanford Medical Science Training Program (G.S.G.). K.R-H. is a New York Stem Cell Foundation – Robertson Investigator. K.R-H. is also supported by the NIH (R01-HL128503)

AUTHOR CONTRIBUTIONS

S.S.S. and M.F.C. conceived and designed the study. S.S.S. and J.A. performed experiments and analyzed data with supervision from M.F.C. G.S.G. analyzed single-cell and bulk RNA sequencing data with assistance from C.B.S. and supervision from A.M.N. A.H.K. assisted with the PDX studies. W.H.D.H. assisted with the metastasis experiments. S.D. performed staining for visualization of tumor vasculature under the supervision of K.R-H. T.A.P. assisted with the circulating cells experiment under the supervision of P.B. D.Q. provided technical support. F.D. assisted with the collection of patient specimens. J.P.T. assessed the enrichment of genes in $LMO2^{+}$ cells and provided guidance in the project. T.R provided the *Lmo2*^{CreERT2} mice. S.S.S., J.A., G.S.G., A.M.N. and M.F.C., wrote the manuscript. All authors commented on the manuscript.

Supplementary Materials:

Materials and Methods

Figures S1-S8

Tables S1-S5

Figures and Figure Legends

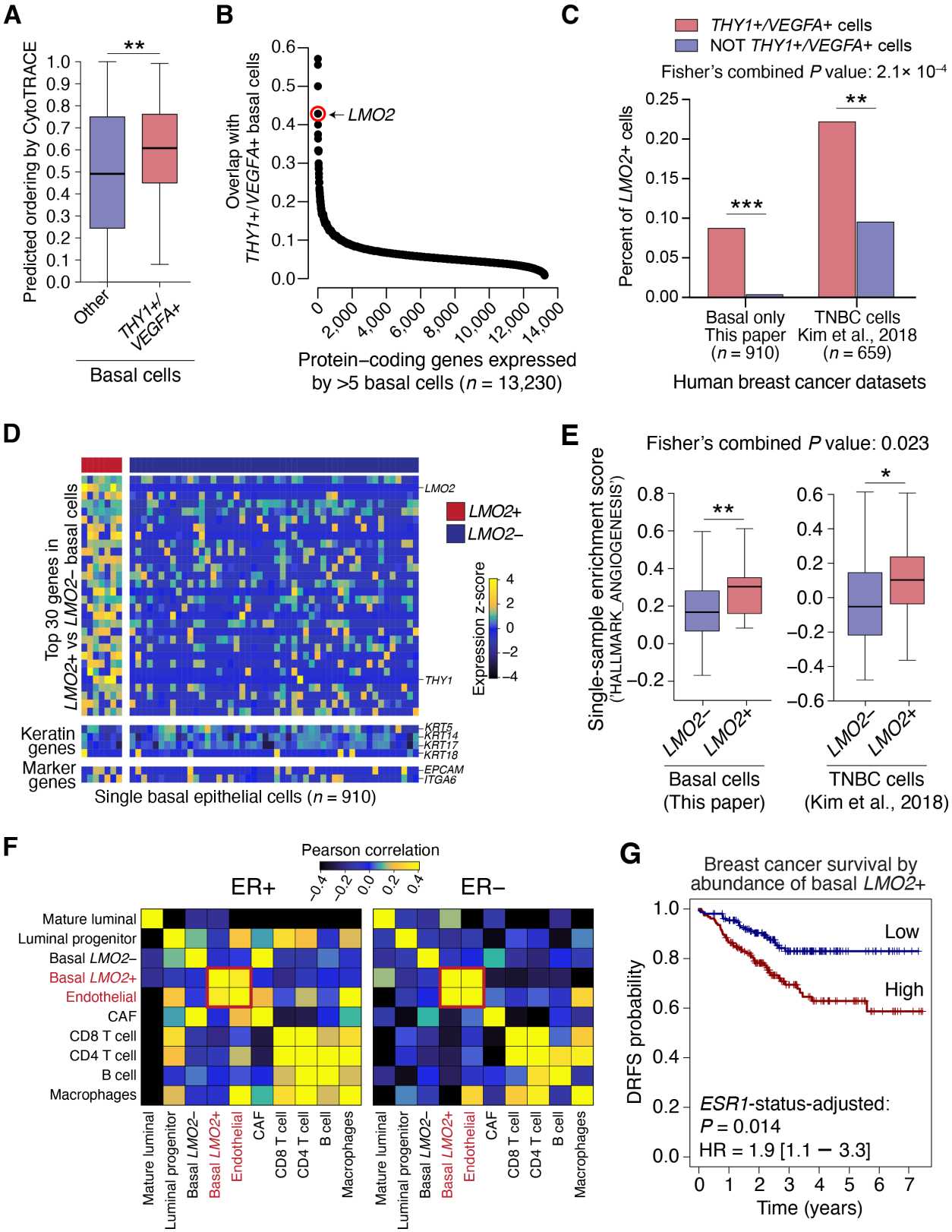


Figure 1: Identification of an immature basal epithelial population associated with pro-angiogenic signaling and poor survival in human breast cancer.

(**A**) Differentiation scores of basal epithelial cells from 17 human breast tumors profiled by scRNA-seq (all but 'SU196' contained basal cells). Differentiation scores were determined by CytoTRACE (Gulati et al., 2020). Statistical significance between *THY1*$^+$/ *VEGFA*$^+$ basal cells and other basal cells was calculated using an unpaired two-tailed *t*-test. **P*<0.1; ***P*<0.05; ****P*<0.01.

(**B**) Plot showing protein-coding genes ordered by their enrichment in *THY1*$^+$/ *VEGFA*$^+$ basal cells from human breast tumors profiled by scRNA-seq. Enrichment was defined as the number of *THY1*$^+$/*VEGFA*$^+$ basal cells expressing a given gene (TPM > 0) divided by the total number of cells expressing that gene. Only genes expressed by at least 5 basal cells were considered. *LMO2* is highlighted in red.

