

Cancer Biology

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint posted

July 3, 2023 (this version)

Sent for peer review

March 30, 2023

Posted to bioRxiv

March 18, 2023

The integration of Tgf β and Egfr signaling programs confers the ability to lead heterogeneous collective invasion

Apsra Nasir, Sharon Camacho, Alec T. McIntosh, Garrett T. Graham, Raneen Rahhal, Molly E. Huysman, Fahda Alsharief, Anna T. Riegel, Gray W. Pearson ✉

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA

 (https://en.wikipedia.org/wiki/Open_access)

 (<https://creativecommons.org/licenses/by/4.0/>)

Abstract

Phenotypic heterogeneity promotes tumor evolution and confounds treatment. Minority subpopulations of trailblazer cells enhance the heterogeneity of invading populations by creating paths in extracellular matrix (ECM) that permit the invasion of phenotypically diverse siblings. The regulatory programs that induce a trailblazer state are poorly understood. Here, we define a new Tgf β induced trailblazer population that is more aggressive than previously characterized Keratin 14 expressing trailblazer cells. Rather than triggering a binary switch to a single trailblazer state, Tgf β induced multiple unique states that were distinguished by their expression of regulatory transcription factors, genes involved in ECM reorganization and capacity to initiate collective invasion. The integration of a parallel Egfr signaling program was necessary to induce pro-motility genes and could be targeted with clinically approved drugs to prevent trailblazer invasion. Surprisingly, Egfr pathway activity also had the collateral consequence of antagonizing the expression of a cohort of Tgf β induced genes, including a subset involved in ECM remodeling. Together, our results reveal a new compromise mode of signal integration that promotes a trailblazer state and can be therapeutically targeted to prevent collective invasion.

eLife assessment

This represents an **important** study that demonstrates a high degree of heterogeneity within trailblazer cells in clusters that participate in collective migration. **Solid** methods highlight this heterogeneity and show that in TNBC cancers, trailblazer cells are defined by vimentin (and not Keratin 14) and are dependent on both TGFbeta and EGFR signaling.

Introduction

Tumor cell phenotypic heterogeneity is the product of genetic and epigenetic variability, localized differences in the microenvironment and stochastic fluctuations in intracellular

signaling pathways (1). This diversity in cell phenotypes drives tumor evolution (2). Cells with traits that provide a competitive advantage over siblings become predominant, improving the fitness of the population (3). High variance in cell phenotypes also increases the odds that a subpopulation has the potential to survive when the local environment changes, such as during treatment or when disseminating cells seed other organs (4, 5). Interestingly, analysis of intrinsic and genetically engineered heterogeneity has revealed that phenotypic diversity also promotes functional interactions between tumor cells (6, 7). Rare enabler tumor subpopulations can promote the growth, immune evasion, invasion and metastasis of abundant siblings harboring different phenotypic traits (8-13). However, the regulatory programs that promote functional relationships between phenotypically distinct subpopulations remain poorly understood. This lack of mechanistic insight limits our ability to define the presence of functional tumor cell interactions, evaluate their impact on patient outcome, and develop new treatment options designed to target rare enabler cells (14).

To better understand how subpopulation relationships are induced during tumor progression, we sought to understand how tumor cells are conferred with the ability to promote the collective invasion of intrinsically less invasive siblings. Collective invasion is the predominant mode of invasion in carcinomas and promotes the dispersion of tumor clusters that seed metastases (15-19). Subpopulations of trailblazer cells detected in breast, lung and thyroid tumor populations have an enhanced ability to reorganize the extracellular matrix (ECM) into micro-tracks that facilitate the initiation of collective invasion (20-23). Importantly, sibling opportunist populations invade through the micro-tracks created by minority trailblazer populations in three-dimensional culture models and xenograft tumors (21, 24, 25). Moreover, opportunist cells can have greater proliferative capacity than trailblazers in vitro and in xenografts, indicating that trailblazer-opportunist relationships have the potential to increase the proliferative fitness of invading cell populations (24, 25). Thus, trailblazer cells promote phenotypic diversity in invading cell cohorts, which may increase the risk of metastasis and elevate the risk of resistance to treatment (26, 27).

Trailblazer cells are distinguished from the bulk tumor population by the increased expression of trailblazer genes, including regulators of the actin cytoskeleton, integrins, receptor tyrosine kinases, keratins, and cancer testes antigens (20, 21, 23, 24, 28). Trailblazer genes contribute to the formation of actin-rich cellular protrusions that provide traction and exert tensile forces necessary for ECM remodeling and other undefined functions (21, 23, 25). Interestingly, trailblazer genes are not required for other forms of cell movement, including migration along two-dimensional surfaces, movement within spheroids and opportunistic collective invasion (21). While genes that specifically contribute to unique invasive characteristics of trailblazer cells have begun to be defined, the regulatory programs underpinning the conversion to a trailblazer state remain largely unknown. Trailblazer cells in breast and lung cancer cell line models are epigenetically distinct from opportunist siblings, although the processes that initiate the conversion to the trailblazer state are not known (21, 23). Breast cancer trailblazer cells have basal characteristics, with Keratin 14 (Krt14) expression proposed as a universal feature for identifying trailblazer cells in mammary tumors (18, 20, 29, 30). Yet, the program controlling the induction of a basal trailblazer state is undefined. Breast cancer trailblazer cells also have features of epithelial-mesenchymal transition (EMT) re-programming (21). However not all EMT programs yield trailblazer activity and whether EMT programs regulate the expression of trailblazer genes is not known (25, 31). Thus, if and how the EMT process imbues a trailblazer state is unclear. Our poor understanding of how trailblazer populations are induced limits options for potential interventions targeting trailblazer cells and advances in employing subpopulation identification to refine patient outcome predictions.

To define factors regulating trailblazer program activity, we determined how cells were induced to collectively invade in the C3(1)/SV40 T-antigen (C3-TAg) genetically engineered

mouse model (GEMM) of breast cancer. Based on histological analysis, breast cancer is defined as Estrogen Receptor-positive (ER-positive), Human Epidermal Growth Factor Receptor-positive (HER2-pos) or triple-negative (TNBC), which have low to undetectable levels of ER, HER2 and the Progesterone Receptor (32). Breast cancer is further stratified into luminal A, luminal B, HER2-enriched and basal-like based on mRNA expression. Luminal A are Estrogen Receptor-positive (ER-pos) while luminal B are more proliferative and are ER-pos or HER2-pos (33). The HER2-enriched subtype is generally HER2-pos and 70-80% of the basal-like breast cancers are classified as TNBC (33). The C3-Tag model is most similar to human basal-like TNBC (34, 35).

Through the analysis of tumor organoid invasion dynamics, a data driven functional dissection of cellular re-programming and quantitative imaging of fixed tumor sections, we discovered a Tgf β regulated program that was necessary to induce and sustain a trailblazer state in C3-TAg tumors and validated this discovery in breast cancer patient tumor samples. Surprisingly, the Tgf β regulated trailblazer cells were not specified by Krt14 expression and were more aggressive than a previously defined Krt14 expressing trailblazer population detected in a luminal B GEMM, indicating that there are multiple different programs for conferring trailblazer ability in breast tumors. In addition, trailblazer cells induced by Tgf β signaling were themselves heterogeneous, having different degrees of re-programming and capacities to initiate invasion. Tgf β signaling collaborated with a parallel Egfr-Mek1/2-Erk1/2-Fra1 pathway to coordinate the expression of a subset of pro-motility genes, including Axl. Interestingly, the integration was also equivalently antagonistic in scope, with Egfr activity restricting the Tgf β induced expression of a cohort of pro-invasive genes. Notably, clinically approved inhibitors of Egfr, Mek1/2 and Axl suppressed collective invasion. This indicated that the collateral inhibition of subset of pro-invasive genes by the Egfr pathway was a feature of the trailblazer state. Thus, our results reveal how a new type of trailblazer cell is induced by a therapeutically addressable compromise between Tgf β and Egfr signaling programs.

Results

The trailblazer cells in C3-TAg basal-like tumors are distinct from trailblazer cells in PyMT luminal B tumors

To define factors that initiate the conversion to a trailblazer state, we investigated the invasive properties of tumor cells isolated from the C3-TAg GEMM of basal-like TNBC (34, 35) and (**Supplementary file 1**). We first determined if there was heterogeneity with respect to the ability to invade in the tumor populations, which would indicate the potential for trailblazer-opportunist relationships. Indeed, there was extensive heterogeneity in the invasive ability of tumor organoids and single dissociated tumor cells (**Figure 1A, Figure 1—figure supplement 1A-B, Figure 1—Videos 1-2**). Clonal invasive (D6 and C7) and clonal non-invasive cells (B6 and G2) were also derived from a single C3-TAg tumor (**Figure 1—figure supplement C-E, Supplementary file 1**). To test whether the invasive subpopulations in C3-TAg tumors could lead the invasion of siblings, we generated heterogeneous clusters containing C3-TAg tumor cells and intrinsically non-invasive B6 cells (**Figure 1B, Figure 1—Videos 3-5**). The C3-TAg tumor cells and the D6 cells led the collective invasion of B6 cells (**Figure 1B, Figure 1—figure supplement 1E, Figure 1—Videos 3 and 6**). These combined results indicated that there were trailblazer subpopulations in C3-TAg tumors with an enhanced ability to lead the collective invasion of intrinsically less invasive siblings.

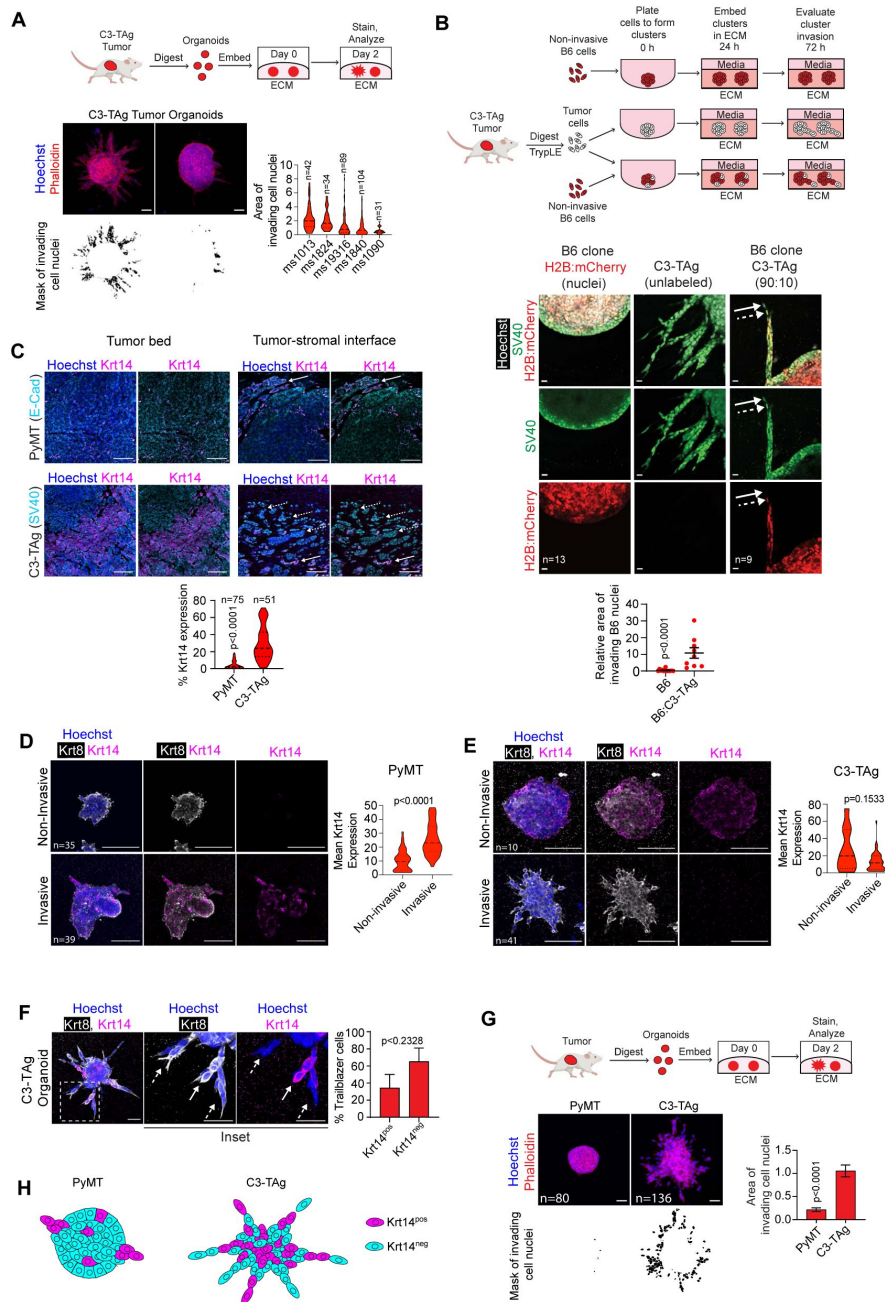


Figure 1 with 1 supplement.

Trailblazer cells in C3-TAG basal-like tumors are distinct from trailblazer cells in PyMT luminal B tumors.

A. There is intra- and inter-tumor invasive heterogeneity in C3-TAG organoids. Graphic shows the process for the derivation and analysis of tumor organoids. Representative images show the invasive heterogeneity of C3-TAG organoids derived from a single tumor. The area of invasion is determined using an image mask of the nuclei of invading cells. Violin plot shows the quantification of organoid invasion within 48 h of tumor isolation. (Mann Whitney test, n =organoids). Scale bars, 50 μ m. **B.** C3-TAG tumors contain trailblazer cells that can lead the collective invasion of intrinsically less invasive cells. Graphic shows the process for generating homogeneous and heterogeneous clusters. Images show clusters immunostained with anti-SV40 antibody (detects TAG in all tumor cells), counterstained with Hoechst (nuclei) and H2B:mCherry fluorescence (B6 cells only). C3-TAG tumor cells (unlabeled) promoted the invasion of B6 cells (H2B:mCherry). Solid arrow indicates a C3-TAG tumor cell leading invasion. Dashed arrow indicates a B6 cell invading behind the C3-TAG cell. Graph

shows the relative area of B6 cell invasion from 3 independent experiments (mean $\pm$ SEM, Mann Whitney test, n =clusters). Scale bars, 20 μ m. **C.** Krt14 expression does not correlate with C3-TAG tumor invasion. Krt14, SV40 (C3-TAG) and E-Cad (PyMT) expression in the tumor core and tumor-stromal interface of C3-TAG and PyMT tumors. Solid and dashed arrows indicate Krt14^{pos} and Krt14^{neg} cells respectively leading collective invasion. Violin plot shows the percentage of tumor cell area expressing Krt14. (Mann Whitney Test, n =ROIs from 5 tumors). Scale bars, 100 μ m. **D.** Krt14 expression positively correlates with PyMT organoid invasion. Violin plot shows mean Krt14 expression in invasive and non-invasive PyMT organoids. (Mann Whitney Test, n =organoids from 3 tumors). Scale bars, 100 μ m. **E.** Krt14 expression does not correlate with C3-TAG organoid invasion. Violin plot shows mean Krt14 expression in invasive and non-invasive C3-TAG organoids. (Mann Whitney Test, n =organoids from 3 tumors). Scale bars, 100 μ m. **F.** Krt14 expression does not correlate with trailblazer phenotype in C3-TAG organoids. Solid and dashed arrows indicate Krt14^{neg} and Krt14^{pos} trailblazer cells respectively. Bar graph shows Krt14 expression status in trailblazer cells. (mean $\pm$ SEM, unpaired Student's t test, n =organoids from 3 tumors) Scale bars, 50 μ m. **G.** C3-TAG organoids are more invasive than PyMT organoids. Bar graph shows the

area of invasion (mean±SEM, Mann Whitney test, n=clusters, r=2) Scale bars, 50 µm. **H.** Model depicting the relationship between Krt14 expression and trailblazer phenotype in PyMT and C3-Tag tumors.

To understand how the trailblazer state was induced in C3-Tag tumor cells, we next evaluated the relationship between Krt14 expression and the ability of cells to lead collective invasion. Krt14 identifies cells that have converted from a luminal non-invasive state to a basal trailblazer state in the MMTV-PyMT (PyMT) GEMM model of luminal B breast cancer, which is commonly used to study breast cancer invasion and metastasis (20, 36). Based on these results, it has been suggested that a basal gene expression program exemplified by Krt14 expression is a universal feature of breast cancer trailblazer cells (20, 29). Consistent with prior results, Krt14 was expressed in PyMT trailblazer cells and Krt14 expression correlated with PyMT tumor organoid invasion (**Figure 1C-D**). Krt14 was expressed in a greater proportion of cells in C3-Tag tumors compared to PyMT tumors (**Figure 1C**). The high Krt14 expression was consistent with the early acquisition of basal traits detected in C3-Tag and TNBC tumors during the initial stages of growth within mammary ducts (37, 38). However, Krt14 expression did not correlate with C3-Tag tumor collective invasion or a trailblazer state in our analysis of primary tumors, tumor organoids and trailblazer and opportunist cell lines (**Figure 1C, E-F, Figure 1—figure supplement 1F**). In fact, extensive collective invasion was detected in C3-Tag tumor organoids with few, if any, Krt14 expressing cells (**Figure 1E**). Krt14^{neg} C3-Tag tumor cells also led collective invasion as frequently as Krt14^{pos} cells (**Figure 1F**). C3-Tag organoids were relatively more invasive than PyMT organoids, demonstrating that the C3-Tag tumor invasion was more aggressive than the Krt14 specified invasion of PyMT tumors (**Figure 1G, 1H, Figure 1—figure supplement 1G, Figure 1—Videos 7-8**). Together, these results suggested that essential features of the C3-Tag trailblazer program were different than those of the PyMT trailblazer program, and that more than a single Krt14 specified trailblazer program functioned in breast cancer patient tumors (**Figure 1H**).

Trailblazer cells in C3-Tag GEMM tumors are distinguished by a subset of EMT traits

Given the lack of concordance with the established Krt14 specified trailblazer program, we next took an unbiased approach to identify the key trailblazer regulatory program features in C3-Tag tumors. To do this, we used next generation RNA sequencing to define differentially expressed genes (DEGs) in tumor organoids from 4 C3-Tag tumors compared with non-invasive spheroids formed by 3 clonal C3-Tag tumor derived cell lines (**Figure 2A, Figure 2—figure supplement 1A**). Among epithelial keratins, Krt14 expression was reduced in C3-Tag tumor organoids compared to the non-invasive clonal C3-Tag cells, consistent with a lack of Krt14 involvement in C3-Tag collective invasion (**Figure 2B, Supplementary file 2**). Expression of the transcription factor p63 was also reduced in C3-Tag tumor organoids and trailblazer clones, further indicating that canonical traits of normal basal mammary epithelial cells and Krt14 specified trailblazer cells (20) were not essential features of C3-Tag trailblazer cells (**Figure 2B, Figure 2—figure supplement 1B, Supplementary file 2**). In addition, Krt14 expression was only detected in 1 out of 4 trailblazer populations in human breast cancer cell lines, which were all classified as TNBC (**Figure 2—figure supplement 1C**). The expression of p63 was also low in all 4 trailblazer populations (**Figure 2—figure supplement 1C**), consistent with our previous results showing that p63 expression is sufficient to suppress the induction of a trailblazer state (25).

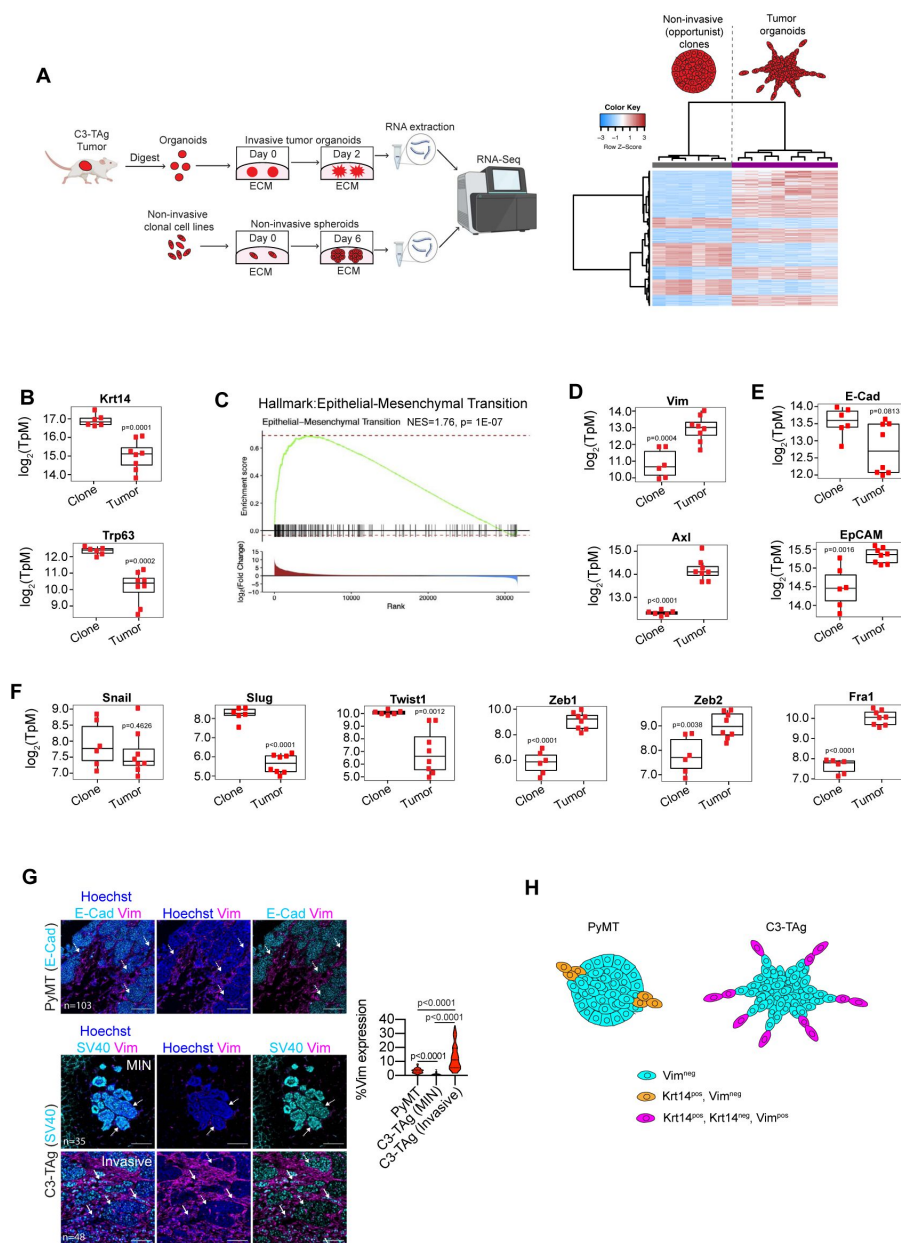


Figure 2 with 1 supplement.