(**C**) Paired bar plots showing percent of *LMO2*$^+$ cells in *THY1*$^+$/*VEGFA*$^+$ cells (red) and all other cells (blue) in two human breast cancer datasets, including Kim et al., 2018 (4 primary triple-negative breast cancers, single nucleus RNA-sequencing, tumor only, *n* = 659) (Kim et al., 2018), and the basal cells (see methods for details*; n* = 910) from this study. Statistical analysis was performed by Fisher's Exact Test for association of *LMO2*$^+$ cells with *THY1*$^+$/*VEGFA*$^+$ cells. Individual and combined *P* values by Fisher's method are shown in the graph. **P*<0.1; ***P*<0.05; ****P*<0.01.

(**D**) Heatmap depicting the top 30 differentially expressed genes, along with selected keratin and lineage markers, in *LMO2*$^+$ (*n* = 7 cells) vs. *LMO2*$^-$ (*n* = 903 cells) basal epithelial cells from primary breast tumors. A random subsample of 50 *LMO2*$^-$ basal cell transcriptomes is shown for clarity. Color scale (above) represents *z*-score-normalized expression per gene.

(**E**) Differential enrichment of the 'HALLMARK_ANGIOGENESIS' pathway in *LMO2*$^+$ vs. $^-$ in two independent human breast cancer datasets described in **C.** To ensure a fair comparison between *LMO2* positive and negative populations, an empirical *P* value was calculated by comparing the mean enrichment in *LMO2*$^+$ basal cells versus a size-matched collection of *LMO2*$^-$ basal cells randomly sampled 10,000 times. A combined *P* value by Fisher's method is also shown. **P*<0.1; ***P*<0.05; ****P*<0.01.

(F and G) Cell-type and survival association of *LMO2*$^+$ basal cells across 508 bulk human breast tumor transcriptomes (Esserman et al., 2012) deconvolved using CIBERSORTx.

(F) Co-association patterns among cell type abundance profiles in bulk breast tumors, as quantified by Pearson correlation. Basal *LMO2*$^+$ cells and endothelial cells are highlighted.

(**G**) Kaplan Meier curves showing differences in distant recurrence-free survival (DRFS) in 508 breast cancer patients stratified by the median abundance of *LMO2*$^+$ basal epithelial cells. DRFS was modeled as a function of *LMO2*$^+$ basal cell status and *ESR1* status (**Methods**). The adjusted log-rank *P* value and hazard ratio with 95% confidence interval for *LMO2*$^+$ basal cell status is shown.

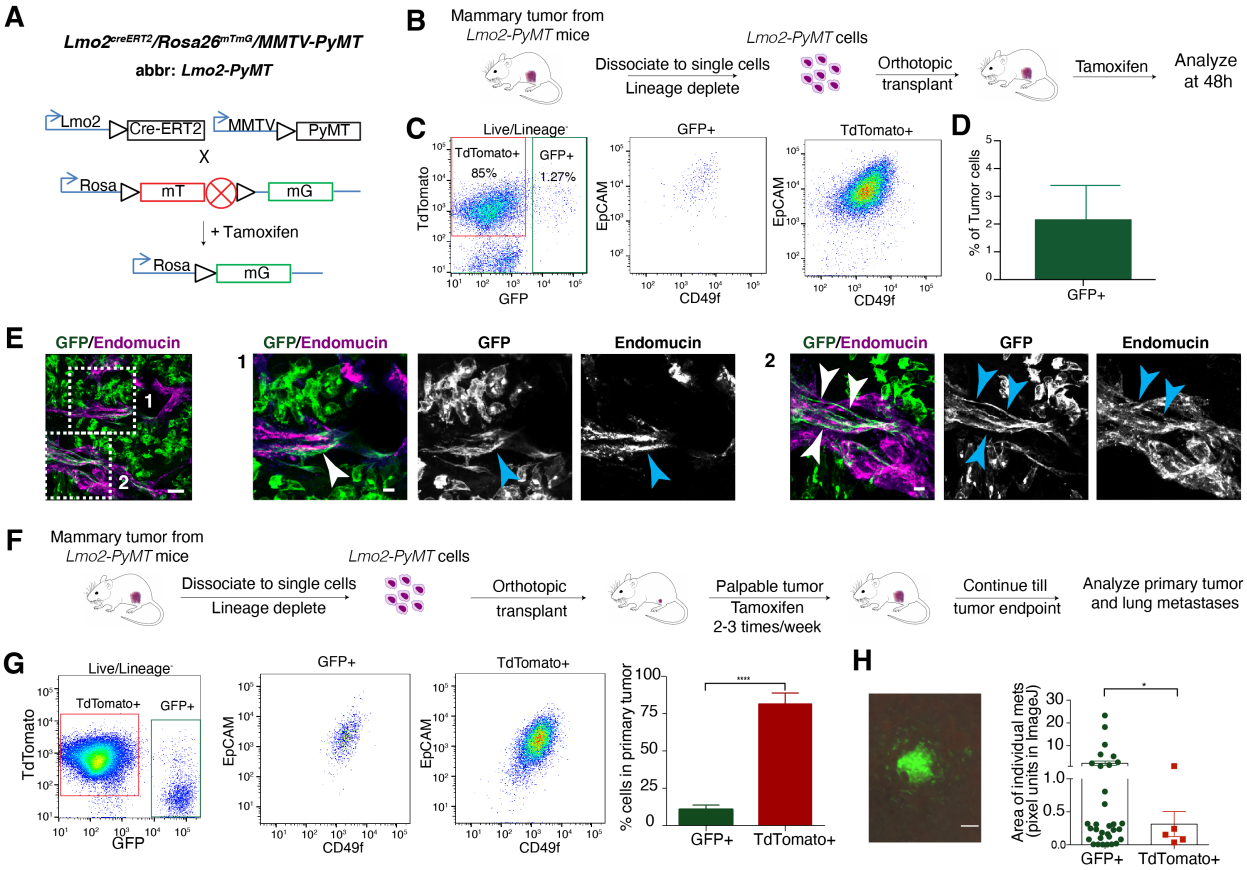


Figure 2: *Lmo2* lineage-traced tumor epithelial cells integrate into the vasculature and can form metastasis in *PyMT* tumors.

(**A**) Schematic diagram showing generation of the triple transgenic *Rosa26*^{mTmG} reporter with *MMTV-PyMT* and *Lmo2-CreERT2* mice (referred to as *Lmo2-PyMT*).

(**B**) Schematic diagram showing the experimental scheme for *Lmo2-PyMT* tumors treated with tamoxifen.

(**C**) *Panel 1*: FACS analysis of *Lmo2-PyMT* tumors 48h after Tamoxifen pulse. Cells are gated on lineage$^-$ (CD45$^-$, CD31$^-$, Ter119$^-$), DAPI$^-$ cells (See **Fig. S2**) and analyzed using TdTomato$^+$ and GFP$^+$. *Panels 2 and 3*: EpCAM and CD49f expression status in GFP$^+$ and TdTomato$^+$ cells.