Trailblazer cells are distinguished by a subset of EMT traits in C3-TAG basal-like mammary tumors.

A. Graphic shows preparation of C3-TAG organoids and non-invasive clonal cell line spheroids for RNA-seq analysis. Heatmap shows unsupervised clustering and differentially expressed genes in non-invasive C3-TAG clones (B6, C6, E8) and invasive organoids from 4 C3-TAG tumors. **B.** Boxplots showing the expression of the basal genes Krt14 and Trp63. **C.** There was increased expression of Hallmark EMT genes in C3-TAG tumor organoids compared to non-invasive C3-TAG clones. Plot shows the details of the Hallmark EMT generated by GSEA. **D-F.** Boxplots showing the expression of canonical mesenchymal (Vimentin, Axl), epithelial (E-cadherin, EpCAM) and EMT-TF (Snai1, Slug, Twist1, Zeb1, Zeb2, Fra1) genes in C3-TAG tumor organoids compared to non-invasive clonal spheroids from RNA-seq data (Mann Whitney test for E-cad, Student's t-test for all others, $n=6$ for clones, (3 biological replicates); $n=8$ for tumor (4 biological replicates). **G.** Vimentin is more highly expressed in invasive

C3-TAG primary tumors compared to noninvasive C3-TAG tumors and PyMT tumors. Solid arrows indicate Vim^{pos} tumor cells leading collective invasion. Dashed arrows indicate Vim^{neg} tumor cells. Violin plot shows percentage of tumor cell area expressing Vimentin (Mann Whitney Test, $n=ROIs$ from 5 tumors). Scale bars, 100 μm . **H.** Model showing the distinct features of trailblazer cells in PyMT and C3-TAG tumors.

Genes highly expressed in C3-TAG tumor organoids were associated with biological processes that contribute to invasion (ECM Receptor Interaction, Focal Adhesions and Regulation of the Actin Cytoskeleton) and EMT (Figure 2C-D, Figure 2—figure supplement 1D-E, Supplementary file 2). The C3-TAG tumor organoids retained canonical epithelial gene expression as well, indicating that the majority of cells were in a hybrid EMT state (Figure 2E, Supplementary file 2). The EMT inducing transcription factors (EMT-TFs) Zeb1, Zeb2 and Fra1 were more highly expressed in C3-TAG organoids, consistent with their high expression in trailblazer cells in human breast cancer cell lines (Figure 2F, Figure 2—figure

supplement 1C). However, the EMT-TFs Snail, Slug and Twist1 were expressed at similar or higher levels in the noninvasive clones, which was consistent with a lack of concordance between the induction of a trailblazer state and Snail, Slug and Twist1 expression in human breast cancer cell lines (**Figure 2F**, **Figure 2—figure supplement 1C**). These results indicated that the specific features of a trailblazer EMT program determined the invasive phenotype of C3-TAg tumor cells, rather than simply the presence or absence of EMT traits, consistent with results obtained analyzing human breast cancer cell lines (25, 31). Our combined results demonstrated that an EMT program exemplified by Vimentin expression conferred a trailblazer state in C3-TAg tumor subpopulations (**Figure 2G**).

To understand the relationship between the EMT program features of the C3-TAg organoids and collective invasion, we evaluated the expression of the microfilament protein Vimentin, a common EMT marker that was more highly expressed in C3-TAg organoids (**Figure 2D**). Vimentin expression was more frequently and extensively detected in invasive C3-TAg tumor cells compared to non-invasive C3-TAg mammary intraepithelial neoplasia (MIN) cells, consistent with Vimentin expression specifying a trailblazer EMT state in C3-TAg tumors. Notably, Vimentin expression was rarely detected in invasive PyMT tumor cells, further indicating that the regulatory programs inducing a trailblazer state in C3-TAg and PyMT tumors were distinct (**Figure 2G**, **Figure 2—figure supplement 1F**). The heterogeneity of Vimentin expression in C3-TAg tumors was consistent with the heterogeneity in the extent of C3-TAg organoid invasion (**Figure 1A**). Vimentin^{pos} and Vimentin^{neg} cells also collectively invaded together in C3-TAg primary tumors and organoids, suggesting that there was heterogeneity in collectively invading cell populations with respect to EMT state (**Figure 2G**, **Figure 2—figure supplement 1F**).

Trailblazer cells are specified by Vimentin expression in basal-like breast cancer patient tumors

We next investigated the characteristics of collectively invading cells in breast cancer patient tumors. KRT14 was expressed at a high level in basal like patient tumors compared to luminal B type breast cancer patient tumors, consistent with our results comparing the C3-TAg basal like and PyMT luminal B breast cancer GEMMs (**Figure 3A**). However, KRT14 expression did not correspond with the induction of collective invasion in TNBC patients, identical to the lack of association between Krt14 expression and invasion in C3-TAg primary tumor and organoids (**Figure 3B**, **Figure 3—figure supplement 1A**). In contrast, Vimentin expression was higher in the basal-like breast cancer compared to other breast cancer subtypes. Importantly, Vimentin expression was heterogeneous in TNBC tumor sections, with a trend towards more extensive Vimentin expression in regions of collective invasion (**Figure 3C-D**, **Figure 3—figure supplement 1B**). These combined results indicated that the difference in trailblazer program features in C3-TAg (basal-like) and PyMT (luminal B) tumors reflected a feature of inter-tumor heterogeneity found in human breast cancer.

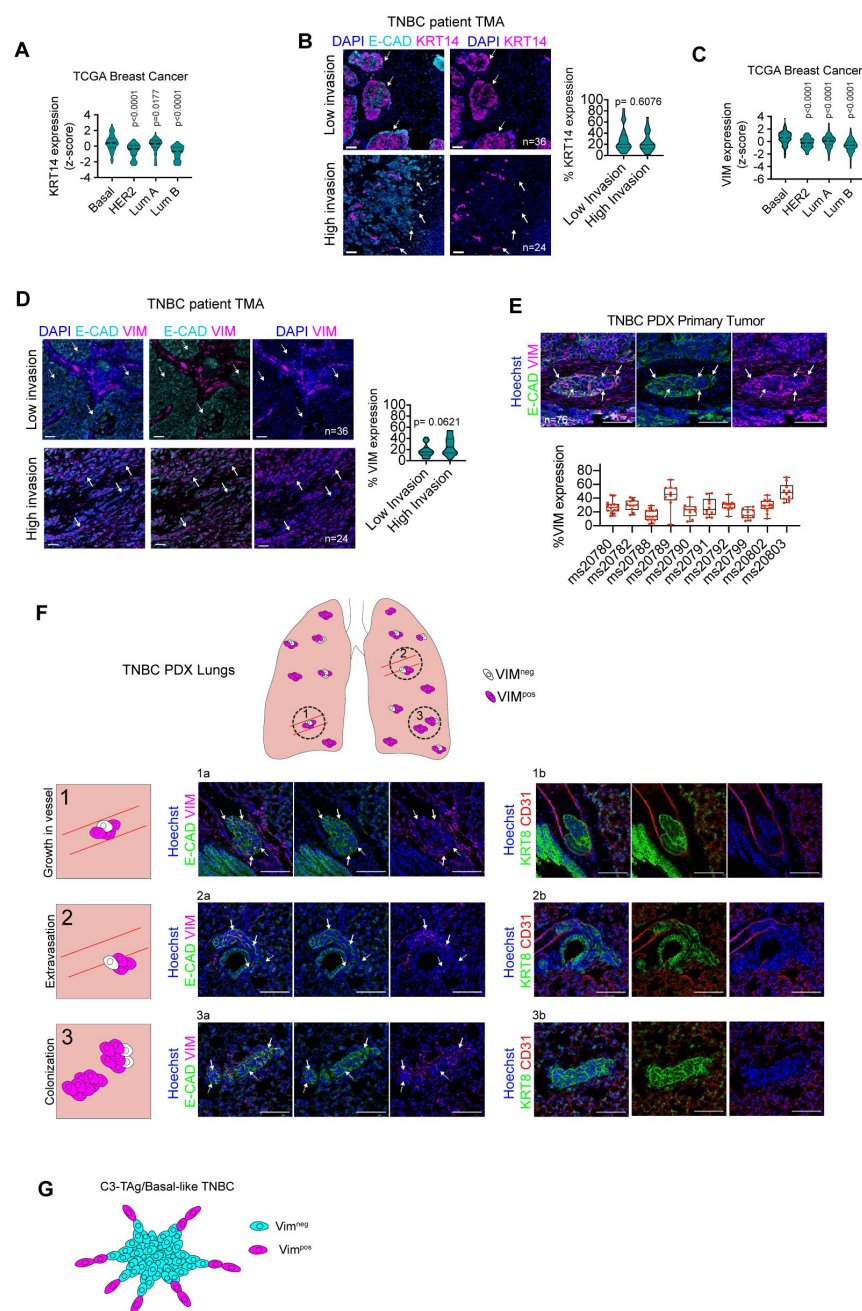


Figure 3 with 1 supplement.

Trailblazer cells are distinguished by a subset of EMT traits in basal-like subtype of breast cancer.

A. KRT14 mRNA is more highly expressed in the basal-like breast cancer patient tumors compared to other breast cancer subtypes. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test) From the TCGA. P-values are in comparison to the basal-like patient tumors. **B.** Krt14 expression in TNBC patient tumors does not correlate with invasion. Violin plot shows percentage of Krt14 expressing cells in regions of low (dashed arrows) and high invasion (solid arrows) (Mann Whitney test, n=tumor sections). Scale bars, 50 μ m. **C.** VIM mRNA is more highly expressed in the basal-like breast cancer patient tumors compared to other breast cancer subtypes. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test) From the TCGA. P-values are in comparison to basal-like tumors. **D.** Invasive regions in TNBC patient tumors trend towards higher Vimentin expression (solid arrows) compared to regions of low invasion (dashed arrows). Violin plot shows percentage of Vimentin expressing tumor cells (Mann Whitney test, n=tumor sections). Scale bars, 50 μ m. **E.** Vimentin is expressed in trailblazer cells in HCI-001 basal-like TNBC patient derived

xenografts. Solid arrows indicate high Vimentin expression. Dashed arrows indicate low Vimentin expression. Violin plot shows percentage of tumor cell area expressing Vimentin (Mann Whitney Test, n=ROIs from 10 HCI-001 tumors). Scale bars, 100 μ m. **F.** metastasis data. **G.** Model showing that Vimentin expression specifies trailblazer cells in basal-like C3-Tag tumors and TNBC patient tumors.

Metastasis is rarely detected in female mice bearing C3-Tag mammary tumors, likely because the rapid rate growth in primary tumors requires mice to be euthanized before metastases have time to develop (34). Therefore, to understand the relationship between the induction of Vimentin high cells and metastasis, we evaluated Vimentin expression in metastatic HCI-001 patient derived xenograft (PDX) TNBC tumors and their corresponding lung metastases (39). Intra- and inter-tumor heterogeneity in Vimentin expression was

detected in the HCI-001 primary tumors (**Figure 3E**). The heterogeneity and range in Vimentin expression was similar to what was observed in C3-TAg tumors and TNBC patient tumors (**Figure 2G**, **Figure 3D**). An evaluation of the lungs revealed disseminated HCI-001 tumor cell clusters in the pulmonary vasculature (**Figure 3F**, **top row**), partially localized within the vasculature and lung tissue (**Figure 3F**, **middle row**) and fully localized within lung tissue (**Figure 3F**, **bottom row**). Notably, heterogeneous Vimentin expression was detected in each category of disseminated cell clusters, consistent with the induction of Vimentin-positive trailblazer cells metastasizing to the lungs (**Figure 3F**). Together, our analysis indicated that Vimentin expression specified a trailblazer state in basal-like TNBC tumors, consistent with the features of the C3-TAg trailblazer populations.

Tgf β induces a trailblazer state in C3-TAg tumors

We next sought to define regulatory factors that activated the Vimentin specified trailblazer program. Gene set enrichment analysis (GSEA) indicated that Tgf β signaling was more active in C3-TAg tumor organoids than non-invasive spheroids (**Figure 4A**, **Supplementary file 2**). Thus, we evaluated the contribution of Tgf β towards regulation of a trailblazer state. A pharmacological inhibitor of Tgf β receptor 1 (Tgfr1), A83-01, suppressed the invasion of freshly derived C3-TAg tumor organoids (**Figure 4B**). 1863 cells derived in the presence of A83-01 were also less invasive than 1863T cells established from the same tumor in the absence of A83-01 (**Figure 4—figure supplement 1A**, **Supplementary file 1**). Depletion of the Tgf β regulated transcription factor Smad3 reduced the invasion of TGF β 1 stimulated 1863 cells and 1863T cells, indicating that TGF β regulated gene expression contributed to the observed phenotypes (**Figure 4—figure supplement 1B-C**). These results indicated that Tgf β signaling was necessary to sustain a trailblazer state in C3-TAg cells.

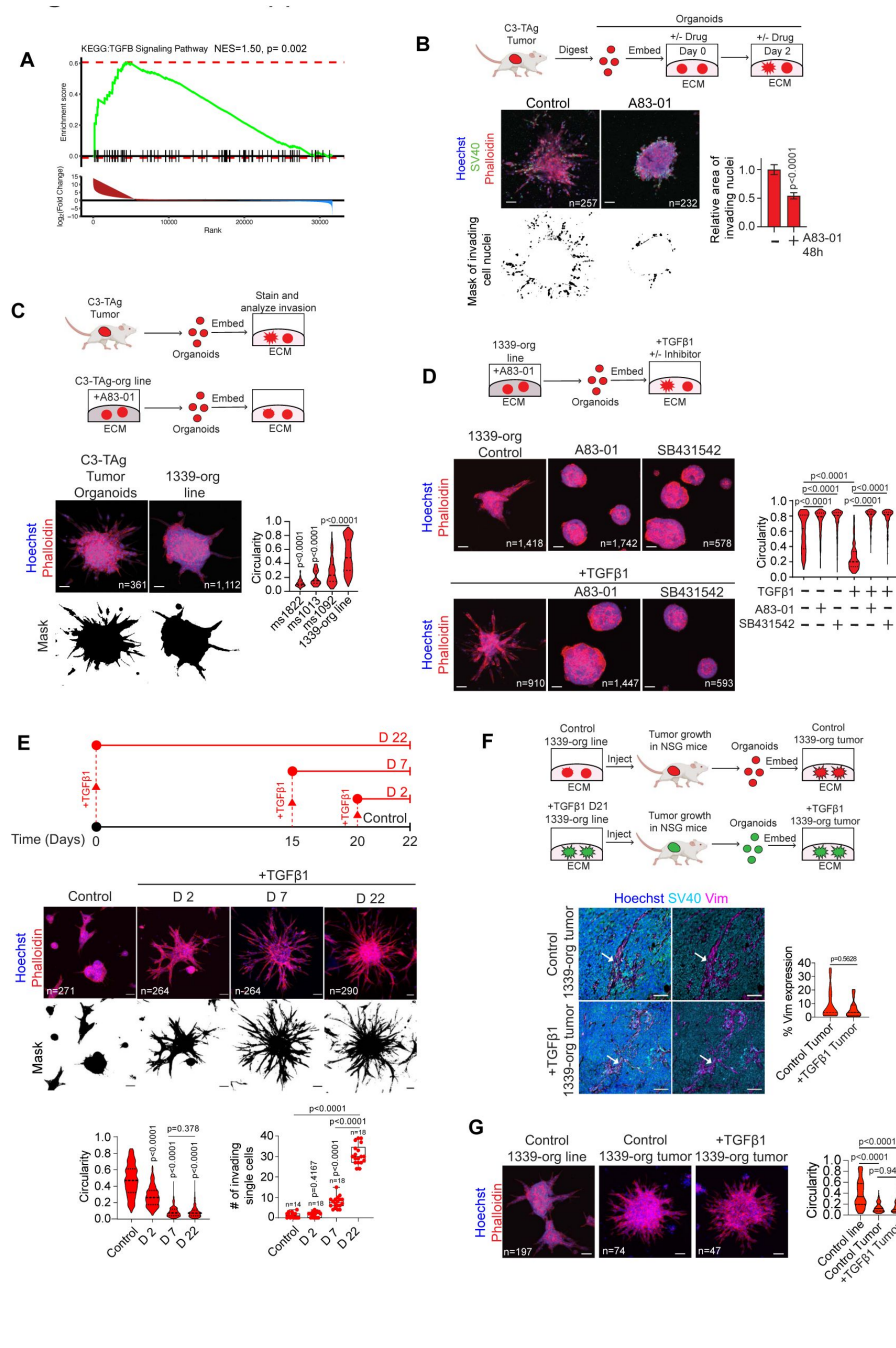


Figure 4 with 3 supplements.

TGFβ induces a trailblazer state in C3-TAG tumors.

A. GSEA was performed on RNA-seq data from C3-TAG organoids and non-invasive C3-TAG clonal cell line spheroids to identify pathway enrichment in C3-TAG organoids. Plot shows the GSEA for the KEGG: TGFβ signaling pathway. **B.** The Tgfr1 inhibitor A83-01 suppresses C3-TAG organoid trailblazer phenotype within 48 h of tumor isolation. Graph shows the quantification of invasion (mean±SEM, Mann Whitney test, n=organoids from 4 tumors). Scale bars, 50µm. **C.** 1339-orgs grown in A83-01 were less invasive than organoids tested immediately after isolation from primary tumors. Graph shows organoid circularity (mean±SEM, Mann Whitney test, n=organoids). Scale bars, 50 µm. **D.** Tgfr1 inhibitors suppress the autocrine and exogenous TGFβ1 induced trailblazer phenotypes in 1339-orgs. Organoids were treated with the Tgfr1 inhibitors A83-01 (500 nM) or SB431542 (1 µM) and exogenous TGFβ1 (2 ng/ml) as indicated. Violin plots show the quantification of organoid circularity (Mann Whitney test, n=organoids, r=3). Scale bars, 50 µm. **E.** Persistent exposure to exogenous TGFβ1 progressively enhanced the invasive phenotype in C3-TAG organoid lines. Graphs show the quantification of organoid circularity and the number of invading single cells (Mann Whitney test, n=organoids, r=3) Scale bars, 50 µm. **F.** Graphic shows the process for generating new primary tumors from control 1339-org line cells (Control 1339-org tumors) and 1339-org cells treated with TGFβ1 for 21 days (+TGFβ1 1339-org tumors). The expression of Vimentin was similar in Control 1339-org tumors and +TGFβ1 1339-org tumors. Violin plot shows quantification of Vimentin expression (Mann Whitney Test, 4 tumors from each group). Scale bars, 100 µm. **G.** The trailblazer phenotype of control 1339-org line cells was enhanced after tumor growth and similar to +TGFβ1 1339-org tumor cells. Violin plot shows quantification of organoid circularity (Mann Whitney Test, n=organoids representative of 4 tumors from each group).

To determine if Tgf β could initiate the conversion from an opportunist to a trailblazer state, we first established organoid lines from 3 different C3-TAg tumors in the presence of A83-01 for at least 4 weeks (**Figure 4—figure supplement 2A, Supplementary file 1**). These conditions yielded 1339-org, 1863-org and 1788-org tumor organoid lines that had a diminished Tgf β activity and a reduced ability to invade relative to tumor organoids tested immediately after isolation from C3-TAg tumors (**Figure 4C, Figure 4—figure supplement 2A-C, Supplementary file 1**). Removal of A83-01 or SB431542, a second inhibitor of Tgf β 1, promoted tumor organoid line invasion (**Figure 4D, Figure 4—figure supplement 2B-C**). The invasion of the tumor organoid lines was further enhanced by exogenous TGF β 1 within 48 h (**Figure 4D, Figure 4—figure supplement 2B-C, Figure 4—Videos 1-3**). Persistent exposure of 1339-org cells to Tgf β signaling for up to 22 days progressively increased the rate of collective invasion (**Figure 4E, Figure 4—Videos 1-4**). TGF β signaling for at least 7 days also increased the frequency that trailblazer cells detached from collectively invading cohorts and invaded as single cells, suggesting that at least some single cell invasion reflected a loss of cohesion rather than the activation of a separate invasion program (**Figure 4E, Figure 4—Video 4**).

To understand how the induction of a trailblazer state by Tgf β influenced the properties of cells during primary tumor growth, we injected control 1339-org cells cultured in A83-01 (Control) or 1339-org cells stimulated with Tgf β for 21 days (+Tgf β) into the mammary fat pads of immune-compromised mice. Both Control and +Tgf β 1339-org cells initiated the growth of primary tumors (**Figure 4F**). Interestingly, newly isolated organoids from the Control and +Tgf β tumors were equivalently invasive (**Figure 4G**). This similarity was a consequence of the induction of a trailblazer state in the Control 1339-org primary tumors, which was indicated by the new organoids derived from Control 1339-org tumors being more invasive than Control 1339-org line cells prior to tumor growth (**Figure 4G**). This induction of a trailblazer state was potentially due to stimulation of primary tumor cells by Tgf β in the tumor microenvironment produced my nontumor cells, such as macrophages (40). Consistent with the equivalent invasion of the new primary tumor organoids, the tumors that developed from Control and +Tgf β organoids contained a similar fraction of Vimentin expressing cells (**Figure 4F**). The equivalent invasive behavior of the tumors also corresponded with a similar number of disseminating tumor cells detected in the blood (**Figure 4—figure supplement 3A**). Metastatic tumor cells were not detected in the lung, indicating that the extent of dissemination was not yet sufficient to initiate colonization, consistent with metastasis being a highly inefficient process (**Figure 4—figure supplement 3B**). Thus, Tgf β signaling promoted the conversion of C3-TAg cells to a trailblazer state.