(**D**) Quantification of GFP$^+$ cells from *Lmo2-PyMT* tumors (*n*=5 mice).

(**E**) Representative immunofluorescence image of *Lmo2* lineage-traced cells (GFP$^+$ green) co-localizing and integrating with endomucin (magenta) stained tumor vasculature. High resolution magnification of Inset 1 and 2 are presented, Scale bar = 50μm.

(**F**) Schematic diagram showing the experimental scheme for *Lmo2-PyMT* tumors treated with tamoxifen to trace metastatic cells.

(**G**) *Panel 1*: FACS analysis of *Lmo2-PyMT* tumors at tumor end point from (**F**). Cells are gated on lineage$^-$ (CD45$^-$, CD31$^-$, Ter119$^-$), DAPI$^-$ cells (See **Fig. S2**) and analyzed using TdTomato$^+$ and GFP$^+$. *Panels 2 and 3*: EpCAM and CD49f expression status in GFP$^+$ and TdTomato$^+$ cells. *Panels 4:* Quantification of TdTomato$^+$ and GFP$^+$ cells from *Lmo2-PyMT* tumors (*n*=4 mice).

(**H**) *Panel 1*: Representative image of metastasis shown, Scale bar = 100μm. *Panel 2*: Quantification of total number and area of GFP^+ and $TdTomato^+$ lung metastasis in *Lmo2-PyMT* tumors. (*n*=4 mice) Data are shown as mean ± SD, and statistical analysis was performed by unpaired, two-sided Wilcoxon rank sum test * $P<0.05$.