Tgf β induces multiple trailblazer states through distinct phases of re-programming

To understand how Tgf β induced a trailblazer state, we used RNA-sequencing (RNA-seq) to analyze gene expression changes 1339-org cells after 2, 7 and 22 days after Tgf β pathway activation. Hierarchical clustering revealed 4 different groups of genes regulated by TGF β (**Figure 5A, Supplementary file 3**). If cells were undergoing binary conversion between states, the population averaged mRNA expression obtained by RNA-seq would have shown 2 groups of genes (suppressed and induced) undergoing a gradual change in expression as percentage of cells in the 2 different states shifted (41). The 4 different gene groups, each with a unique timing in their onset of change and maximal expression difference, therefore indicated that the Tgf β stimulated C3-TAg tumor cells progressed through different stages of re-programming. Consistent with inducing multiple different states, TGF β promoted a shift in the distribution of EpCAM expression rather than a switch between two unique EpCAM-high and EpCAM-low states, (**Figure 5—figure supplement 1A**). TGF β also induced a shift in

the distribution of Vimentin expression intensity in organoids at day 7 and day 22, as opposed to a switch between two distinct Vimentin^{neg} and Vimentin^{pos} states (Figure 5—figure supplement 1B).

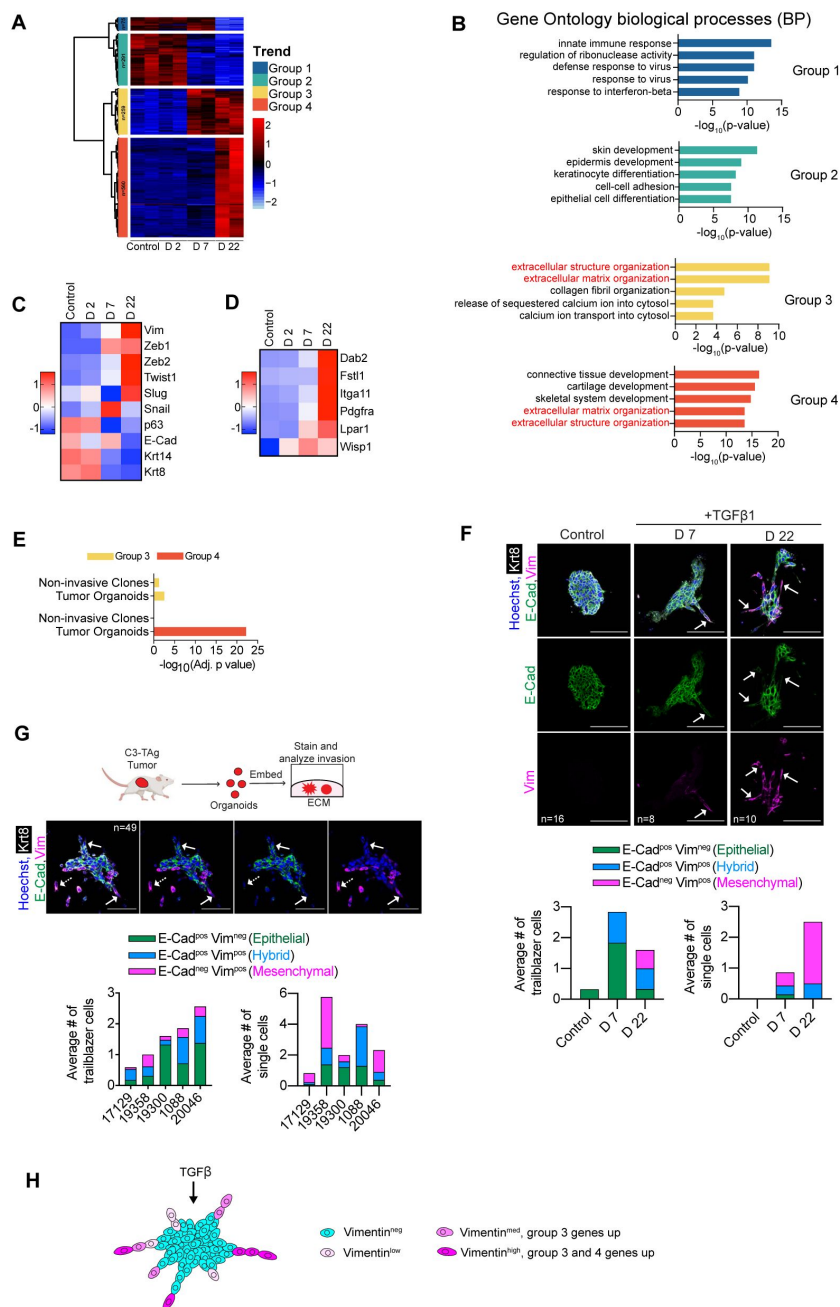


Figure 5 with 1 supplement.

TGFβ1 induces multiple trailblazer states through distinct phases of re-programming.

A. RNA-seq was performed on 1339-org cells treated for 2, 7 and 22 days with TGFβ1. Heatmap shows unsupervised clustering of control and TGFβ1 treated cells along with 4 groups of DEGs. **B.** Graphs show the top 5 Gene Ontology terms associated with each group. **C.** Heatmap shows changes in expression of mesenchymal (Vim), EMT-TF (Zeb1 and 2, Twist1, Snail, Slug), epithelial (E-Cad), basal (p63, Krt14) and luminal (Krt8) markers after TGFβ1 exposure for 0, 2, 7 and 22 days. **D.** Heatmap shows that persistent activation of Tgfβ signaling increases the expression of previously defined trailblazer genes. **E.** Graph shows the enrichment of Group 3 and Group 4 induced genes in non-invasive clones and invasive C3-TAG tumor organoids. **F-G.** Trailblazer cells in multiple different EMT states lead collective invasion in freshly isolated C3-TAG tumor organoids and TGFβ1 treated 1339-org line organoids. Sections from paraffin-embedded organoids were immunostained with Vimentin and E-cadherin antibodies to define different EMT states. The phenotype of trailblazer cells leading invasion or invading single cells was quantified. Bar graphs indicate the mean number of trailblazer or single invading cells per organoid and their EMT phenotype. **H.** Model showing the het-

erogeneity in trailblazer states induced by TGFβ1.

The gene groups identified by RNA-seq were associated with distinct combinations of biological processes (Figure 5B, Supplementary file 3). Of note, genes in Groups 3 and 4 were increased in expression with different kinetics in response to Tgfβ and involved with ECM regulation (Figure 5B, Supplementary file 3). This indicated that there were different mechanisms for regulating genes associated with the same biological functions in response

to TGF β . Group 3 and 4 genes were also connected with unique functions, which indicated that the nature of reprogramming changed over time and was controlled by distinct mechanisms. Interestingly, genes commonly associated with EMT were included in Group 3 (Vcam1, Itgb3, Fn1) and Group 4 (Vimentin, N-cadherin, Periostin) (**Supplementary file 3**). EMT-TFs were also split between Group 3 (Zeb1), Group 4 (Zeb2, Twist), or not included in any group (Snail) (**Figure 5C**). E-cadherin (Group 1) and the mammary epithelial lineage markers p63, Krt8 and Krt14 (Group 2) were also decreased in expression on different time scales (**Figure 5C**). Consistent with these results, high Vimentin and Zeb1 expression and reduced E-cadherin expression was observed in D6 and 1863T trailblazer cells, indicating that these cell lines reflected features of trailblazer cells after persistent activation of Tgf β signaling (**Figure 5—figure supplement 1C-E**). Previously defined trailblazer genes were in Group 3 (Wisp1) and Group 4 (Dab2, Fstl1, Itga11, Pdgfra, Lpar1) (**Figure 5D**). Notably, Group 3 and Group 4 genes were enriched in the genes that were highly expressed in C3-TAG organoids, which indicated that Tgf β re-activated a program that was present in C3-TAG tumors (**Figure 2A, 5E, Supplementary file 3**).

The distinct expression profile of EMT-TFs, EMT effectors and ECM regulatory genes indicated that persistent Tgf β pathway activation induced re-programming yielded multiple different trailblazer states in response to regulatory mechanisms that evolve over time. To test this possibility we immunostained 1339-org and C3-TAG tumor organoids with Vimentin and E-Cadherin antibodies, to specify different states within the spectrum of EMT re-programming. Vimentin^{neg}/E-cadherin^{pos}, Vimentin^{pos}/E-cadherin^{pos} and Vimentin^{pos}/E-cadherin^{neg} cells were all capable of leading collective invasion (**Figure 5F-G**). Notably, the presence of more Vimentin^{pos}/E-cadherin^{pos} and Vimentin^{pos}/E-cadherin^{neg} trailblazer cells with increasing time after Tgf β pathway activation correlated with an increased extent of 1339-org collective invasion (**Figure 5F**). Our combined results suggested that there were multiple trailblazer states induced by Tgf β and that they had different capacities to initiate collective invasion (**Figure 5H**).

The transcription factors Zeb1, Zeb2 and Fra1 confer a trailblazer state

To further understand how Tgf β induced a trailblazer state, we next determined how siRNAs targeting 12 EMT regulatory transcription factors influenced the invasion of 1863T trailblazer cells. Smad3 siRNA transfection served as a positive control. The 1863T trailblazer cells test the requirements for trailblazer cell invasion after long-term activation of the Tgf β pathway (**Figure 6—figure supplement 1A**). The siRNAs targeting Fra1, Zeb1 and Zeb2 suppressed 1863T invasion with a $p < 0.05$ (**Figure 6A**), consistent with Fra1, Zeb1 and Zeb2 being the EMT-TFs that were more highly expressed in C3-TAG tumor organoids (**Figure 2F**). Fra1, Zeb1 and Zeb2 siRNAs also suppressed invasion to a similar extent as Smad3 siRNAs (**Figure 6A**). Thus we prioritized Fra1, Zeb1 and Zeb2 for further investigation.

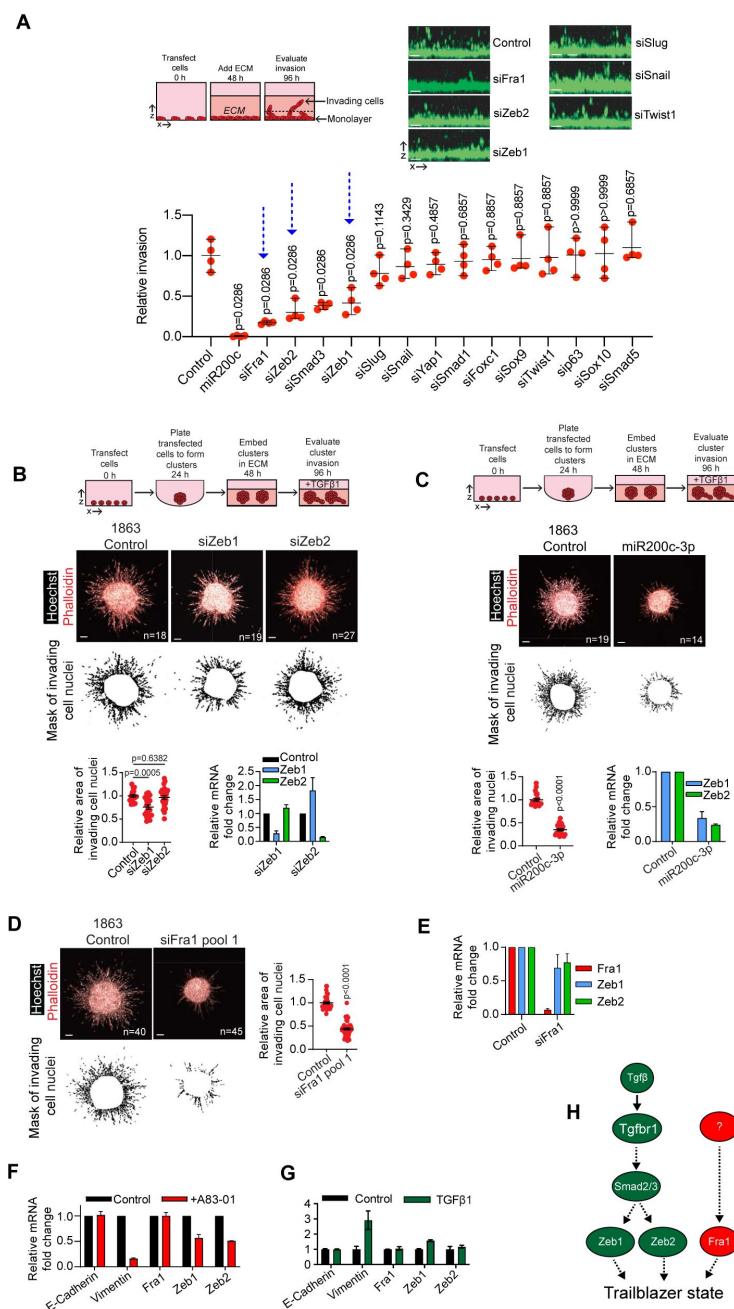


Figure 6 with 4 supplements.

The transcription factors Fra1, Zeb1 and Zeb2 confer a trailblazer state.

A. Invasion of 1863T cells transfected with siRNAs targeting TFs associated with EMT and invasion. Graph shows the quantification of cells vertically invading $\geq 40 \mu\text{m}$ into the ECM (mean $\pm$ SEM, Mann Whitney test, $n=4$ wells from 2 independent experiments). Scale bars, $50 \mu\text{m}$. Smad3 siRNA serves as a positive control for suppression of invasion. Blue arrows indicate genes prioritized for further analysis.

B. Zeb1 siRNA transfection suppressed the invasion of TGF β 1 treated 1863 clusters. Dot plot shows the quantification of invasion (mean $\pm$ SEM, Mann Whitney test, $n=\text{clusters}$, $r=2$). Scale bars, $50 \mu\text{m}$. Bar graph shows Zeb1 and Zeb 2 expression (mean $\pm$ SEM, $n=2$). **C.** Transfection with a miR200c-3p mimic decreased TGF β 1 treated 1863 cluster invasion and Zeb1 and Zeb2 expression. Dot plot shows the quantification of invasion (mean $\pm$ SEM, Mann Whitney test, $n=\text{clusters}$, $r=2$). Scale bars, $50 \mu\text{m}$. Bar graph shows Zeb1 and Zeb 2 expression, (mean $\pm$ SEM, $n=2$). **D.** Fra1 siRNA pool 1 transfection suppressed GF β 1 treated 1863 cluster invasion (mean $\pm$ SEM, Mann Whitney test $n=\text{clusters}$, $r=5$) Scale bars, $50 \mu\text{m}$. **E.** Fra1 depletion did not suppress Zeb1 and Zeb2 expression (mean $\pm$ SEM, $n=2$). **F-G.** E-cadherin, Vimentin, Fra1, Zeb1 and Zeb2 mRNA expression after 48 h treatment with (F) Tgfr1 inhibitor A8301 and (G) TGF β 1 (mean $\pm$ SEM, $n=2$). **I.** Model showing the parallel regulation of the trailblazer state by the Tgf β signaling and a second pathway that induces Fra1.

The Zeb1 and Zeb2 siRNAs suppressed 1863T cluster invasion, demonstrating that Zeb1 and Zeb2 depletion influence 1863T invasion in multiple contexts (Figure 6—figure supplement 1B). Our analysis in 1863T cells determined how Zeb1 and Zeb2 depletion influenced the invasion of trailblazer cells after long-term Tgf β activation. To determine how Zeb1 and Zeb2 depletion influenced the induction of invasion upon the initial Tgf β pathway stimulation, we analyzed the invasion of 1863 cells that were established and propagated in the presence of the Tgfr1 inhibitor A83-01 before experimentation (Figure 6—figure supplement 1A). Interestingly, only Zeb1 depletion suppressed TGF β stimulated 1863 cluster invasion, whereas Zeb2 depletion had a negligible effect (Figure 6B, Figure 6—figure supplement 1A). A miR200a-3p mimic, which suppressed Zeb1 and 2 expression, inhibited 1863 and 1863T invasion, further indicating Zeb1 and 2 contribute to the induction of a trailblazer state (Figure 6C, Figure 6—figure supplement 1C). The requirement of Zeb1, but not Zeb2, for the invasion of 1863 cells upon initial TGF β stimulation was consistent with our results

showing that Zeb1 expression rapidly increased within 7 days while increased Zeb2 expression occurred after day 7 of extrinsic TGF β exposure (**Figure 5C**). Thus, our results suggested that the requirement of Zeb1 and Zeb2 may be dependent on the specific trailblazer state induced by Tgf β signaling.

Additional investigation of Fra1 revealed that Fra1 depletion by 4 individual siRNAs and 2 different siRNA pools reduced 1863T and TGF β stimulated 1863 invasion to a similar degree (**Figure 6D**, **Figure 6—figure supplement 2A-D**). Fra1 depletion also reduced the invasion of D6 trailblazer cells and Tgf β stimulated 1339 cells, which were derived from a different C3-TAg tumor than the 1863 and 1863T cells (**Figure 6—figure supplement 2E-G**). Fra1 interacts with Jun to regulate gene expression as an AP-1 complex (42). Indeed, Jun was required for 1863T and Tgf β stimulated 1863 cell invasion, suggesting Fra1 containing AP-1 complexes regulate the expression of genes necessary for the induction of a trailblazer state (**Figure 6—figure supplement 3A-C**).

Exogenous Fra1 expression promotes Zeb1 and Zeb2 expression in a mouse mammary epithelial cell line, suggesting a potential connection between the functions of Fra1, Zeb1 and Zeb2 (43). However, Fra1 depletion in 1863T cells did not alter Zeb1 or Zeb2 expression, indicating that Zeb1 and Zeb2 are not regulated by endogenous Fra1 in trailblazer cells (**Figure 6E**). In addition, Zeb1, Zeb2 and Vimentin expression were regulated by autocrine TGF β signaling or exogenous TGF β treatment whereas Fra1 expression was not influenced by autocrine or extrinsic TGF β pathway stimulation (**Figure 5C**, **Figure 6F-G**, **Supplementary file 3**). These results indicated that the regulation of Fra1 expression was uncoupled from Tgf β , Zeb1 and Zeb2. Nevertheless, Fra1 was expressed at a higher level in C3-TAg tumor organoids compared to non-invasive spheroids, consistent with increased Fra1 expression contributing to the induction of a trailblazer state (**Figure 2F**). In addition, FRA1 expression was higher in basal like breast cancer compared to other breast cancer subtypes (**Figure 6—figure supplement 4A**). FRA1 expression also correlated with VIM (Vimentin) expression in breast cancer patients, suggesting that Fra1 potentially contributes to induction of a Vimentin specified trailblazer state during tumor progression in basal-like breast cancer patients (**Figure 6—figure supplement 4B**). Thus, our data indicated that the induction of Fra1 and a trailblazer state in C3-TAg tumors required the activity of at least one additional signaling pathway acting in parallel with Tgf β signaling (**Figure 6H**). We therefore selected Fra1 for further analysis to more fully understand the global regulatory framework that empowered a trailblazer state.

Egfr pathway activity promotes Fra1 expression and a trailblazer state

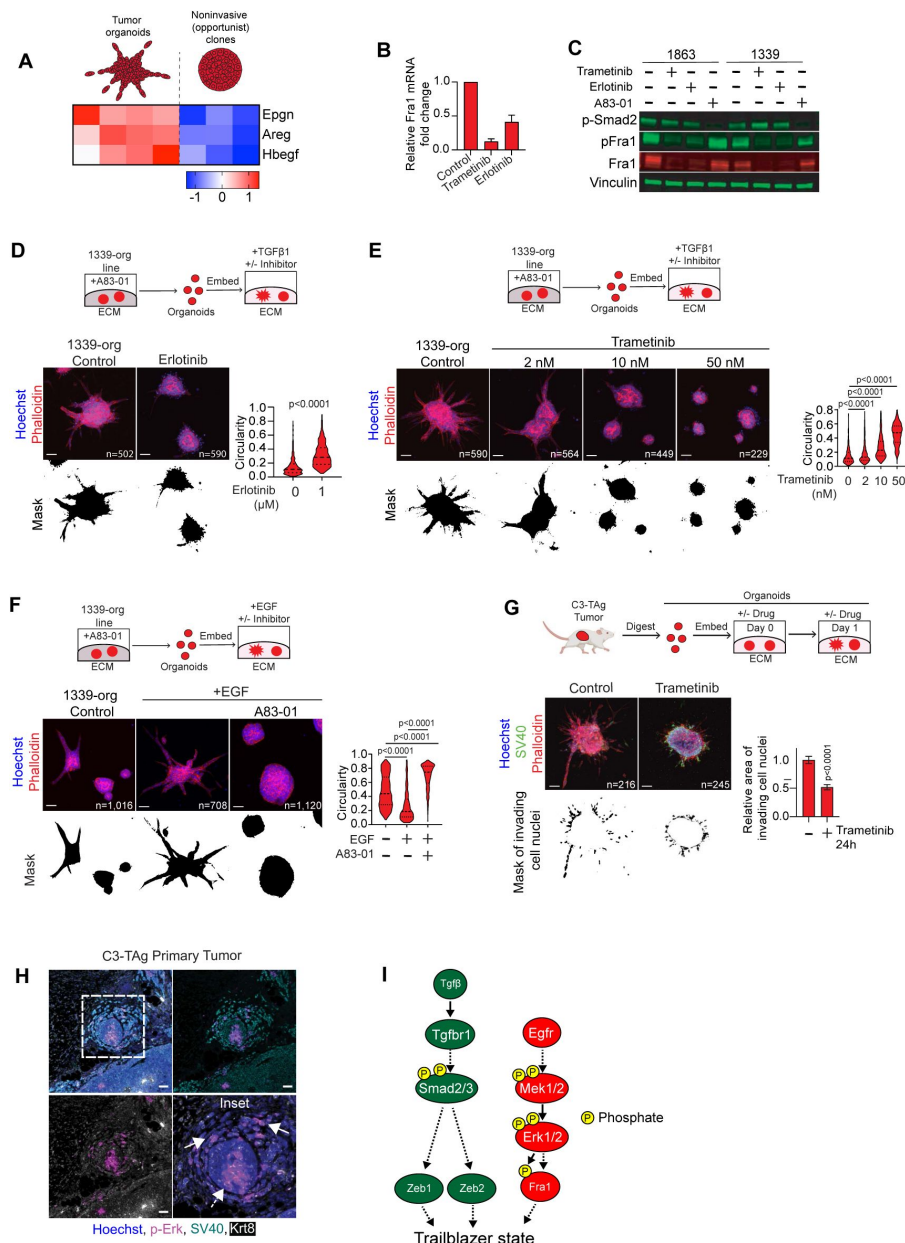
We next determined how Fra1 expression was regulated in C3-TAg tumor cells. Fra1 mRNA was more highly expressed in C3-TAg tumor organoids compared to opportunist spheroids in our RNA-seq analysis, indicating that the signaling pathway responsible for the control of Fra1 was more active in the C3-TAg organoids (**Figure 2F**). GSEA identified multiple active KRAS regulated expression signatures that were more active in the C3-TAg tumor organoids (**Figure 7—figure supplement 1A-B**, **Supplementary file 2**). The Kras gene can be amplified, but is not frequently mutated in C3-TAg tumors, indicating that pathway activation occurred upstream of Kras (44). Analysis of the genes in the active KRAS signature showed that the Egfr ligand Hbegf was more highly expressed in C3-TAg organoids than opportunist clones (**Figure 7A**, **Supplementary file 2**). Like FRA1, HBEGF was more highly expressed in Basal type breast tumors (**Figure 7—figure supplement 1C**). HBEGF expression correlated with FRA1 expression in breast cancer patient tumors as well, consistent with the possibility that Hbegf promotes Fra1 expression during tumor development (**Figure 7—figure supplement 1D**). Further inspection found that the Egfr ligands Areg and Epgn were

also expressed at a higher level in C3-Tag organoids (Figure 7A, Supplementary file 2). These combined results prompted us to investigate how the Egfr-Ras-Mek1/2-Erk1/2 signaling pathway contributed to Fra1 expression and the induction of a trailblazer state.