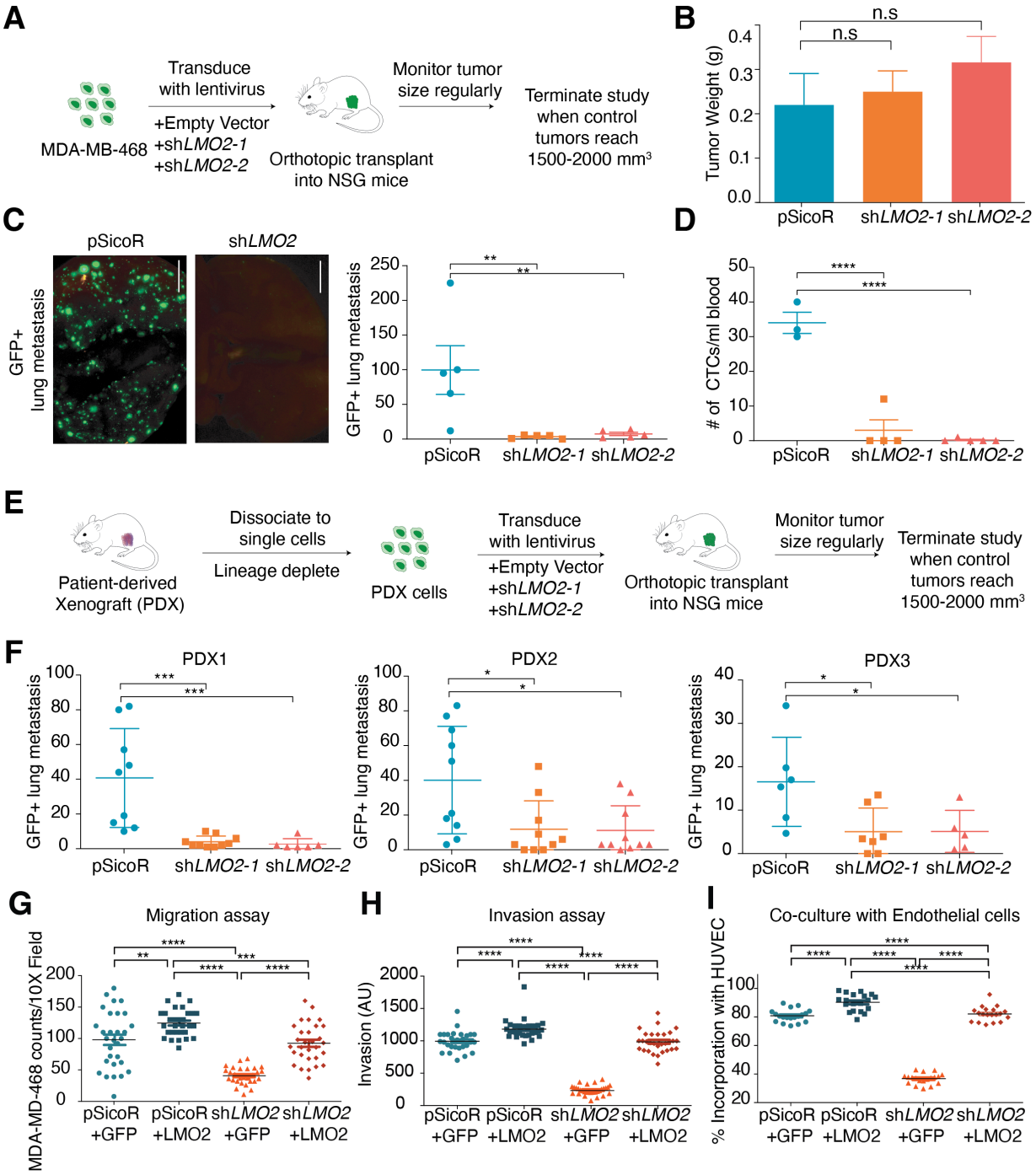
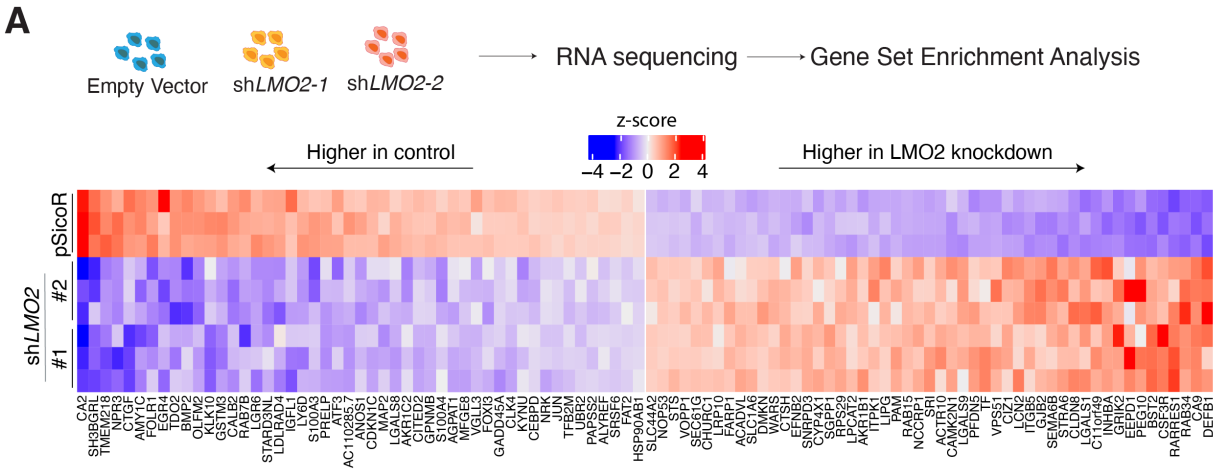


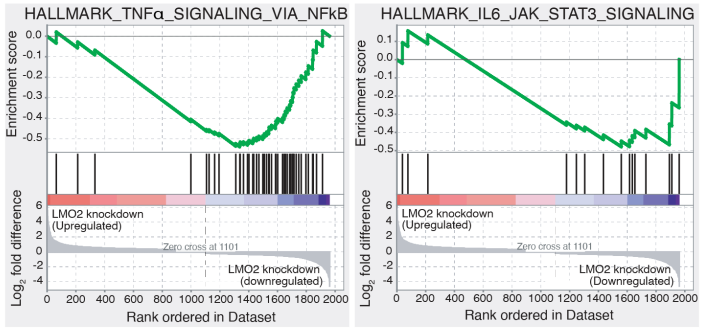
Figure 3. Knockdown of LMO2 reduces lung metastasis in human breast cancer. (**A**) Schematic of LMO2 knockdown in MDA-MB-468 cells followed by orthotopic transplant in NSG mice to evaluate tumor burden and metastases. (**B**) LMO2 knockdown in MDA-MB-468 cells. Tumor weight is shown with no significant difference between the control and LMO2 knockdown (*n*=5 mice/group). Data are shown as mean ± SD, and statistical analysis was performed by ANOVA with Dunnett's adjustment, n.s *P*>0.05