Figure 7 with 4 supplements.

Egfr signaling regulates Fra1 expression and C3-Tag trailblazer state.

A. Heatmap shows the expression of the Egfr ligands Epigen (Epgn), Amphiregulin (Areg) and Heparin-binding epidermal growth factor (Hbegf) in C3-Tag organoids and non-invasive clones. **B.** The Mek1/2 inhibitor trametinib (10 nM) and Egfr inhibitor erlotinib (1 μ M) reduce Fra1 expression in the 1339-org line (mean $\pm$ SEM, n=2). **C.** Immunoblot shows the response of 1863 and 1339 cells to trametinib, erlotinib and A83-01 treatment (n=2). **D.** Erlotinib (1 μ M) suppresses the TGF β 1 induced invasion the 1339-org line. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=3) Scale bars, 50 μ m. **E.** Trametinib suppresses the TGF β 1 induced invasion of the 1339-org line in a dose dependent manner. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=3). Scale bars, 50 μ m. **F.** A83-01 suppresses the EGF induced invasion in the 1339-org line. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=4) Scale bars, 50 μ m. **G.** Trametinib suppresses the invasion of freshly derived C3-



Tag organoids within 24 h of original tumor processing. Graph shows the quantification of invasion from 4 different tumors (mean $\pm$ SEM, Mann Whitney test, n=organoids) Scale bars, 50 μ m. **H.** Images show dually phosphorylated Erk1/2 immunostaining in regions of high invasion (solid arrows) and low invasion (dotted arrow) in primary C3-Tag tumors. Scale bars, 50 μ m. **I.** Model showing the parallel regulation of trailblazer cells by signaling pathways coordinated by Tgf β and Egfr.

Erlotinib (Egfr inhibitor) and trametinib (Mek1/2 inhibitor) suppressed Fra1 mRNA and protein expression in C3-Tag cell lines and organoid lines (Figure 7B-C, Figure 7—figure

supplement 2A). By comparison, Jun expression was unchanged by treatment with either inhibitor (**Figure 7—figure supplement 2B**). In addition, erlotinib and trametinib did not inhibit the activating phosphorylation of Smad2 or Zeb1 expression, indicating that the Egfr1-Mek1/2-Erk1/2-Fra1 did not influence cell state through autocrine Tgf β signaling (**Figure 7C**, **Figure 7—figure supplement 2C**). A83-01 suppressed Smad2 phosphorylation, but did not influence Fra1 expression and phosphorylation, further indicating that Tgf β signaling did not regulate Fra1 (**Figure 7C**).

Erlotinib treatment reduced the TGF β stimulated invasion of 1339-org cells (**Figure 7D**). Trametinib treatment also blocked the initial stimulation of 1339-org invasion by Tgf β , the invasion of 1339-org trailblazer cells following persistent Tgf β treatment and the invasion D6 trailblazer cells (**Figure 7E**, **Figure 7—figure supplement 2D-E**). Moreover, extrinsic EGF induced the invasion of 1339-org, 1863-org and 1788-org cells and this enhancement was blocked by A83-01, further indicating that the integration of Tgfbr1 and Egfr pathways promoted a trailblazer state (**Figure 7F**, **Figure 7—figure supplement 2F-G**). Trametinib inhibited the invasion of C3-Tag primary tumor organoids and phosphorylated Erk1/2 was detected in collectively invading cells in C3-TAg primary tumors (**Figure 7G-H**). Erk1/2 activity is pulsatile in response to EGF stimulation, and these pulses promote Fra1 expression (45). Thus, persistent Erk1/2 phosphorylation was not necessarily expected in all collectively invading cells or to promote Fra1 expression. Consistent with these results, A83-01, erlotinib and trametinib suppressed the invasion of human MCF10A-Traillblazer cells and MCFDCIS cells, which harbor features of basal-like breast cancer (25, 46). This indicated that the integration of Tgf β and Egfr pathway signaling was necessary for the induction of a trailblazer state in multiple contexts (**Figure 7—figure supplement 3A-C**, **Figure 7—figure supplement 4A-B**). Together, these results indicate that the regulation of Fra1 by Egfr and Erk1/2 promotes the induction of a trailblazer EMT (**Figure 7I**).

A compromise between Tgfbr1 and Egfr regulated signaling programs confers the trailblazer state

To understand the contribution of the Tgfbr1 and Egfr regulated signaling pathways towards the induction of the trailblazer state, we performed RNA-seq on organoids treated with diluent, TGF β , TGF β +erlotinib, EGF and EGF+A83-01 (**Figure 8A**, **Figure 8—figure supplement 1A**). Polar plots were used to cluster genes based on the nature and the degree of integration between the Tgfbr1 and Egfr signaling pathways (47). This analysis revealed multiple modes of interaction between the two pathways in response to either extrinsic TGF β or EGF and that the type of interaction (positive or negative) was associated with different biological processes (**Figure 8B-C**, **Figure 8—figure supplement 1B-F**, **Supplementary files 4 and 5**). Thus, the relationship was far more complex than the simple positive signal integration previously suggested by phenotypic analysis and the expression of specific canonical EMT program genes (48, 49). In addition, the gene expression changes induced by extrinsic TGF β and EGF were associated with unique combinations of biological processes, indicating that extrinsic TGF β and extrinsic EGF induced distinct types of trailblazer cells (**Figure 8C**, **Figure 8—figure supplement 1F**). Nevertheless, there were conserved features of pathway integration that influenced the regulation of processes associated with invasion regardless of the external stimulus. Genes involved in controlling the organization and composition of the ECM, including previously defined trailblazer genes, were predominantly induced by Tgfbr1 and not positively impacted by Egfr signaling (**Figure 8C**, **Figure 8—figure supplement 1F**, **Supplementary files 4 and 5**). Notably, autocrine and extrinsic Egfr activation suppressed a number of Tgfbr1 regulated ECM reorganization genes and Hallmark EMT genes, including Postn, Bmp2, Itgb3, Orai2, Mmp11 and Erg (**Figure 8D**, **Figure 8—figure supplement 1G**, **Supplementary files 4 and 5**), which can contribute to invasion and metastasis (50-55). By comparison, autocrine or extrinsic EGF

TGF β 1+erlotinib. Also see [Figure 8—figure supplement 1C-D](#) and [Supplementary file 5](#) for the assignment of polar plot clusters and detailed biological processes used to define the summary integration groups. **D.** Egfr suppressed the expression of a subset of genes induced by extrinsic TGF β 1. Line plots indicate the expression of genes in Cluster 1. Dashed line indicates the average scaled expression of all Cluster 1 genes. Lower bar graph shows the top 10 biological processes (GO-BP) associated with Cluster 1 genes (Fisher's exact test). **E.** Egfr was necessary for the expression of a subset of genes induced by extrinsic TGF β . Line plots show the expression of genes in Cluster 5. Dashed line indicates the average scaled expression of all Cluster 5 genes. Lower bar graph shows the top 10 biological processes (GO-BP) associated with Cluster 5 genes (Fisher's exact test). **F.** Graph of Axl expression (mean $\pm$ SEM, $n=2$). **G.** The Axl inhibitor BGB324 (1 μ M) suppressed freshly isolated C3-TAg tumor organoid invasion. Graph shows the quantification of invasion (mean $\pm$ SEM, Mann Whitney test, n =organoids, $r=2$). **H.** Model showing how the induction of the trailblazer state involves a compromise between the Tgf β and Egfr regulatory programs. The Egfr dependent induction of genes, such as Axl, that are required for trailblazer cell motility has the cost of restricting the expression ECM remodeling genes known to promote invasion and metastasis.

Discussion

Our findings indicate that defining how cells metastasize requires an understanding of how trailblazer cells are induced in different breast tumors ([Figure 9](#)). The Tgf β regulated trailblazer program that we have defined in the basal-like C3-TAg mammary tumor cells induced a more aggressive form of collective invasion than was detected in luminal B PyMT tumors and functioned independently of increased Krt14 and p63 expression. This indicated that the Tgf β program was mechanistically and phenotypically distinct from a previously described Krt14-high trailblazer program (20). In addition, PyMT tumor cells rarely expressed detectable Vimentin, which suggested that the Tgf β regulated trailblazer program that we identified was not active in PyMT tumors. Thus, our results indicate that there are at least 2 different trailblazer programs and that their activity and requirement for invasion varies between different types of mammary tumors associated with previously defined breast cancer intrinsic subtypes ([Figure 9](#)). Genes induced as part of the C3-TAg trailblazer program are more frequently detected at a high level in basal-like breast cancer patients. Vimentin expression also correlated with collective invasion in TNBC patients, which are most often classified as basal-like. Vimentin expression was detected in basal-like TNBC PDX metastases as well. In addition, genes induced as part of the Tgf β trailblazer program have previously been shown to correlate with reduced odds of survival in TNBC, but not patients classified as ER-positive or HER2-positive (21). These combined results indicate that the Tgf β trailblazer program is active in basal-like TNBC and influences patient outcome. Tumor subtype specific mechanisms for inducing a trailblazer state are potentially a general property of cancer progression, as molecular subtypes are a feature of many carcinomas (61-64). Further investigation is required to determine the extent to which the Tgf β and Krt14 trailblazer programs align with existing tumor subtypes or define new categories of inter-tumor heterogeneity. Thus, our results provide new insight into approaches used to stratify patients into outcome groups.

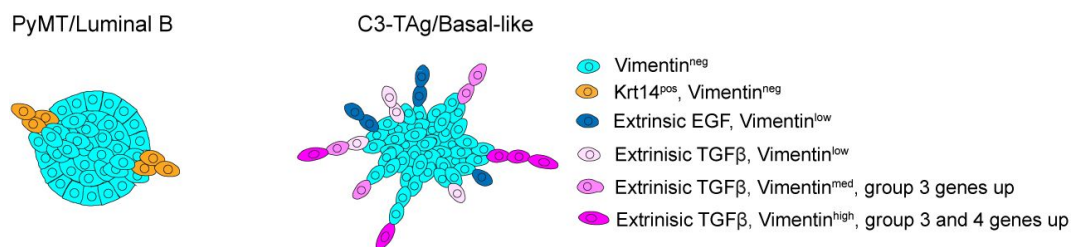


Figure 9.

Model of inter- and intra-tumor trailblazer cell heterogeneity.

The characteristic features and functional requirements of the trailblazer cells induced by Tgfβ and Egfr signaling programs in C3-TAg/basal-like tumors are fundamentally distinct from the trailblazer cells in PyMT/luminal B tumors.

The heterogeneity in C3-TAg tumor organoid invasion indicated that the trailblazer population itself can be heterogeneous within tumors with the same primary driving genetic factors and molecular subtype (Figure 9). Previous analysis of other tumor cell models had suggested that trailblazer cells were in a single uniform state (20, 21, 24). Notably, our analysis of 1339-org cells revealed that invasive variability can be a consequence of differences in cellular responses to a single stimulus, in this case TGFβ signaling. Therefore trailblazer cell phenotypic heterogeneity within a population did not necessarily reflect differences between EMT programs that were initiated by different activating signals (65). Gene expression changes in response to TGFβ are heterogeneous and temporally regulated in other systems as well (66). However, the influence of this gene expression heterogeneity on cell phenotype has not been clear and has often been modeled as a progressive accumulation of new functions (67). The phased changes in gene expression we defined in 1339-org cells suggested how trailblazer phenotypes evolved over time in response to TGFβ activity, with early onset gene expression changes inducing trailblazer cells that were further modulated by later stage reprogramming events (Figure 9). Thus, differences between tumor cells in the duration of TGFβ pathway activation may contribute to the heterogeneity we detect in trailblazer populations. Interestingly, long-term TGFβ induced changes in gene expression that were associated with other cell lineages. This suggests that the enhanced invasion of trailblazer cells in response to persistent TGFβ stimulation may be associated with a more general lineage plasticity and not simply a re-balancing of epithelial and mesenchymal gene expression.

Our gene expression analysis and functional studies have begun to unravel the details of the transcriptional regulation of the trailblazer state. We detected an increase in Fra1, Zeb1 and Zeb2 expression in C3-TAg trailblazer cells and human breast cancer trailblazer cell lines, with Fra1 regulated by Egfr signaling and Zeb1 and Zeb2 regulated by Tgfβ. Fra1 and Zeb1 were required for the initial Tgfβ induced induction of a trailblazer invasion as well as invasion of cells after they had converted to trailblazer state. Interestingly, Zeb2, which showed a delayed induction in response to Tgfβ, was only required for the invasion of a trailblazer cell line with the characteristics of long-term Tgfβ pathway activation. These results suggest that the transcriptional regulation of trailblazer populations may be different depending on the specific trailblazer state. We also determined how siRNAs targeting 10 additional transcription factors, including the canonical EMT TFs Snail (Snai1), Slug (Snai2) and Twist1, influenced trailblazer cell invasion. The siRNAs targeting the 10 TFs did not yield a clear suppression of invasion and we did not detect a correlation between the increased

expression of the TFs and the induction of a trailblazer state. However, our results do not exclude the possibility that these TFs contribute to the Tgf β dependent induction of a trailblazer state in some context.

Our results revealed that a new compromise mode of integration between Tgfbr1 and Egfr regulated gene expression programs. The integration of parallel signaling pathways is frequently required to confer specific cell states and is typically classified from the standpoint of being a positive or negative interaction (49, 68, 69). Indeed, prior studies have suggested that Tgf β and Egfr-Ras regulated pathways positively integrate at the point of EMT-TF expression to initiate an EMT program (48, 49, 70, 71). While the combined activity of the Tgfbr1 and Egfr signaling pathways was necessary to induce a trailblazer state, there was an equivalent degree of positive and antagonistic interaction between the programs in C3-TAg tumor cells. This finding revealed a new type of integration between the two signaling programs in which the amplitude of a subset of the gene expression changes induced by either pathway alone are compromised. The selective activation and restriction of Tgf β and Egfr dependent gene expression suggests that compromise integration is dictated by differences in the expression or activity of co-factors that modulate chromatin accessibility or DNA binding. Tgf β can increase Egfr expression and activate Erk1/2 signaling in a cell line dependent manner (72, 73). Similarly, Egfr-Mek1/2-Erk1/2 signaling can activate Tgf β signaling through the phosphorylation of SMAD2/3 and increased expression of Tgf β expression depending on cell context (43, 74, 75). However, this manner of integration is unlikely contributing to the compromise integration in the C3-TAg tumor cells given that Tgfbr1 inhibition did not alter ERK1/2 phosphorylation or Fra1 expression and Egfr and Mek1/2 inhibition failed to impact Smad2/3 phosphorylation. The direct control of the activity of one pathway by the other would also be expected to yield a more uniform positive or negative integration. Defining the mechanistic basis of selective antagonism of gene expression by the Tgf β and Egfr pathways is an important line of future investigation.

The requirement of Tgf β and Egfr pathway activity for trailblazer cell invasion indicated that the compromise mode of integration was a feature of the trailblazer state. Egfr pathway stimulation can also promote the motility of non-trailblazer cells, raising the possibility that Egfr signaling has distinct functions in different subpopulations of invading cells, with the precise regulatory contribution dictated by Tgf β pathway activity (76, 77). The requirement of Tgf β and Egfr signaling for the collective invasion of MCF10A-Tailblazer and MCFDCIS cells indicates that the integration of the pathways is essential for the induction of a trailblazer state in multiple contexts. However, the Egfr ligand Epgn restricts the invasion of luminal B PyMT organoids, potentially due to the ability of Epgn to suppress the expression of EMT-TFs and ECM reorganization genes in PyMT cells (78). This result suggests that integration of Egfr signaling is not essential for the induction of Krt14 specified trailblazer cells, like those detected in PyMT tumors (20). The differences in how Egfr signaling influences invasion may be a consequence of how tumor cells obtain the ability to resist TGF β induced growth suppression and apoptosis (79). This is accomplished either through the clonal selection of mutants that inactivate Tgf β signaling, which does not occur in C3-TAg tumor cells, or an intrinsic reprogramming of cellular responses to TGF β (80). The type of reprogramming to TGF β stimulation that is selected for during tumor progression may determine whether Tgfbr1 and Egfr signaling engage in a compromise mode of integration or if there is a predominantly positive or negative interaction between the pathways.

Our findings indicate that defining the nature of Tgf β and Egfr pathway integration is necessary to understand tumor cell properties that influence metastasis and treatment response. In C3-TAg cells in which Tgf β and Egfr activity combined to promote a trailblazer state, Egfr also restricted the expression of genes that could endow trailblazer cells with additional invasive and metastatic properties (53, 81-84). Similarly, Tgf β limited the Egfr induced expression of chemokines that increase invasion and the recruitment of tumor

promoting immune cells (85, 86). Thus, the compromise mode of integration limited the expression of genes that could promote tumor progression whereas a strictly positive integration would have no such restrictions. The type of interaction between the Tgf β and Egfr programs also influences the rationale for the therapeutic targeting of trailblazer cells. Inhibiting the Egfr pathway with the clinically approved drugs erlotinib and trametinib suppressed C3-TAg tumor cell invasion (87-89). However, combining either drug with newly developed anti-Tgf β antibodies (90) may have additional benefit due to the fact that the Egfr and Tgf β pathways function in parallel and each independently regulate large cohorts of genes. By comparison, cells in which Tgf β activates Egfr-ERK1/2 signaling or vice versa may derive little or no benefit from combination therapy (72, 73). Thus, understanding how the mode of Tgf β and Egfr pathway integration influences invasion, the risk of metastasis and options for treatment is an important line of future investigation.

Together, our results indicate that an improved understanding of how trailblazer cells are regulated and function in different breast cancer subtypes has the potential to improve the accuracy of patient diagnosis and reveal new treatment options.

Materials and methods

Mice

The C3-TAg [FVB-Tg(C3-1-TAg)cJeg/Jeg], PyMT [FVB/N-Tg(MMTV-PyVT)634Mul/J] and NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice were purchased from The Jackson Laboratory (C3-TAg: 013591, PyMT: 002374 and NSG: 005557) and (**Supplementary file 1**). Mice were housed, bred and euthanized in accordance with a protocol approved by the Institutional Animal Use and Care Committee at Georgetown University (IRB# 2017-0076) and in compliance with the NIH Guide for the Care and Use of Laboratory animals. Female mice were used for all experiments.

Patient derived xenografts

Triple negative breast cancer (TNBC) PDX fragments (**Supplementary file 1**) were obtained from Dr. Alana Welm at University of Utah (39). A single fragment of HCI-001 PDX was transplanted into the cleared 4th mammary fat pad of 4 weeks old NSG female mice. Tumor measurements were taken every week to monitor tumor growth. All surgical procedures and euthanasia was performed in accordance with IACUC approved protocols at Georgetown University. Tumor pieces and lungs were fixed in formalin followed by paraffin embedding and sectioning by Histology and Tissue Shared Resource (HTSR) at Georgetown University.

Tumor organoid derivation

The largest tumors from female C3-TAg and PyMT mice were minced and dissociated for up to 120 min at 37°C in a mixture of 1 mg/ml Collagenase (Worthington, LS004188), 20 U/ml DNase (Invitrogen, 18047-019), 5% Fetal Bovine Serum (FBS) (Peak Serum, PS-FB2), ITS (Lonza, 17-838Z), non-essential amino acids (Sigma, M7145) and 10 ng/ml FGF-2 (Peprotech, 100-18C) in DMEM/F-12 (Corning, 10-092-CV). Dissociated tumors were pelleted at 80 x g for 2 min and the supernatant was discarded. Tumor organoids were then rinsed up to 5 times in 5% FBS in DMEM/F-12 followed by filtering through a 250 μ m tissue strainer. Organoids were then tested in invasion assays or to establish organoid lines and cell lines (**supplementary file 1**).

Cell and organoid culture

Cell and organoid lines were generated from C3-TAg tumors (**supplementary file 1**) and tested for mycoplasma (Lonza, LT07-703) prior to use and the creation of frozen stocks. Cells and organoids were routinely used within 25 passages. Culture conditions for all cell and organoid lines are summarized in **Supplementary file 1**. Organoids were passaged at least once per week by dissociating tumor organoids into single cell suspensions with Dispase (Sigma, SCM133) and TryPLE (Gibco, 12605-010). The cell suspensions were plated at a density of 200,000-500,000 cells per well in a 24-well ultra-low adhesion plate (Sigma, CLS3474) to form multicellular aggregates for at least 16 h. The aggregates were then embedded in Matrigel (Corning, 354230) or Cultrex (Biotechne, 3533-005-02) and overlaid with Organoid Media (**Supplementary file 1**) in 6-well ultra-low adhesion plates (Sigma, CLS3471). H2B:GFP retrovirus (91) and H2B:mCherry and GFP lentivirus were produced and cells were infected as described previously (77).

Orthotopic Tumor Models

For orthotopic transplantation experiments, 1339-org cells were first cultured in Organoid media (Control) or Organoid media without A83-01 and supplemented with 2 ng/ml TGF β 1 (+TGF β 1). After 21 days of treatment, 500,000 control or +TGF β 1 1339-org cells were injected in the right 4th fat pad of 8-12 weeks old female NSG mice. Seven mice per group from 2 independent experiments were analyzed. Mice were euthanized and tissue collection was performed 34 days post injection. Representative portions of primary tumors and complete lungs were fixed in formalin followed by paraffin embedding and sectioning by Histology and Tissue Shared Resource (HTSR) at Georgetown University. The remaining primary tumor portions were used for new organoid derivation and analysis.

Blood collection and Circulating Tumor Cell (CTC) isolation

For CTC isolation, mice bearing Control or +TGF β 1 1339-org orthotopic tumors were anesthetized with isoflurane and 800 μ l of blood was extracted via cardiac puncture using a 25 G needle (BD, 305122). The blood was immediately transferred to EDTA coated micro tubes (Sarstedt, 4111395105) to avoid coagulation. Blood collected from all mice from the same treatment group (Control or +TGF β 1) were pooled together, mixed with red blood cell (RBC) lysis buffer (Thermo Fisher, 00433357) and incubated on ice for 10 min. The samples were then spun down at 4°C for 10 min at 250 g. The supernatant was discarded and the same steps were repeated once more to completely remove the RBCs. The cell pellet was then resuspended in cold 1X PBS and plated in poly-l-lysine pre-coated 8 well chamber slides (0.4 mL per well). The slide was placed in the 37°C incubator for at least 30 min. The cells were then fixed, permeabilized, blocked and stained with anti-Krt8 antibody to detect circulating C3-TAg tumor cells.

Organoid invasion

Organoids in 30 μ l of ECM were plated onto a 20 μ l base layer of ECM in 8 well chamber slides (Falcon, 354108) and allowed to invade for 48 h unless otherwise indicated. The ECM was a mixture of 2.4 mg/ml rat tail collagen I (Corning, CB-40236) and 2 mg/ml of reconstituted basement membrane (Matrigel or Cultrex) unless otherwise indicated. Tumor organoids were analyzed in a base media of DMEM/F-12 and ITS in **Figure 1A, D, E and G, Figure 1—figure supplement 1B, Figure 4B, C and G, Figure 5G, Figure 7G and Figure 8G**, and 1% FBS + 10 ng/ml FGF2 in **Figure 1B and F**. A base media of DMEM/F-12 and ITS was used for the analysis of 1339-org, 1863-org and 1788-org lines. Growth factors (2 ng/ml

TGF β 1; 50 ng/ml EGF) and inhibitors were added to the base media and tested as indicated in the text and **Supplementary file 1**. Organoids from established organoid lines (**Supplementary file 1**) were tested after dissociation and aggregation for at least 16 h as described for organoid passaging. Organoids were fixed and stained with Hoechst and Phalloidin. Images were acquired with a Zeiss LSM800 laser scanning confocal microscope using 10x/0.45 (Zeiss, 1 420640-9900-000) or 20x/0.8 (Zeiss, 1 420650-9902-000) objectives. To determine the area of invasion and circularity, an image mask for each organoid was generated using ImageJ (NIH) based on the Hoechst (invasion area) or Phalloidin (circularity) signal. The total area of the invading cell nuclei was determined using the “Measure” function in ImageJ. Circularity was quantified using the “Analyze Particle” function.

siRNA experiments

Cells were transfected with 50 nM of siRNA using RNAiMax transfection reagent (Invitrogen, 13778-150) for 48 h before testing. Cells in all conditions designated as “Control” were transfected with a pool of siRNAs that does not target mouse genes. The catalog numbers for each siRNA are located in **Supplementary file 1**.

Vertical invasion assay

Vertical invasion assays were performed and quantified as described previously using a Collagen/Basement Membrane (BM) mix (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex) (25). Cells expressing H2B:GFP (10,000-18,000/well) were transfected in 96-well plates (Thermo, 165305) for 48 h. Media was then replaced with 50 μ l Collagen/BM) ECM and overlaid with fresh media (DMEM-F12, ITS, 1% FBS, 10 ng/ml FGF2) for 48 h. Cells were imaged with a 10X/0.30 objective (Zeiss, 1 420340-9901-000) on a Zeiss LSM800 laser scanning confocal microscope. Twenty-one z-slices at 10 μ m intervals over a total span of 200 μ m were acquired for 9 tiled x, y positions per well. At least 3 wells (27 x,y positions total) were imaged per condition in each experiment. ImageJ software (NIH) was used to process images and quantify invasion using a custom macro as described (21). Invasion was normalized by dividing the total number of cells that had migrated $\geq$ 40 μ m above the monolayer by the number of cells in the monolayer. The normalized invasion value for each condition was divided by the normalized invasion for the control wells transfected with an siRNA pool that does not target any mouse genes to provide a relative invasion value.

Cluster invasion

Cells (10,000-18,000/well) were transfected in 96-well plates (Corning, 3904). After 24 h, the transfected cells were trypsinized and clustered in 96-well Nunclon Sphera U-bottom low adhesion plates (Thermo Scientific, 174925) overnight (1000 cells/well). Homogeneous and mosaic clusters were also generated in U-bottom Nunclon Sphera low adhesion plates. Cells then aggregated into clusters for at least 16 h before embedding in ECM. Clusters were resuspended in 30 μ l Collagen I/BM mix (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex), plated on 20 μ l of a base layer of Matrigel/Collagen I, overlaid with media containing 1% FBS and 2 ng/ml TGF β 1 and allowed to invade for 24 h-48 h unless stated otherwise. The spheroids were fixed and stained with Hoechst and Phalloidin and imaged with a 10x/0.45 objective (Zeiss, 1 420640-9900-000) using a Zeiss LSM800 laser scanning confocal microscope. To determine the area of invasion, Hoechst staining was used to identify and mask nuclei invading into the ECM using ImageJ. The total area of the invading nuclei was then determined using the “Measure” function in ImageJ.

Live cell imaging

Imaging was performed using a Zeiss LSM800 laser scanning confocal microscope enclosed in a 37° C chamber supplemented with humidified CO₂. Images were acquired every 30 min for 14-18 h with a 10X/0.30 objective (Zeiss, 1 420340-9901-000). At least 15 different x,y coordinates with 5 to 7 z-slices over 50-100 µm span for each condition were acquired in parallel.

Immunofluorescence and immunoblotting

Experiments and analysis were performed as described (25) using antibodies detailed in **Supplementary file 1**. Immunoblots were imaged using an Odyssey scanner (Licor, 9120). Immunofluorescence images were acquired using 10x/0.45 (Zeiss, 1 420640-9900-000) and 20x/0.8 (Zeiss, 1 420650-9902-000) objectives. For analysis of the relationship between Krt14 expression and invasion, organoids were classified as invasive if they had at least two protrusions with each protrusion having at least two cells invading outward from the center mass. Images were exported as TIFFs and analyzed with ImageJ. The freehand selection tool was used to define ROIs containing individual organoids for analysis. The Krt14 fluorescence for each organoid was then defined as the “Mean Gray Value”, as determined using the “Measure” function. To define Krt14 expression in cells leading invasion, >10 organoids stained with Phalloidin, Hoechst, Krt14 and Krt8 were analyzed. Seven Z slices at 10 µm intervals over a total span of 100 µm were acquired, and the leading cells of each protrusion were designated as Krt14^{low} or Krt14^{high}. Vimentin and E-cadherin expression were analyzed in 5 µm histology sections. Organoids in 8-well chamber slides (Falcon, 354108) were washed 3x with distilled water. After detaching the chamber walls from the slides, the organoid/ECM gels were gently picked up using a single edge blade (Personna, 94-120-71) and laid out on a strip of parafilm (Bemis, PM-999). The 8-well chamber walls were placed around the gels and 150-200 µl of hydroxyethyl agarose processing gel (Thermo Fisher, HG 4000-012) was poured into the well. After the hydroxyethyl agarose processing gel solidified, it was transferred into histosettes (Simport, M498-3) and stored in 70% EtOH until embedding in paraffin by the Georgetown Histology and Tissue Shared Resource. The 5 µm sections were deparaffinized and immunostained as described (21) to detect Vimentin and E-cadherin expression (**Table S1**). For individual organoids, Vimentin expression was defined as the “Mean Gray Value” using the “Measure” function in ImageJ. Using ImageJ, Vimentin expression was defined as the “Mean Gray Value” for individual organoids using the “Measure” function. For characterization of the trailblazer cell state, the cell leading a collectively invading group of cells was classified as a “trailblazer cell” while cells detached from the tumor organoid were classified as “single cells”. Each cell was assigned E-cadherin^{pos}/Vimentin^{neg}, E-Cadherin^{pos}/Vimentin^{pos} or E-Cadherin^{neg}/Vimentin^{pos} status based on E-Cadherin and Vimentin expression.

Immunohistochemistry

Sample processing and staining was performed as described (25). Images were acquired using 10x/0.45 (Zeiss, 1 420640-9900-000) and 20x/0.8 (Zeiss, 1 420650-9902-000) objectives. At least 10 regions of interest (ROIs) were randomly selected for each tumor. Cell Profiler (Broad Institute) was used to define image masks based on Krt8, Krt14 or Vimentin signal. To define their expression in tumor cells, masks based on the area of overlap between the Krt14 and Krt8 and the Vimentin and Krt8 signals were determined for each ROI. The area of individual channel and overlap masks created in CellProfiler were then loaded onto ImageJ and quantified using the “Measure” function.

Quantitative real-time PCR

Total RNA was isolated using the RNeasy Kit (Qiagen, 74104) or the PureLink™ RNA Mini Kit (Invitrogen, 12183025). Reverse transcription was performed using iScript cDNA Synthesis Kit (Bio-Rad, 1708891) following the manufacturer's protocol. Real time PCR was performed with iQ SYBR Green Supermix (Bio-Rad, 1708882) using a RealPlex 2 thermocycler (Eppendorf, 2894). All primers were designed and purchased from Integrated DNA Technologies and their sequences are listed in **Supplementary file 1**.

Flow Cytometry

Cells were dissociated with TrypLe (Gibco, 12605-010) into single cell suspensions, and then washed 2x with FACs buffer (PBS + 2% FBS). Cells were then incubated with anti-CD326 (EpCAM)-PE/Dazzle 594 (Biolegend, 118236) at 4°C for 30 minutes. After staining, cells were washed with FACs buffer 2x, and stained with Helix NP Blue (Biolegend, 425305) in order to exclude dead cells. Cells were then subjected to flow cytometry analysis using BD LSR Fortessa. Results were analyzed using FCSEXPRESS 7 (De Novo Software).

RNA sequencing

For the analysis of C3-TAG tumor organoids and non-invasive spheroids in **Figure 2** and **Figure 2-figure supplement 1**, at least 1000 C3-TAG tumor organoids or 60,000-80,000 non-invasive cells were embedded in a 500 µl ECM mixture (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex) and plated on top of 500 µl base layer of ECM in 6 cm dishes (Falcon, 353004). C3-TAG tumor organoids were allowed to invade for 48 h. Non-invasive cells were allowed to grow and form spheroids for 6 days. Experiments were performed twice using frozen stocks of tumors and cell lines. For the analysis of 1339-org responses to growth factors and inhibitors in **Figure 4**, **Figure 8** and **Figure 8-figure supplement 1**, 250,000 1339-org cells were embedded in a 100 µl ECM mixture (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex) and plated on top of a 200 µl ECM base layer in 6 cm dishes (Falcon, 353004). The gels were overlaid with control or treatment media. Organoid/ECM gels were collected in 15 mL conical tubes (Fisher, 339651), snap frozen and stored in -80° C freezer. Experiments were performed twice. For RNA isolation in all experiments, the gels were thawed on ice and disintegrated in 1 mL Trizol (Invitrogen, 15596026) by repeatedly passing the mixture through a 23G needle (BD, 305143). After the gels were homogenized, the Trizol-gel mixture was allowed to sit at RT for 5 min. Then 200 µL of chloroform (Sigma-Aldrich, C2432-25ML) was added. Following vigorous shaking for 15 sec, the mix was left for 2 min at RT to allow for phase separation. The mix was then centrifuged at 12,000 x g for 15 min at 4° C. The upper aqueous phase was then transferred to a new tube and RNA was then further purified using the RNeasy Kit (Qiagen, 74104) and or PureLink™ RNA Mini Kit (Invitrogen, 12183025). For **Figure 2** and **Figure 2-figure supplement 1**, RNA sequencing, was performed by the Genomics and Epigenomics Shared Resource at Georgetown University. The cDNA libraries were generated using the Illumina TruSeq Stranded Total RNA Library Preparation Kit (Illumina, 20020597). Sequencing was performed using an Illumina NextSeq 500 at a depth of 12 million paired-end 75 base pair reads per sample. Data is available from GEO at GSE188312. Reads were quantified using Salmon with mouse reference genome (GRCm38/mm10). Counts were normalized and differential expression was calculated using DESeq2. For **Figure 4**, **Figure 8** and **Figure 8-figure supplement 1**, RNA sequencing was performed by Novogene. RNA quality analysis was assessed by Bioanalyzer using the RNA Nano 6000 Assay Kit (Agilent, 5067-1511). The cDNA libraries were prepared and indexed using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, E7530L). Libraries were sequenced with the NovaSeq 6000 at a depth of 20 million paired-end 150 base pair reads

per sample. Sequencing data available from the GEO at GSE188300 and GSE188323. Reads were mapped to the mouse reference genome (GRCm38/mm10) using HISAT2 and reads per gene determined using HTSeq. Counts were normalized and differential expression was calculated using DESeq2.

Polar plot analysis

Expression data was normalized and differentially expressed genes with a ≥ 2 -fold change in expression with $p < 0.05$ after Benjamini-Hochberg (BH) correction were identified using the *DESeq2* package in R (92, 93). Normalized expression data was collapsed for all genes as the average between condition replicates and were subset to only include genes that were differentially expressed between at least one of the three conditions (control, growth factor, growth factor plus inhibitor) assessed. To define the integration between the pathways, we modified a previously described approach for generating polar plots that has been used to identify dynamic trends in gene expression (47, 94). Relative mean signal changes over successive conditions, of control, growth factor, and growth factor plus inhibitor treatment (T^0 , T^1 , and T^2 , respectively), were transformed to an angle between -180 and $+180$ degrees (θ). This angle was further transformed to a scale from 0 to 360 degrees and then interpolated into a color based on the angle. Differentially expressed genes were plotted using polar coordinates (r, θ) with the radius (r), or distance from the focus, being equal to the value of largest absolute $\log_2$ fold change between conditions and theta (θ) being equal to the angle corresponding to the trend of change and calculated using the following equations:

$$I = \max(T^0, T^1, T^2)$$

$$R = 1 - \frac{\min(T^0, T^1, T^2)}{I}$$

$$\theta = \left(\frac{T^0 + T^1 - T^2 - I}{I \times R} + 2 \right) \times \text{sign}(T^1 - T^0) \times 60$$

The intensity (I) is equivalent to the maximum expression for a gene across conditions and the ratio (R) between the minimum expression of a gene across conditions over the intensity, are used to calculate the angle of expression trend (θ). As such, genes are plotted as a point with an angle (θ) corresponding to relative changes in a gene's expression with subsequent treatments at a distance from the focus corresponding to the largest magnitude of change between any pair of conditions in the treatment series. Three-point profile annotations of expression change (corresponding from left to right as T^0 , T^1 , and T^2 , respectively) are indicated on the outer edge of the polar scatter plots every 30 degrees in the color assigned to that particular θ . These profiles are accompanied by a radial histogram denoting the density of differentially changed genes across organoid conditions at particular θ of the polar scatterplot.

To summarize the meaning of genes with particular profiles in the polar scatterplots, genes were grouped by profiles at 30-degree intervals starting at 0 degrees, which denotes genes that were unchanged in expression from T^0 to T^1 but demonstrated a significant increase in expression following growth factor plus inhibitor treatment of the organoids (T^2). Genes from each group were assessed for enrichment of Gene Ontology biological processes (GP) using the EnrichGo function of the *clusterProfiler* package in R (95). Processes with $p < 0.01$ after BH correction are included in **Supplementary files 4 and 5**. Values were transformed ($-\log_{10}$) and displayed alongside polar scatterplots. Genes within each profile group of the polar scatterplot and utilized for GO enrichment were scaled and displayed as a line plot. A

dotted line was added to these linear plots to denote the mean scaled expression pattern across organoid treatment conditions for all genes within a particular interval.

Breast Cancer Tissue Microarray

The construction of the TNBC tissue microarray (TMA) containing duplicate cores from 50 patients was performed by the Histopathology and Tissue Shared Resource (HTSR) at Georgetown University. Details of the methodology, patient consent, de-identification and demographics have been previously described (96). Multiplex staining of deparaffinized and rehydrated cores was performed through sequential rounds of staining, stripping and re-probing with antibodies recognizing KRT14, Vimentin and E-cadherin and OPAL fluorescent secondary antibodies (**Supplementary file 1**). The TMA cores were imaged with a Vectra3 Multi-Spectral Imaging microscope (Perkin Elmer) and analyzed with inForm software (Perkin Elmer). Tumor and stroma were segmented based on E-cadherin expression (tumor). Cores with folded sections or no clear tumor-stomal interface by manual inspection were excluded from analysis. Cores were classified “high invasion” if invasive strands were detected in >50% of the tumorstomal interface by manual inspection. The KRT14 and Vimentin signals were not visible when classifying invasion status.

Human Breast Cancer Gene Expression Analysis

The mRNA expression data for 996 patients in the Breast Invasive Carcinoma (The Cancer Genome Atlas (TCGA), PanCancer Atlas) was analyzed and downloaded from the cBioPortal (97, 98). Tumor subtype specific gene expression values are Z-scores relative to all samples. Correlative expression values are Log2 transformed. The mRNA expression for human breast cancer cells was derived from a prior analyses (21, 31) and available at the GEO (GSE58643 and GSE62569).

Statistical Methods

Data was analyzed with Prism 9.0.2 (Graphpad) and *DESeq2*, *GSEA* and *ClusterProfiler* packages in R. Data with a normal distribution, as determined by Shapiro-Wilk test, was analyzed by two tailed Student’s t-test. Data that did not pass a normality test were analyzed by Mann-Whitney U test or Kruskal Wallis test with Dunn’s Multiple comparison test. Differentially expressed genes with $p < 0.05$ after Benjamini-Hochberg correction were identified using the *DESeq2* package in R. Fisher’s exact tests to determine the enrichment of Gene Ontology biological processes (BP) were performed using the *EnrichGo* function of the *ClusterProfiler* package in R. The specific statistical tests are indicated in the figure legends. Sample numbers are defined and indicated in the figure legends or on the figure panels. All experiments include a minimum of 2 biological replicates from separate biological preparations.

Data Availability

The RNA sequencing data generated in this study are publicly available at the GEO at GSE188300, GSE188312, GSE188323, GSE58643 and GSE62569. All other data are available within the article and its supplementary data files.

Materials Availability

Cell and organoid lines derived from C3-TAg tumors will be made available by the corresponding author upon reasonable request.

Disclosure of potential conflicts of interest

There are no potential conflicts of interest.

Authors contributions

A. Nasir: Conceptualization, investigation, writing-review and editing. **S. Camacho:** Conceptualization, investigation, writing-review and editing. **A.T. McIntosh:** Investigation, writing-review and editing. **G. Graham:** Investigation, writing-review and editing. **R. Rahhal:** Investigation, writing-review and editing. **M.E. Huysman:** Investigation. **F. Alsharief:** Investigation. **A.T. Riegel:** Resources, funding acquisition. **G.W. Pearson:** Conceptualization, investigation, writing-review and editing, funding acquisition.

Acknowledgements

Work was supported by NIH R01CA218670, Georgetown Women and Wine (to G.W. Pearson), NIH T32CA009686 (to S. Camacho, G. Graham) and NIH P30CA051008.