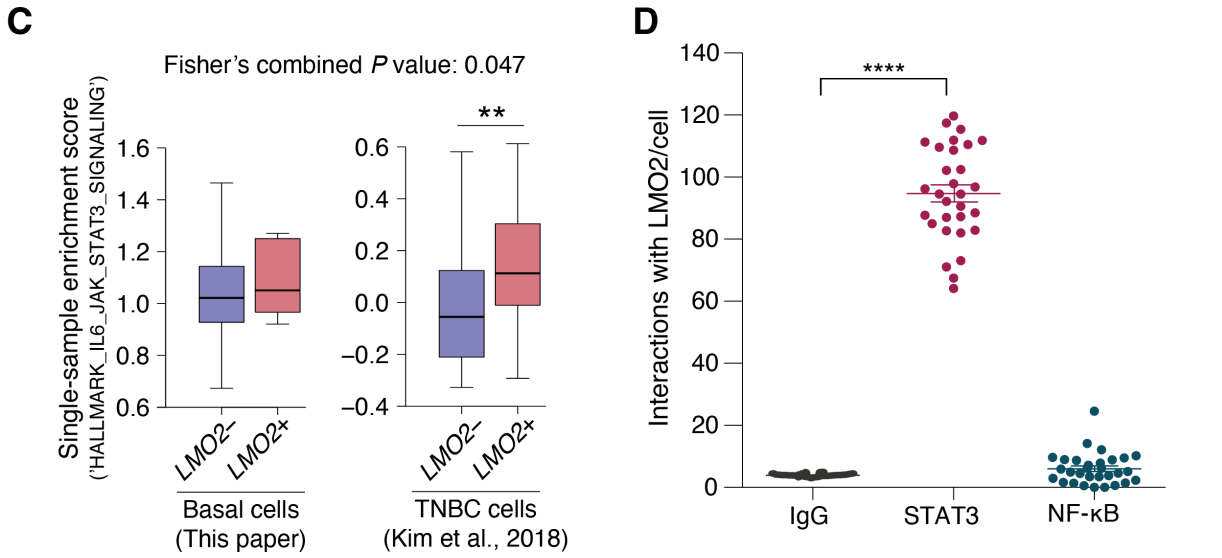
(**C**) LMO2 knockdown decreases the number of spontaneous GFP^+ lung metastasis in MDA-MB-468 cells (*n*=5mice/group). *Left panel*: representative immunofluorescence image with scale bar = 5mm, *right panel*: quantification. Data are shown as mean ± SD, and statistical analysis was performed by ANOVA with Dunnett's adjustment, ** *P*<0.01.
(**D**) LMO2 knockdown decreases the number of circulating tumor cells in MDA-MB-468 cells (*n*=3mice in pSicoR, 4 in sh*LMO2-1*, 5 in sh*LMO2-2*). Data are shown as mean ± SD, and statistical analysis was performed by ANOVA with Dunnett's adjustment, **** *P*<0.0001.
(**E**) Schematic of LMO2 knockdown in patient derived xenografts (PDXs) followed by orthotopic transplant in NSG mice to evaluate tumor burden and metastases.
(**F**) LMO2 knockdown decreased number of spontaneous GFP^+ lung metastasis in PDX samples. Data are combined from 3 independent experiments for PDX1, PDX3 and from 2 independent experiment for PDX2 (*n*=9 mice/ group for PDX1, *n*=6 mice/group for PDX2, *n*=10mice/group for PDX3). Data are shown as mean ± SD, and statistical analysis was performed by ANOVA with Dunnett's adjustment, * *P*<0.05, ** *P*<0.01, *** *P*<0.001, **** *P*<0.0001.
(**G**) MDA-MB-468 cells infected with shRNA targeting 3' UTR of LMO2 or a control shRNA pSicoR were infected with either an empty vector control 'GFP' or an LMO2-overexpression vector '+LMO2' to generate pSicoR +GFP, pSicoR +LMO2, shLMO2 +GFP, shLMO2 +LMO2. Transwell migration was quantification at 24 hours.
(**H**) Spheroid invasion assay was performed and quantified at Day 5 using the breast cancer cells from (**G**).
(**I**) The breast cancer cells from (**G**) were co-cultured with HUVEC cells and the percentage of breast cancer cells that are co-localizing with HUVEC tubes was quantified using ImageJ.
For all experiments in (**G-I**), *n*=3 and 10 images were analyzed per condition per *n*. Statistical analysis was performed by ANOVA with Dunnett's adjustment, and significance is indicated as ** *P*<0.01, *** *P*<0.001, **** *P*<0.0001



B

Pathway	NES	FDR *Q*-value
TNFα Signaling Via NFκB	-3.89	0.00
E2F Targets	-1.97	0.05
IL6-JAK-STAT3 Signaling	-1.89	0.06
Interferon Gamma Response	-1.85	0.06
Apoptosis	-1.80	0.05
Hypoxia	-1.80	0.05
KRAS Signaling UP	-1.75	0.06





641 **Figure 4: LMO2 regulates the IL6-JAK-STAT3 pathway and binds to STAT3**
642 (**A**) *Top:* Schematic of bulk RNA-sequencing analysis in MDA-MB-468 cells infected
643 with shRNAs targeting LMO2 or a control pSicoR. *Bottom*: Heatmap showing top and
644 bottom 50 genes differentially expressed between control and LMO2 knockdown
645 conditions, ordered by *P*-adjusted value.

(**B**) *Left*: Hallmark gene sets found to be significantly enriched by GSEA analysis. Normalized enrichment scores (corresponding to control pSicoR vs LMO2 knockdown) and FDR *Q*-values are determined by the GSEA software. An FDR *Q*-value cutoff of <0.25 was used to select significant gene sets. *Right*: Enrichment plots for 'HALLMARK_TNFα_SIGNALING_VIA _NFkB' and 'HALLMARK_IL6_JAK_STAT3_SIGNALING' are depicted.
(**C**) Differential enrichment of the 'HALLMARK_IL6_JAK_STAT3_SIGNALING' pathway in *LMO2*$^{+}$ vs. $^{-}$ cells from two independent human breast cancer datasets as described in **Fig. 1C.**
(**D**) Proximity mediated ligation assay showed that LMO2 had a stronger interaction with STAT3 compared to NF-kB *in vitro* (*n*=3, 10 images were analyzed per condition per *n*). Statistical analysis was performed by ANOVA with Dunnett's adjustment. **** *P*<0.0001