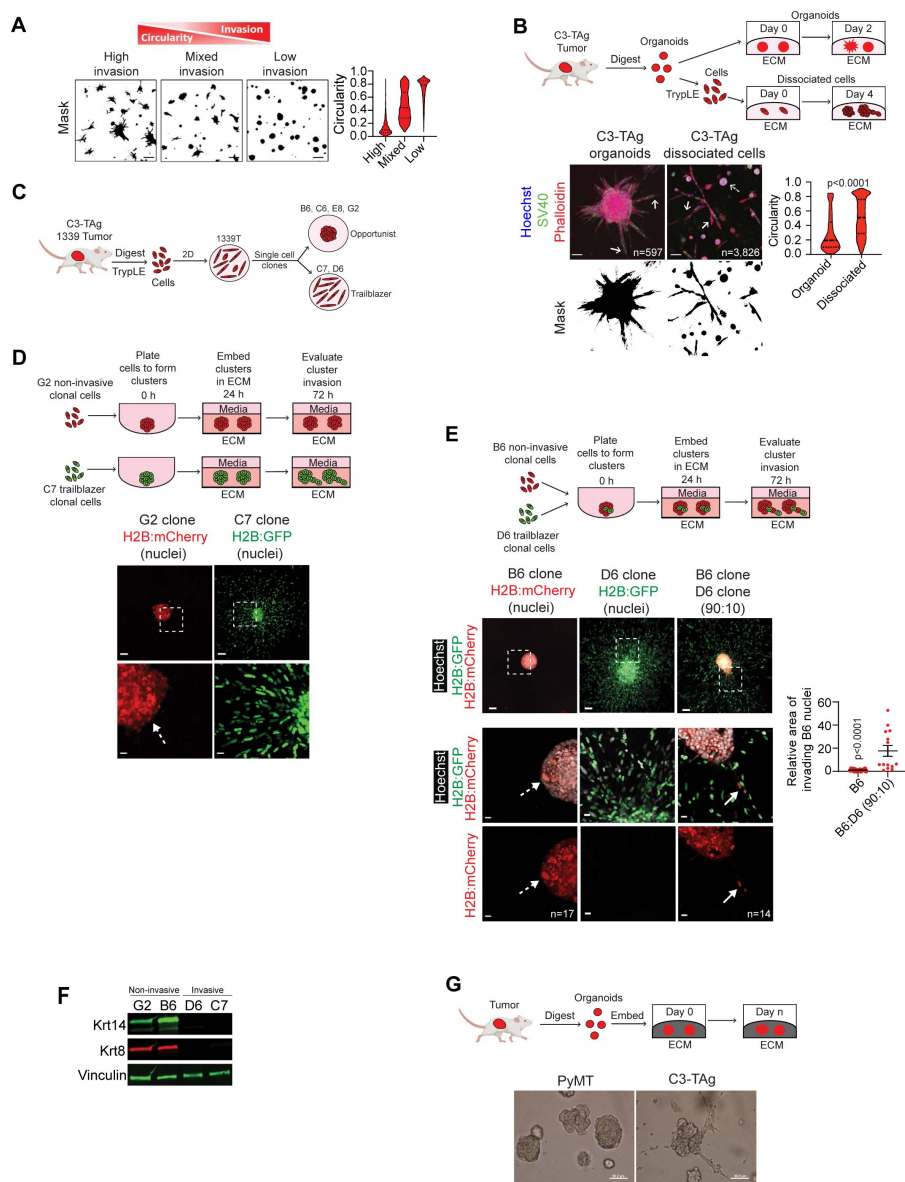


Figure 1—figure supplement 1.

A. Representative masked images of C3-TAG organoids showing reduced organoid invasion which is represented by increased organoid circularity. Scale bars, 50 μ m. **B.** C3-TAG organoids and dissociated cells from the same tumors were grown in ECM for 2 and 4 days respectively to evaluate invasion. Solid arrows indicate representative invasive cells. Dashed arrow indicates a representative non-invasive spheroid. Violin plot shows the quantification of circularity (Mann Whitney test, n =organoids, dissociated cell spheroids or individual dissociated cells, r =2). Scale bars, 50 μ m. **C.** Graphic shows the methodology for deriving opportunist and trailblazer clonal cell lines. **D.** Representative images of non-invasive (G2) and invasive (C7) clonal cell line clusters. Scale bars, 100 μ m (top), 20 μ m (bottom, inset). **E.** Trailblazer D6-H2B:GFP cells promoted the collective invasion of non-invasive opportunist B6-H2B:mCherry cells. Top row, 2 $\times$ 2 tiled LED fluorescence images. Inset images are confocal sections of the ROIs indicated by the dashed line boxes. Graph shows the quantification of B6-H2B:mCherry cell invasion (mean $\pm$ SEM, Mann Whitney test,

n =clusters, r =2). Scale bars, top row 100 μ m; inset images 20 μ m. **F.** Immunoblots showing the expression of Krt14 and Krt8 in non-invasive (G2, B6) and invasive (D6, C7) C3-TAG clonal cell lines. Vinculin serves as a loading control. (n =2) **G.** Representative images from 2 independent experiments show that C3-TAG organoids are intrinsically more invasive than PyMT organoids in ECM composed of low Collagen I.

Figure 1—Video 1. Live imaging of C3-TAG derived tumor organoids in organotypic culture (GFP, green, tumor cell). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with invasive heterogeneity in C3-TAG organoids shown in [Figure 1A](#) and [Figure 1—figure supplement 1B](#).

Figure 1—Video 2. Live imaging of C3-TAG derived tumor dissociated organoids in organotypic culture. Bright field images were acquired for 5 h at 30 min intervals. The video corresponds with the intratumor invasive heterogeneity in C3-TAG tumor dissociated organoids shown in [Figure 1—figure supplement 1B](#).

Figure 1—Video 3. Live imaging of homogenous clusters of C3-TAg tumor derived noninvasive clonal cells B6 (H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with non-invasive B6 clusters shown in [Figure 1B](#) and [Figure 1—figure supplement 1E](#).

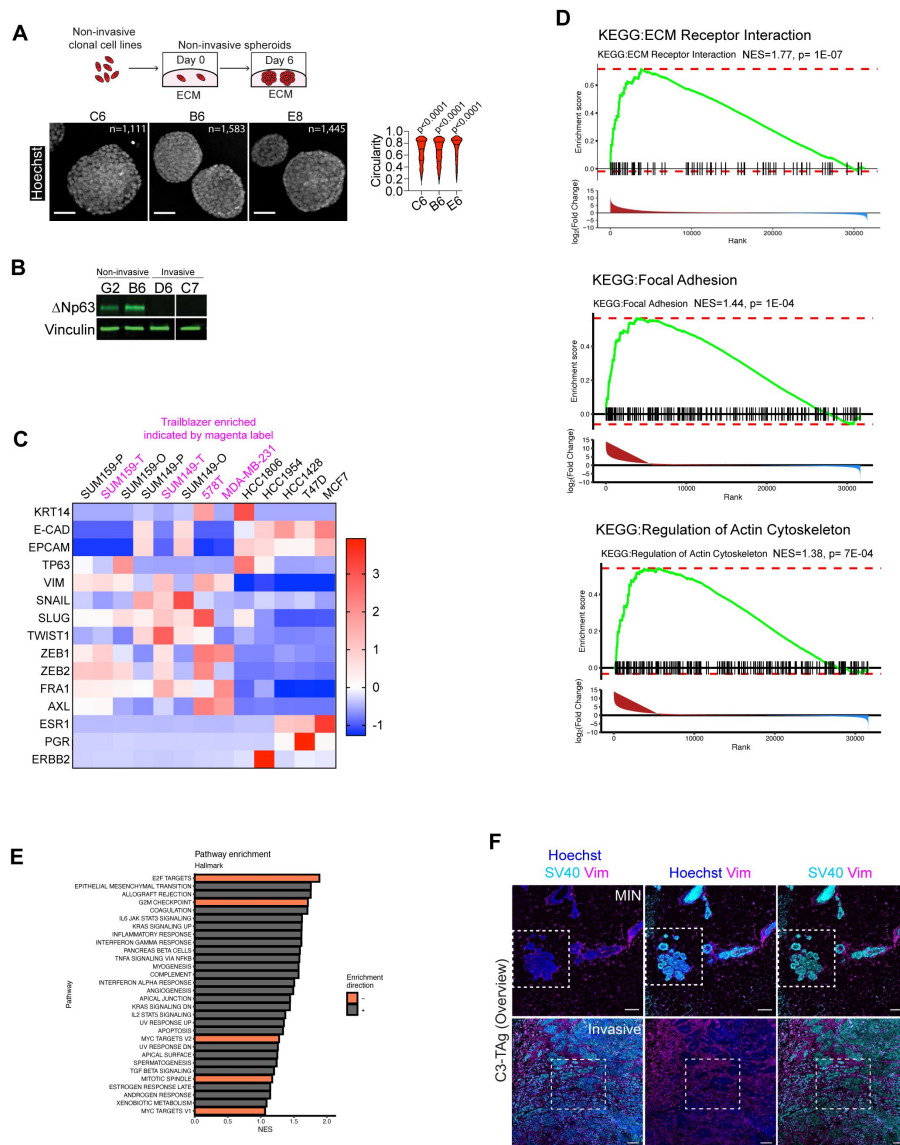
Figure 1—Video 4. Live imaging of homogenous clusters of C3-TAg derived tumor cells. Bright field images were acquired for 14 h at 30 min intervals. The video corresponds with invasive C3-TAg tumor cells clusters shown in [Figure 1B](#).

Figure 1—Video 5. Live imaging of heterogeneous clusters of C3-TAg derived tumor cells and non-invasive clonal cells B6 (H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with opportunistic invasion of B6 cells induced by C3-TAg tumor cells as shown in [Figure 1B](#).

Figure 1—Video 6. Live imaging of heterogeneous clusters of C3-TAg derived invasive (D6) and non-invasive (B6) clonal cells (H2B: GFP, green, nuclei; H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with opportunistic invasion of B6 cells induced by D6 cells as shown in [Figure 1—figure supplement 1E](#).

Figure 1—Video 7. Live imaging of PyMT derived tumor organoids in organotypic culture. Bright field images were acquired for 18 h at 30 min intervals. The video corresponds with less invasive PyMT organoids as shown in [Figure 1F](#).

Figure 1—Video 8. Live imaging of C3-TAg derived tumor organoids in organotypic culture. Bright field images were acquired for 18 h at 30 min intervals. The video corresponds with more highly invasive C3-TAg organoids as shown in [Figure 1F](#).



the ROIs from Figure 2G. Scale bars, 100 μ m.

Figure 2—figure supplement 1.

A. Non-invasive opportunist C3-TAg clonal cell lines were grown in ECM for at least 6 days. Violin plot shows the quantification of spheroid circularity (Mann Whitney test, n=spheroids, r=2). Scale bars, 50 μ m. **B.** Immunoblots showing the expression of Δ Np63 in non-invasive (opponentist) (G2, B6) and invasive (trailblazer) (D6, C7) C3-TAg clonal cell lines (n=2). Lysates are from the same gel. **C.** Heatmap showing the expression of the indicated genes in human breast cancer cell lines. Trailblazer enriched populations are indicated in magenta. P=parental, T=trailblazer and O=opponentist when parental cell lines and their daughter trailblazer and opponentist subpopulations are shown (SUM159 and SUM149). **D.** GSEA was performed on RNA-seq data from C3-TAg organoids and non-invasive C3-TAg clones spheroids to identify pathway enrichment in C3-TAg organoids. Plots show the GSEA for the indicated KEGG pathways. **E.** Graph shows Hallmark pathways with NES ≥ 1 . **F.** Overview images of noninvasive carcinoma in situ (CIS) and invasive carcinoma cells in C3-TAg tumors. Dashed boxes showing

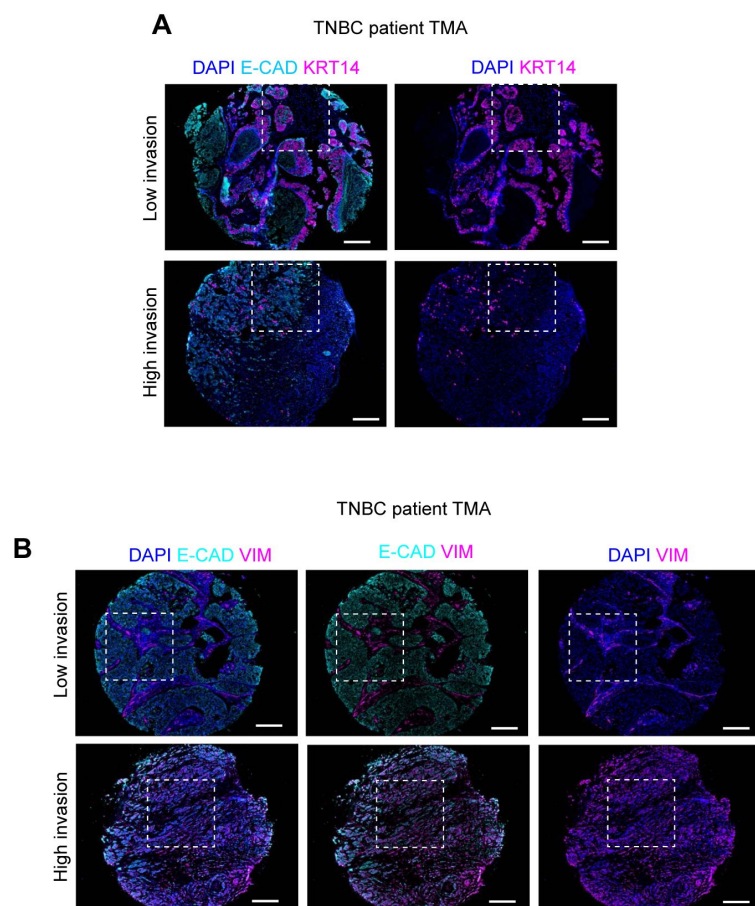


Figure 3—figure supplement 1

A. Full TNBC patient tumors images with dashed boxes showing the ROIs from [Figure 3B](#). Scale bars, 200 μ m. **B.** Full TNBC patient tumor images with dashed boxes showing the ROIs from [Figure 3D](#). Scale bars, 200 μ m.

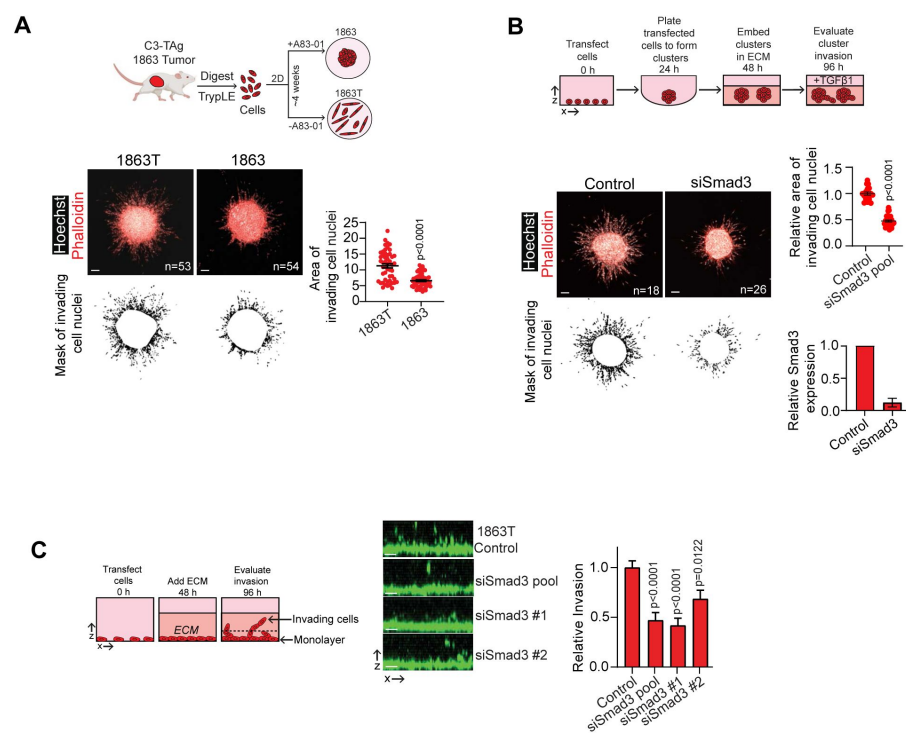


Figure 4—figure supplement 1.

A. 1863 cells derived in the presence of the Tgfr1 inhibitor A83-01 are less invasive than 1863T cells, which were derived from the same tumor but without A83-01. Graph shows 1863 and 1863T cluster invasion (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μ m. **B.** Smad3 siRNA transfection suppresses TGF β 1 treated 1863 cluster invasion. Dot plot shows the quantification of invasion (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μ m. Bar graph shows Smad3 expression (mean $\pm$ SEM, n=2). **C.** Pooled and individual Smad3 siRNAs suppress 1863T vertical invasion (mean $\pm$ SEM, unpaired Student's t test, n=3). Scale bars, 50 μ m.

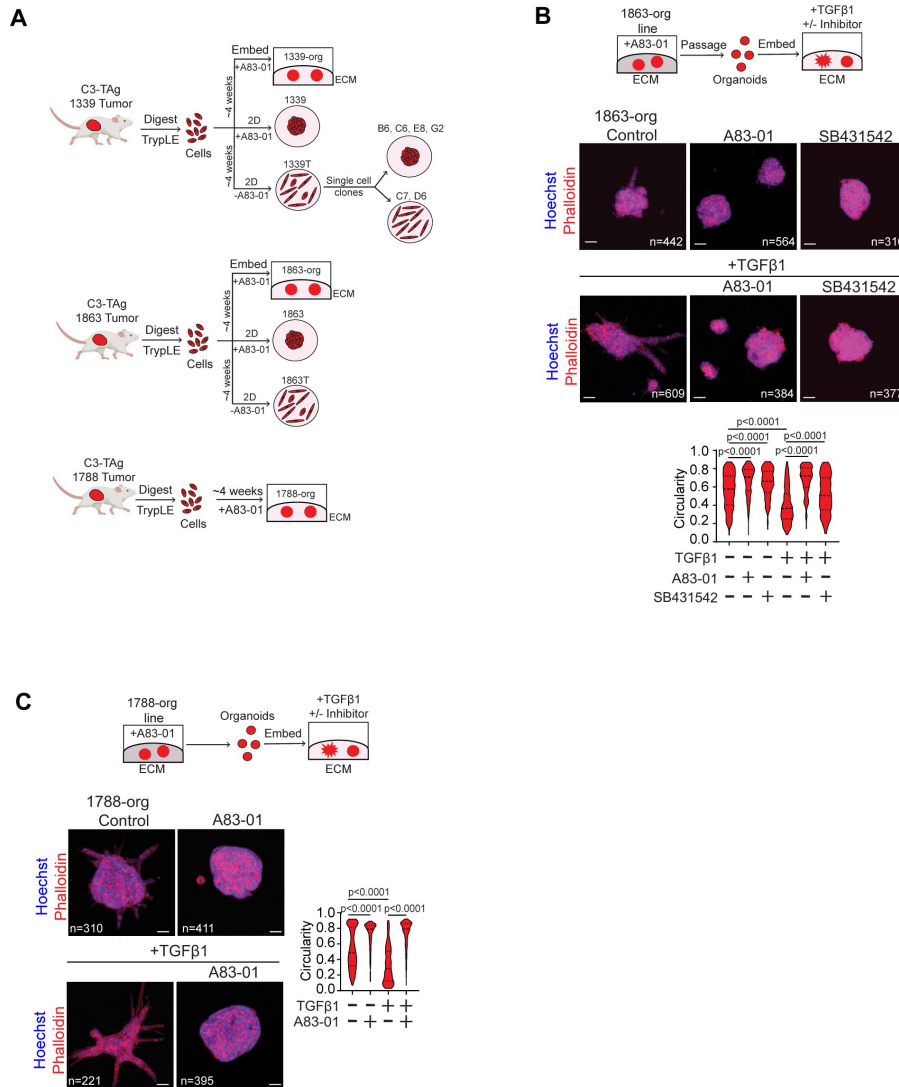


Figure 4—figure supplement 2.

A. Graphics show the process of deriving 1339-org, 1863-org and 1788-org lines and their relationship to 2D cell lines derived from the same tumors. **B-C.** Tgfr1 inhibition suppresses autocrine and exogenous TGFβ1 induced invasion of C3-TAG **(B)** 1863-org and **(C)** 1788-org lines. Organoids were treated with the Tgfr1 inhibitors A83-01 (500 nM), SB431542 (1 μM) or exogenous TGFβ1 (2 ng/ml) as indicated. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=2). Scale bars, 50 μm.

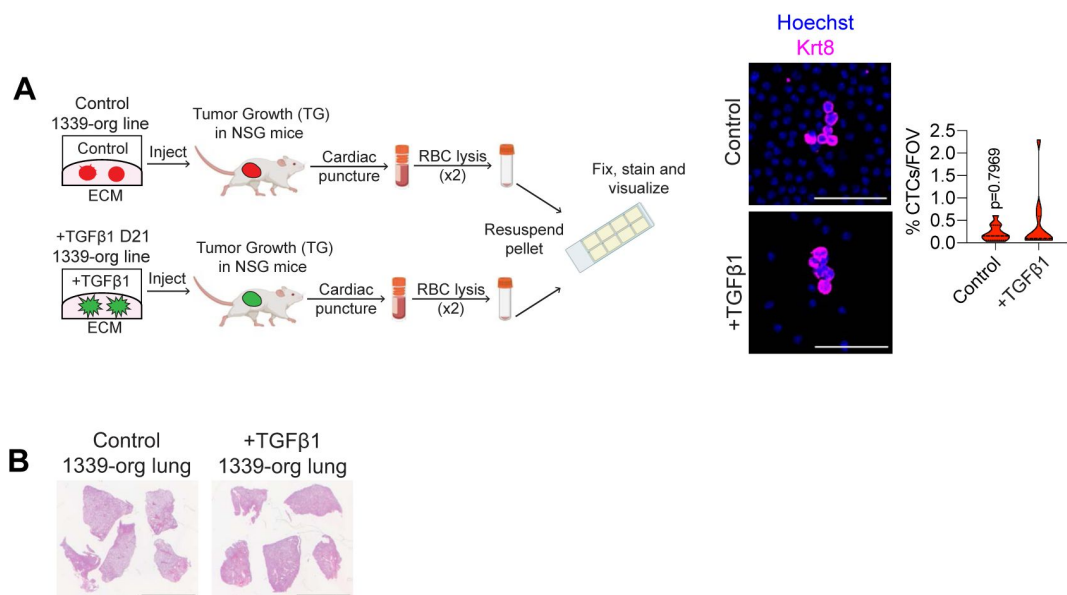


Figure 4—figure supplement 3.

A. Graphic showing the process of CTC isolation from mice. Representative images show Krt8 expressing CTCs detected in mice with control 1339-org tumors and +TGFβ1 1339-org tumors. Violin plot shows percentage of total cells that were CTCs per field of view (Mann Whitney test, representative of blood collected from 7 mice per group, $r=2$). Scale bars, 100 μm; 50 μm (inset).

Figure 4—Video 1. Live imaging of C3-TAg derived 1339-org line organoids in serum starved media (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with the baseline invasion of C3-TAg derived 1339-org line organoids as shown in [Fig. 3C](#), [S3C](#).

Figure 4—Video 2. Live imaging of C3-TAg derived 1339-org line organoids in serum starved media + 500 nM A830-01 (Tgfr1 inhibitor) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with suppression of 1339-org line organoid invasion by inhibiting autocrine TGFβR1 activity as shown in [Fig. S3C](#).

Figure 4—Video 3. Live imaging of C3-TAg derived 1339-org line organoids in presence of exogenous TGFβ1 (48 h) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with TGFβ1 induced invasion in 1339-org line organoids as shown in [Fig. 3C](#), [S3C](#).