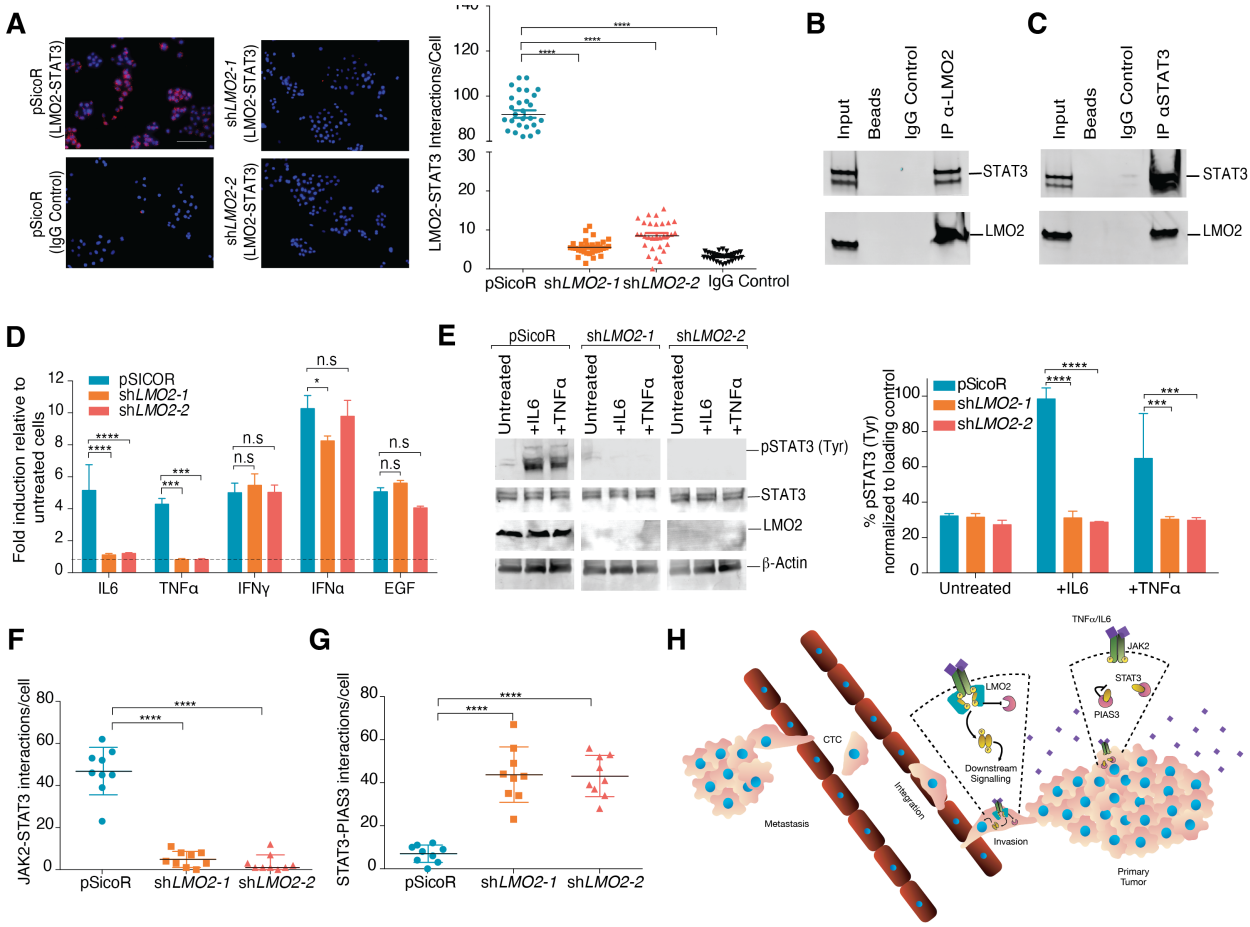


Figure 5: LMO2 stabilizes STAT3 signaling in breast cancer cells

(**A**) *Left panel*: Proximity mediated ligation assay shows that LMO2 binds to STAT3 *in vitro* and this interaction is significantly reduced with LMO2 knockdown indicating specificity of the assay. *Right panel*: Quantification of *n*=3 experiments and 10 images were analyzed per condition per *n*. Scale bar = 60μm. Statistical analysis was performed by ANOVA with Dunnett's adjustment, **** $P<0.0001$.

(**B**) Western blot of the input, immunoprecipitated beads (control), IgG (control) and LMO2 shows that LMO2 is able to pull-down STAT3. One representative blot of three independent experiments in shown.

(**C**) Western blot of the input, immunoprecipitated beads (control), IgG (control) and STAT3 shows that STAT3 is able to pull-down LMO2. One representative blot of three independent experiments in shown.

(**D**) STAT3-luciferase reporter activity shows robust stimulation of luciferase in control but not in cells with LMO2 knockdown when treated with IL6 and TNFα. IFNα, IFNγ, EGF treatment of cells results in robust stimulation in control and knockdown cells suggesting that LMO2 function is specific to IL6 and TNFα. Quantification of *n*=3 experiments. Statistical analysis was performed by 2-way ANOVA with Sidak's correction, and significance is indicated as ** $P<0.01$, *** $P<0.001$, n.s. $P>0.05$.

(**E**) Immunoblotting (*left panel*) and quantification (*right panel*) showed decreased phosphorylation of STAT3 at Tyr705 in LMO2 knockdown cells when treated with IL6

and TNFα indicating that LMO2 knockdown disrupts phosphorylation of STAT3. Quantification of *n*=3 experiments. Statistical analysis was performed by 2-way ANOVA with Sidak's correction, and significance is indicated as ***P<0.001, **** P<0.0001. (**F**) Interactions between STAT3 and JAK2 detected by proximity mediated ligation assay is reduced upon LMO2 knockdown indicating that LMO2 facilitates binding of STAT3 and JAK2. Quantification of n=3 experiments. Statistical analysis was performed by ANOVA with Dunnett's adjustment, and significance is indicated as **** P<0.0001. (**G**) Interactions between STAT3 and PIAS3 detected by Proximity mediated ligation assay is increased upon LMO2 knockdown indicating that LMO2 prevents binding of STAT3 and PIAS3. Quantification of *n*=3 experiments. Statistical analysis was performed by ANOVA with Dunnett's adjustment, and significance is indicated as **** P<0.0001. (**H**) Schematic of proposed mechanism of LMO2 in breast cancer metastasis. Tumor cells that express LMO2 have stabilized STAT3 signaling in response to IL6 and TNFα from the microenvironment, allowing these cells to intravasate into the circulation by incorporating into the vasculature.