Figure 4—Video 4. Live imaging of C3-TAg derived 1339-org line organoids in presence of persistent exogenous TGFβ1 (Day 16) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with progressively increased TGFβ1 induced invasion in 1339-org line organoids as shown in [Fig. 3C](#).

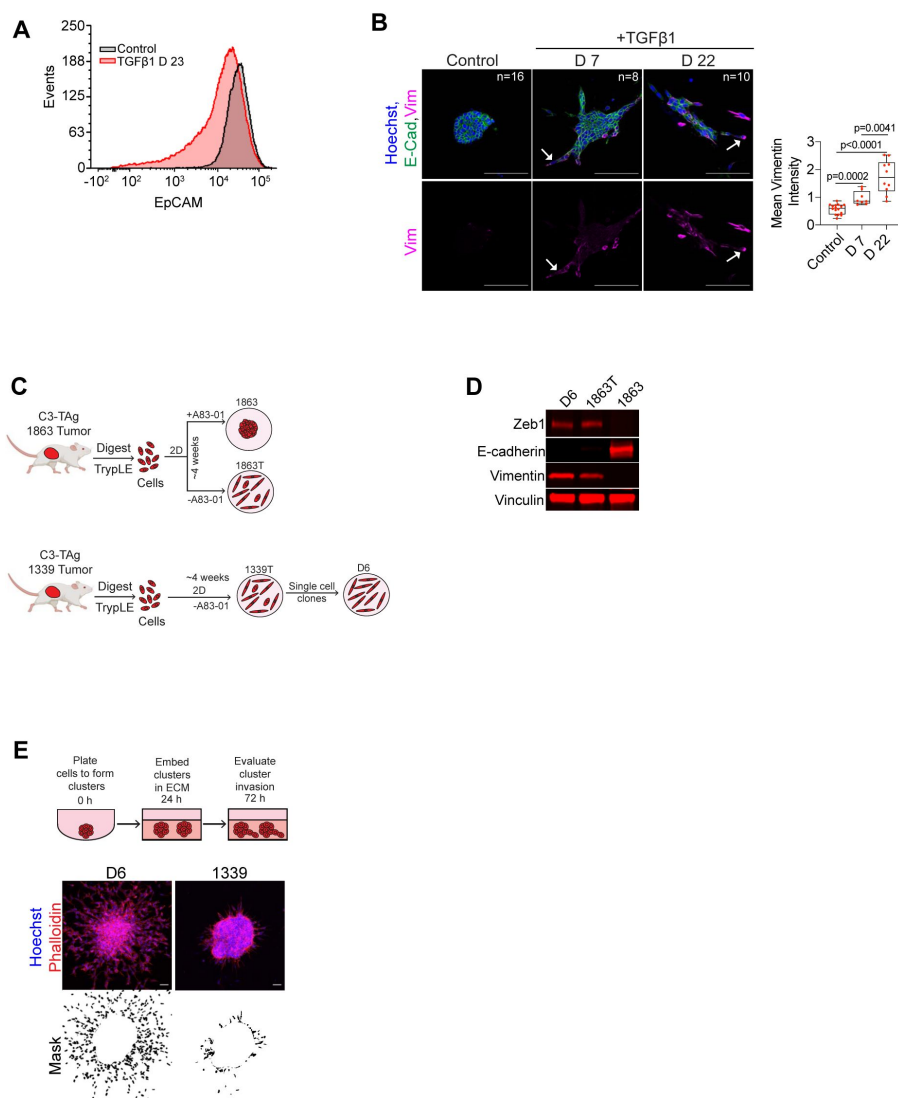


Figure 5—figure supplement 1.

A. FACS showing that TGFβ1 treated 1339-org cells undergo a shift in the distribution of EpCAM expression (representative of 2 independent experiments). **B.** Images showing a progressive increase of Vimentin expression in TGFβ1 treated 1339-org cells. Arrows indicate representative Vimentin expressing cells leading collective invasion. Graph shows Vimentin expression in 1339-org organoids (Student's t-test, n=organoids). Scale bars, 100 μm. **C.** Immunoblots showing the expression of Zeb1, E-cadherin and Vimentin in D6, 1863T and 1863 cell lines. Vinculin serves as a loading control (n=2). **D.** Immunoblots showing the expression of E-cadherin and Vimentin in non-invasive B6 and trailblazer D6 cell lines. Vinculin serves as a loading control. (n=2). **E.** Representative images of D6 and 1339 clusters. Scale bars 50 μm.

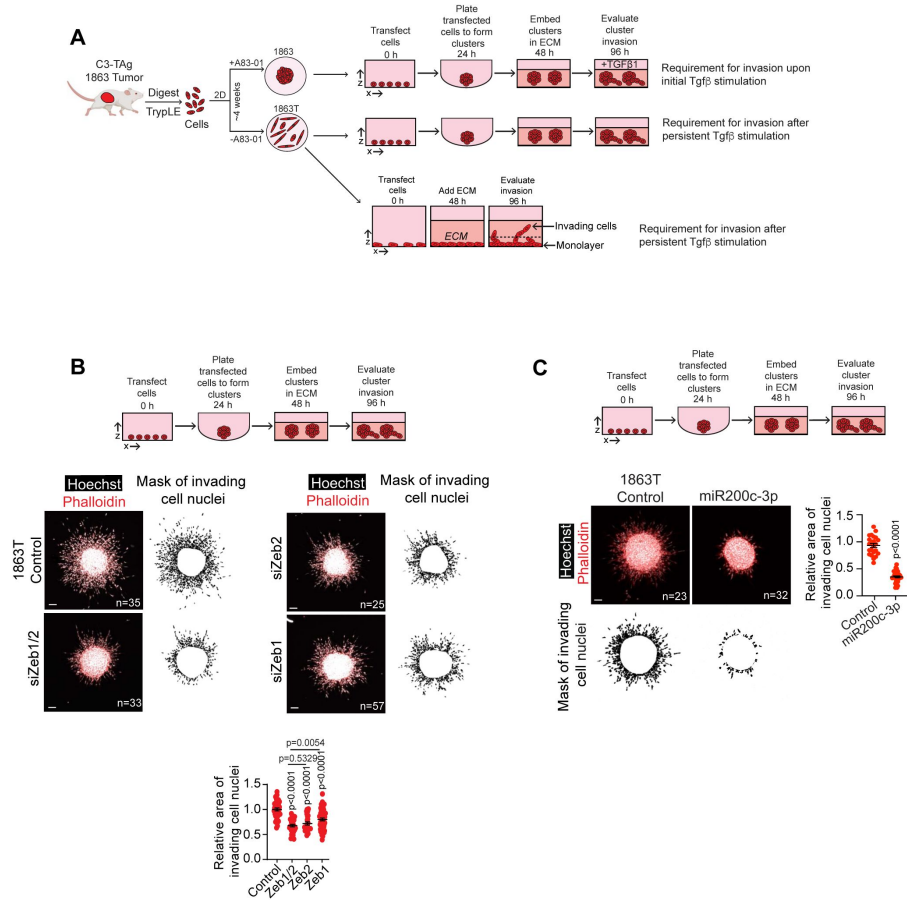


Figure 6—figure supplement 1.

A. Zeb1 and Zeb2 siRNA transfection suppresses 1863T cluster invasion (mean±SEM, Mann Whitney test, n=clusters, r=3). Scale bars, 50 μm. **B.** Transfection with a miR200c-3p mimic decreases 1863T cluster invasion (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μm.

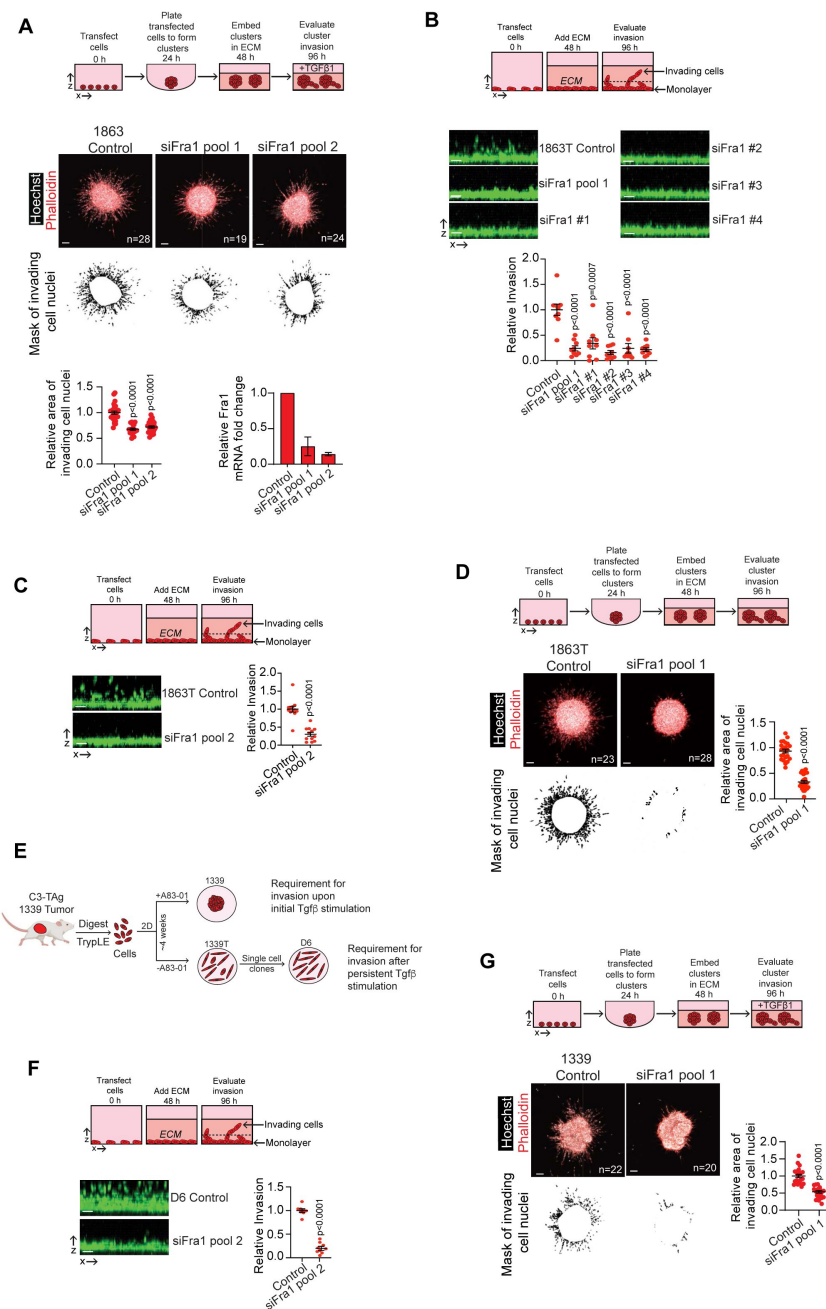


Figure 6—figure supplement 2.

A. Fra1 depletion by a second distinct Fra1 siRNA pool (pool 2) decreases TGF β 1 treated 1863T cluster invasion. Dot plot shows the quantification of invasion (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2). Bar graph shows the depletion of Fra1 expression by two different pools of Fra1 siRNA (mean $\pm$ SEM, n=3 for pool 1 and n=2 for pool 2). Scale bars, 50 μ m. **B.** 1863T vertical invasion is suppressed by each of the four individual sequences from Fra1 siRNA pool 1 (mean $\pm$ SEM, Mann Whitney test, n=9 wells, r=3). Scale bars, 50 μ m. **C.** Fra1 siRNA pool 2 reduces the vertical invasion of 1863T cells (mean $\pm$ SEM, Mann Whitney test, n=12 wells, r=4). Scale bars, 50 μ m. **D.** Fra1 depletion suppresses the invasion of 1863T clusters (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μ m. **E.** Fra1 depletion reduces the vertical invasion of D6 cells (mean $\pm$ SEM, Mann Whitney test, n=9 wells, r=3). Scale bars, 50 μ m. **F.** Fra1 depletion reduces the invasion of TGF β 1 treated 1339 clusters (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μ m.

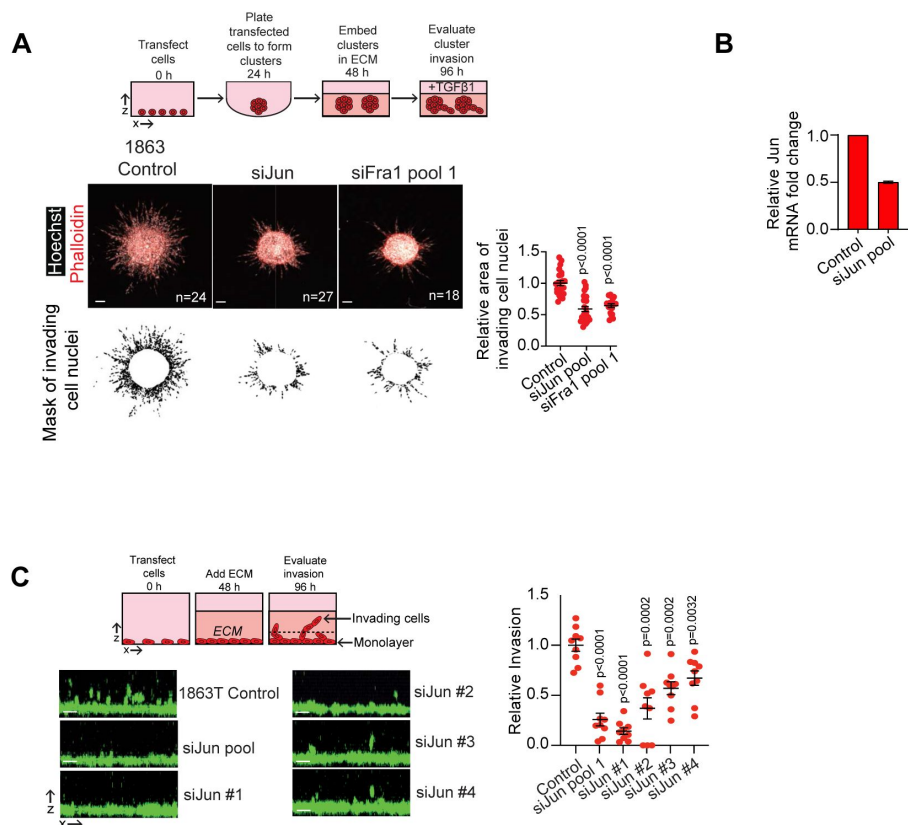


Figure 6—figure supplement 3.

A. Jun depletion suppresses TGF β 1 induced 1863 cluster invasion (mean $\pm$ SEM, Mann Whitney test, n=clusters, r=2) Scale bars, 50 μ m. **B.** Depletion of Jun expression by Jun siRNA transfection (mean $\pm$ SEM, n=2). **C.** Pooled and individual Jun siRNAs suppress 1863T vertical invasion (mean $\pm$ SEM, unpaired Student's t test, n=3). Scale bars, 50 μ m.

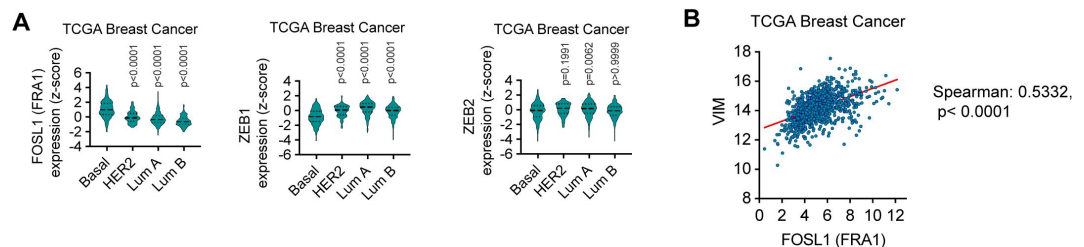


Figure 6—figure supplement 4.

A. FOSL1 (FRA1) mRNA is more highly expressed in the basal-like breast cancer patient tumors compared to other breast cancer subtypes. By comparison, ZEB1 and ZEB2 mRNA is more highly expressed in other breast cancer cancer subtypes compared to basal-like breast cancer. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to basal-like tumors **B.** Correlation of VIM and FOSL1 (FRA1) expression in breast cancer patient tumors.

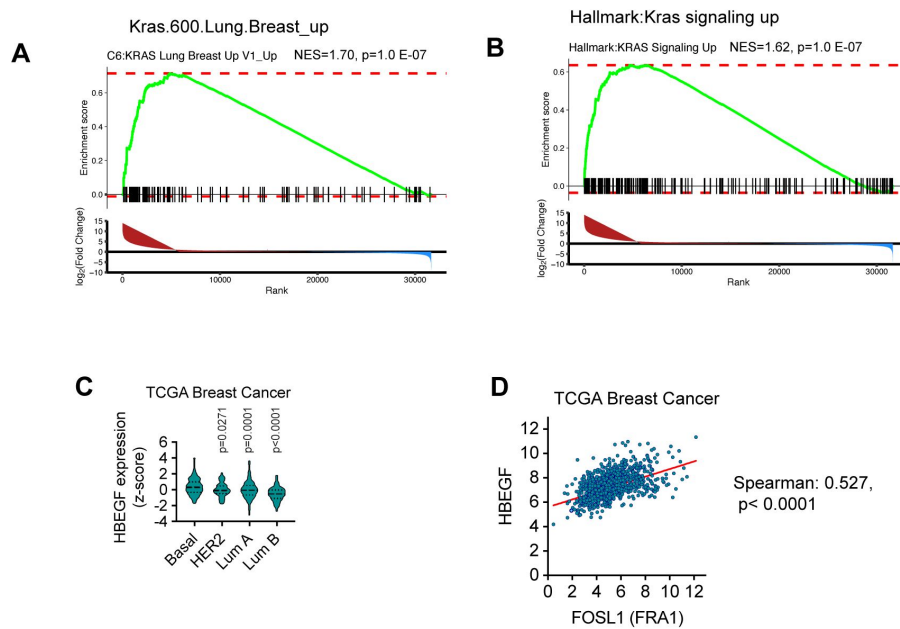


Figure 7—figure supplement 1.

A-B. GSEA was performed on RNA-seq data from C3-TAg organoids and non-invasive C3-TAg clone spheroids to identify pathway enrichment in C3-TAg organoids. Plots show the GSEA associated with Kras signaling. **C.** HBEGF mRNA is more highly expressed in the basal-like breast cancer patient tumors compared to other breast cancer subtypes. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to the basal-like tumors. **D.** Correlation of FOSL1 (FRA1) and HBEGF expression in breast cancer patient tumors.

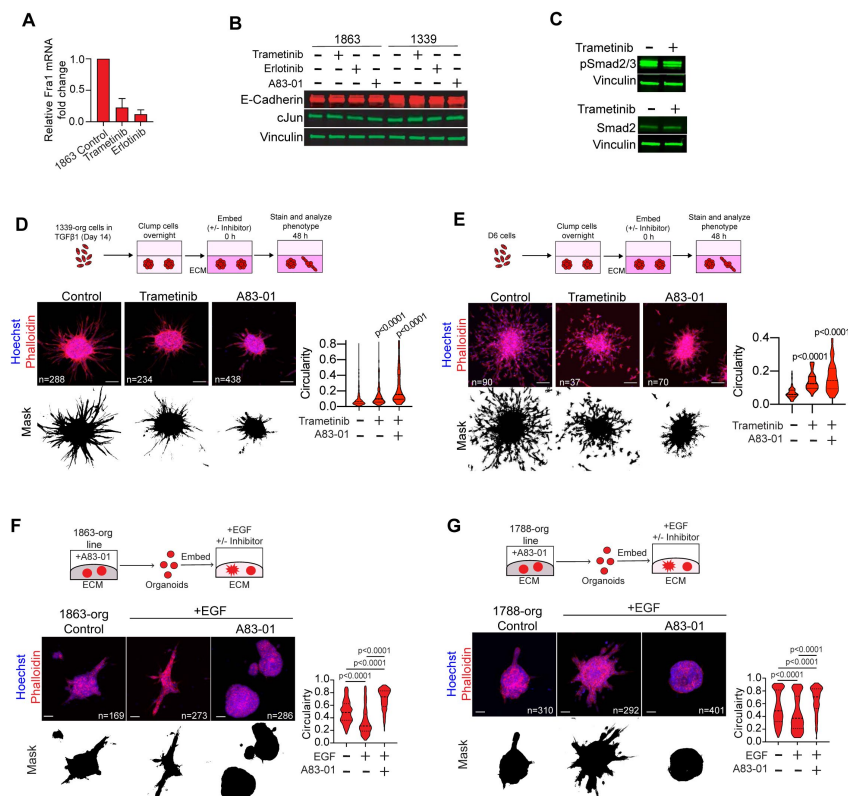


Figure 7—figure supplement 2.

A. Trametinib (10 nM) and erlotinib (1 μM) reduce Fra1 expression in 1863 cells (mean±SEM, n=2). **B.** Immunoblot showing that Jun expression is not perturbed by treatment with trametinib (10 nM), erlotinib (1 μM) or A83-01 (500 nM) in 1863 and 1339 cells (n=2 for erlotinib, A8301; n=3 for trametinib). **C.** Immunoblot shows that trametinib (10 nM) does not reduce Zeb1 expression or Smad2 phosphorylation in C3-TAg cells (n=2). **D.** 1339-org organoids were persistently stimulated with exogenous TGFβ1 for 14 days to induce a trailblazer state (D14 TGFβ1 1339-org). The ability of trametinib and A83-01 to suppress then invasion of the D14 TGFβ1 1339-org cells was then determined. Violin plots show the quantification of invasion (Mann Whitney test, n=organoids, r=2). Scale bars, 100 μm. **E.** Trametinib and A83-01 suppress the invasion of D6 trailblazer cells (Mann Whitney test, n=spheroids, r=2).

F-G. The Tgfr1 inhibitor A83-01 suppresses the EGF induced invasion of **(F)** 1863-org and **(G)** 1788-org organoids. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=2). Scale bars, 50 μm.

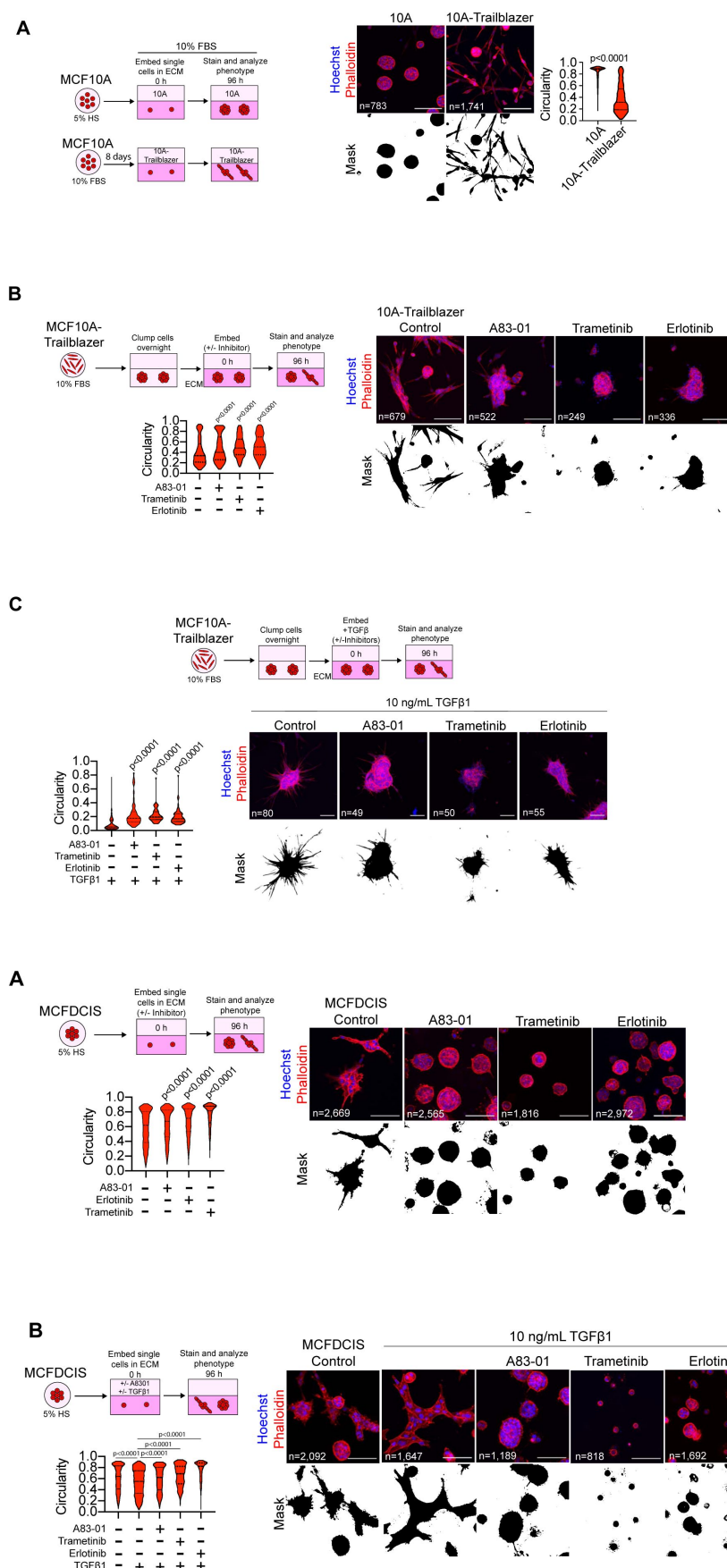


Figure 7—figure supplement 3.

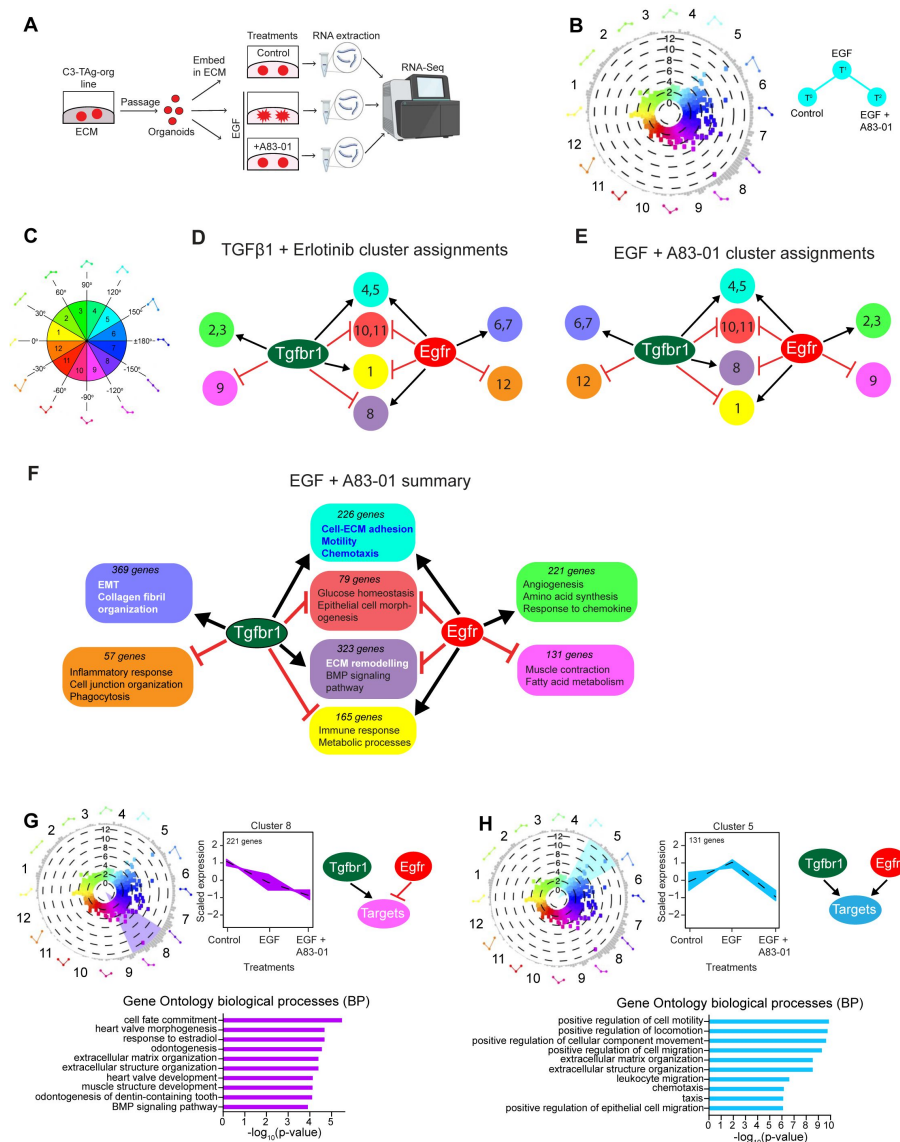
A. Graphic and data shows how a MCF10A-trailblazer cell line was established by treating MCF10A cells with 10% FBS. MCF10A cells were supplemented with 10% FBS for 8 days in monolayer culture. The invasive ability of control and 10% FBS treated MCF10A cells was then determined. The 10% FBS induced a trailblazer state in the MCF10A cells, which we term “MCF10A-Traiblazer”. Violin plot shows quantification of circularity. (Mann Whitney test, n =spheroids, r =2). Scale bars, 100 μ m. **B.** A83-01, trametinib and erlotinib suppresses MCF10A-trailblazer invasion (Mann Whitney test, n =spheroids, r =2). Scale bars, 100 μ m. **C.** A83-01, trametinib and erlotinib suppress the invasion of MCF10A-trailblazer cells treated with exogenous TGF β 1 (Mann Whitney test, n =spheroids, r =2). Scale bars, 100 μ m.

Figure 7—figure supplement 4.

A. A83-01, trametinib and erlotinib suppress MCFDCIS spheroid invasion (Mann Whitney test, n =spheroids, r =2). Scale bars, 100 μ m. **B.** A83-01, trametinib and erlotinib suppress the TGF β 1 induced invasion of MCFDCIS spheroids (Mann Whitney test, n =spheroids, r =2). Scale bars, 100 μ m.

Figure 8—figure supplement 1.

A. Overview of the approach to test the interactions between the Egfr and Tgf β signaling pathways when cells were treated with EGF or EGF+A83-01. **B.** Polar plot organizing genes represented as individual dots based on their change in expression in response to treatment with EGF or co-treatment with EGF and A83-01. Genes that underwent ≥ 2 fold-change with a $p < 0.05$ after Benjamini-Hochberg correction in at least one condition relative to one of the other conditions are shown (3 total comparisons). The distance from the center of the plot represents the $\Delta\log_2$ value of each gene (increased difference in expression, further from the center). The bar plots on the surface of the circle show the number of genes at the degree point in the plot. Genes are grouped into 12 color-coded groups along 30 degree increments over the 360 degree plot. As an example, genes in group 5 were increased in expression by EGF and this increase in expression was inhibited by A83-01. In contrast, genes in cluster 1 were increased in expression by EGF, and this induction is further enhanced by treatment with A83-01. **C.** Plot shows the cluster



locations along the 30 degree increments in the polar plots shown in **Figure 8B** and **Figure 8—figure supplement 1B**. **D.** Clusters assigned to different summary groups for TGF β 1 and TGF β 1+erlotinib treated organoids shown in **Figure 8B**. **E.** Clusters assigned to different summary groups for EGF and EGF+A83-01 treated organoids shown in **Figure 8—figure supplement 1B**. **F.** Summary of the number of genes and represented biological processes associated with different modes of interaction between the Tgfr1 and Egfr pathways in cells treated with EGF and EGF+A83-01. Also see **Supplementary file 5** for the assignment of polar plot clusters and detailed biological processes used to define the summary integration groups. **G.** Egfr restricts the expression of a subset of genes induced by autocrine Tgf β . Line plots show the expression of genes in Cluster 1. Dashed line indicates the average scaled expression of all Cluster 1 genes. Graph shows the top 10 biological processes (GO-BP) associated with Cluster 1 genes (Fisher's exact test). **H.** Tgfr1 activity was necessary for the expression of a subset of genes induced by extrinsic EGF. Line plots show the expression of genes in Cluster 5. Dashed line indicates the average scaled expression of all Cluster 5 genes. Graph shows the top 10 biological processes (GO-BP) associated with Cluster 5 genes (Fisher's exact test).

Supplementary file 1. Description of GEM models, organoid/cell line models and reagents used in the study.

Supplementary file 2. RNA-Seq data (log2(CPM) genes x samples) from comparison between non-invasive clones and C3-TAG tumor organoids. Pathway enrichment analysis from comparison between non-invasive clones and C3-TAG tumor organoids

Supplementary file 3. Gene clusters and enriched biological processes from comparison between control and TGFβ1 treated C3-TAG 1339-org line organoids

Supplementary file 4. Gene clusters and enriched biological processes from comparison between control (T⁰), TGFβ1 (T¹) and TGFβ1 + Erlotinib (T²) treated C3-TAG 1339-org line organoids.

Supplementary file 5. Gene clusters and enriched biological processes from comparison between control (T⁰), EGF (T¹) and EGF + A8301 (T²) treated C3-TAG 1339-org line organoids.

Source Data Legends

Full annotated and raw immunoblots corresponding to [Figure 1—figure supplement 1F](#), [Figure 2—figure supplement 1B](#), [Figure 5—figure supplement 1D](#), [Figure 7C](#) and [Figure 7—figure supplement 2B-C](#).

References

1. Marusyk A , Polyak K. (2010) **Tumor heterogeneity: causes and consequences** *Biochim Biophys Acta* **1805**:105–17
<https://doi.org/10.1016/j.bbcan.2009.11.002>
2. Beca F , Polyak K (2016) **Intratumor Heterogeneity in Breast Cancer** *Adv Exp Med Biol* **882**:169–89
https://doi.org/10.1007/978-3-319-22909-6_7
3. McGranahan N , Swanton C (2015) **Biological and therapeutic impact of intratumor heterogeneity in cancer evolution** *Cancer Cell* **27**:15–26
<https://doi.org/10.1016/j.ccell.2014.12.001>
4. Marusyk A , Janiszewska M , Polyak K (2020) **Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance** *Cancer Cell* **37**:471–84
<https://doi.org/10.1016/j.ccell.2020.03.007>
5. Ganesh K , Massagué J (2021) **Targeting metastatic cancer** *Nat Med* **27**:34–44
<https://doi.org/10.1038/s41591-020-01195-4>
6. Calbo J , van Montfort E , Proost N , van Drunen E , Beverloo HB , Meuwissen R , Berns A (2011) **A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer** *Cancer Cell* **19**:244–56
<https://doi.org/10.1016/j.ccr.2010.12.021>

7. Dongre A , Rashidian M , Reinhardt F , Bagnato A , Keckesova Z , Ploegh HL , Weinberg RA (2017) **Epithelial-to-Mesenchymal Transition Contributes to Immunosuppression in Breast Carcinomas** *Cancer Res* **77**:3982–9
<https://doi.org/10.1158/0008-5472.Can-16-3292>
8. Celià-Terrassa T , Meca-Cortés O , Mateo F , de Paz A Martínez , Rubio N , Arnal-Estapé A , Ell BJ , Bermudo R , Díaz A , Guerra-Rebollo M , Lozano JJ , Estarás C , Ulloa C , Álvarez-Simón D , Milà J , Vilella R , Paciucci R , Martínez-Balbás M , de Herreros AG , Gomis RR , Kang Y , Blanco J , Fernández PL , Thomson TM (2012) **Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells** *J Clin Invest* **122**:1849–68
<https://doi.org/10.1172/jci59218>
9. Marusyk A , Tabassum DP , Altrock PM , Almendro V , Michor F , Polyak K (2014) **Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity** *Nature* **514**:54–8
<https://doi.org/10.1038/nature13556>
10. Zhang M , Tsimelzon A , Chang CH , Fan C , Wolff A , Perou CM , Hilsenbeck SG , Rosen JM (2015) **Intratumoral heterogeneity in a Trp53-null mouse model of human breast cancer** *Cancer Discov* **5**:520–33
<https://doi.org/10.1158/2159-8290.CD-14-1101>
11. Neelakantan D , Zhou H , Oliphant MUJ , Zhang X , Simon LM , Henke DM , Shaw CA , Wu MF , Hilsenbeck SG , White LD , Lewis MT , Ford HL (2017) **EMT cells increase breast cancer metastasis via paracrine GLI activation in neighbouring tumour cells** *Nat Commun* **8**
<https://doi.org/10.1038/ncomms15773>
12. Dongre A , Rashidian M , Eaton EN , Reinhardt F , Thiru P , Zagorulya M , Nepal S , Banaz T , Martner A , Spranger S , Weinberg RA (2021) **Direct and Indirect Regulators of Epithelial-Mesenchymal Transition-Mediated Immunosuppression in Breast Carcinomas** *Cancer Discov* **11**:1286–305
<https://doi.org/10.1158/2159-8290.Cd-20-0603>
13. Sharif GM , Campbell MJ , Nasir A , Sengupta S , Graham GT , Kushner MH , Kietzman WB , Schmidt MO , Pearson GW , Loudig O , Fineberg S , Wellstein A , Riegel AT (2021) **An AIB1 Isoform Alters Enhancer Access and Enables Progression of Early-Stage Triple-Negative Breast Cancer** *Cancer Res* **81**:4230–41
<https://doi.org/10.1158/0008-5472.CAN-20-3625>
14. Zhou H , Neelakantan D , Ford HL (2017) **Clonal cooperativity in heterogenous cancers** *Semin Cell Dev Biol* **64**:79–89
<https://doi.org/10.1016/j.semcdb.2016.08.028>
15. Friedl P , Gilmour D (2009) **Collective cell migration in morphogenesis, regeneration and cancer** *Nat Rev Mol Cell Biol* **10**:445–57
<https://doi.org/10.1038/nrm2720>

16. Nguyen-Ngoc KV , Cheung KJ , Brenot A , Shamir ER , Gray RS , Hines WC , Yaswen P , Werb Z , Ewald AJ (2012) **ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium** *Proc Natl Acad Sci U S A* **109**:E2595–604
<https://doi.org/10.1073/pnas.1212834109>
17. Bronsert P , Enderle-Ammour K , Bader M , Timme S , Kuehs M , Csanadi A , Kayser G , Kohler I , Bausch D , Hoepfner J , Hopt UT , Keck T , Stickeler E , Passlick B , Schilling O , Reiss CP , Vashist Y , Brabletz T , Berger J , Lotz J , Olesch J , Werner M , Wellner UF (2014) **Cancer cell invasion and EMT marker expression: a three-dimensional study of the human cancer-host interface** *J Pathol* **234**:410–22
<https://doi.org/10.1002/path.4416>
18. Cheung KJ , Padmanaban V , Silvestri V , Schipper K , Cohen JD , Fairchild AN , Gorin MA , Verdone JE , Pienta KJ , Bader JS , Ewald AJ (2016) **Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters** *Proc Natl Acad Sci U S A* **113**:E854–63
<https://doi.org/10.1073/pnas.1508541113>
19. Pearson GW (2019) **Control of Invasion by Epithelial-to-Mesenchymal Transition Programs during Metastasis** *J Clin Med* **8**
<https://doi.org/10.3390/jcm8050646>
20. Cheung KJ , Gabrielson E , Werb Z , Ewald AJ (2013) **Collective invasion in breast cancer requires a conserved basal epithelial program** *Cell* **155**:1639–51
<https://doi.org/10.1016/j.cell.2013.11.029>
21. Westcott JM , Precht AM , Maine EA , Dang TT , Esparza MA , Sun H , Zhou Y , Xie Y , Pearson GW (2015) **An epigenetically distinct breast cancer cell subpopulation promotes collective invasion** *J Clin Invest* **125**:1927–43
<https://doi.org/10.1172/JCI77767>
22. Kim YH , Choi YW , Lee J , Soh EY , Kim JH , Park TJ (2017) **Senescent tumor cells lead the collective invasion in thyroid cancer** *Nat Commun* **8**
<https://doi.org/10.1038/ncomms15208>
23. Summerbell ER , Mouw JK , Bell JSK , Knippler CM , Pedro B , Arnst JL , Khatib TO , Commander R , Barwick BG , Konen J , Dwivedi B , Seby S , Kowalski J , Vertino PM , Marcus AI (2020) **Epigenetically heterogeneous tumor cells direct collective invasion through filopodia-driven fibronectin micropatterning** *Sci Adv* **6**
<https://doi.org/10.1126/sci-adv.aaz6197>
24. Konen J , Summerbell E , Dwivedi B , Galior K , Hou Y , Rusnak L , Chen A , Saltz J , Zhou W , Boise LH , Vertino P , Cooper L , Salaita K , Kowalski J , Marcus AI (2017) **Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion** *Nat Commun* **8**
<https://doi.org/10.1038/ncomms15078>

25. Westcott JM , Camacho S , Nasir A , Huysman ME , Rahhal R , Dang TT , Riegel AT , Brekken RA , Pearson GW (2020) **Δ Np63-Regulated Epithelial-to-Mesenchymal Transition State Heterogeneity Confers a Leader-Follower Relationship That Drives Collective Invasion** *Cancer Research* **80**:3933–44
<https://doi.org/10.1158/0008-5472.Can-20-0014>
26. Tabassum DP , Polyak K (2015) **Tumorigenesis: it takes a village** *Nat Rev Cancer* **15**:473–83
<https://doi.org/10.1038/nrc3971>
27. Celia-Terrassa T , Kang Y (2016) **Distinctive properties of metastasis-initiating cells** *Genes Dev* **30**:892–908
<https://doi.org/10.1101/gad.277681.116>
28. Maine EA , Westcott JM , Precht AM , Dang TT , Whitehurst AW , Pearson GW (2016) **The cancer-testis antigens SPANX-A/C/D and CTAG2 promote breast cancer invasion** *Oncotarget* **7**:14708–26
<https://doi.org/10.18632/oncotarget.7408>
29. Padmanaban V , Tsehay Y , Cheung KJ , Ewald AJ , Bader JS (2020) **Between-tumor and within-tumor heterogeneity in invasive potential** *PLoS Comput Biol* **16**
<https://doi.org/10.1371/journal.pcbi.1007464>
30. Winkler J , Abisoye-Ogunniyan A , Metcalf KJ , Werb Z (2020) **Concepts of extracellular matrix remodelling in tumour progression and metastasis** *Nat Commun* **11**
<https://doi.org/10.1038/s41467-020-18794-x>
31. Dang TT , Esparza MA , Maine EA , Westcott JM , Pearson GW (2015) **DeltaNp63alpha Promotes Breast Cancer Cell Motility through the Selective Activation of Components of the Epithelial-to-Mesenchymal Transition Program** *Cancer Res* **75**:3925–35
<https://doi.org/10.1158/0008-5472.CAN-14-3363>
32. Li CI , Uribe DJ , Daling JR (2005) **Clinical characteristics of different histologic types of breast cancer** *British Journal of Cancer* **93**:1046–52
<https://doi.org/10.1038/sj.bjc.6602787>
33. Yersal O , Barutca S (2014) **Biological subtypes of breast cancer: Prognostic and therapeutic implications** *World J Clin Oncol* **5**:412–24
<https://doi.org/10.5306/wjco.v5.i3.412>
34. Green JE , Shibata MA , Yoshidome K , Liu ML , Jorcyk C , Anver MR , Wigginton J , Wilttrout R , Shibata E , Kaczmarczyk S , Wang W , Liu ZY , Calvo A , Couldrey C (2000) **The C3(1)/SV40 T-antigen transgenic mouse model of mammary cancer: ductal epithelial cell targeting with multi-stage progression to carcinoma** *Oncogene* **19**:1020–7
<https://doi.org/10.1038/sj.onc.1203280>

35. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S, Backlund MG, Yin Y, Khramtsov AI, Bastein R, Quackenbush J, Glazer RI, Brown PH, Green JE, Kopelovich L, Furth PA, Palazzo JP, Olopade OI, Bernard PS, Churchill GA, Van Dyke T, Perou CM (2007) **Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors** *Genome Biol* **8**
<https://doi.org/10.1186/gb-2007-8-5-r76>
36. Hwang PY, Brenot A, King AC, Longmore GD, George SC (2019) **Randomly Distributed K14(+) Breast Tumor Cells Polarize to the Leading Edge and Guide Collective Migration in Response to Chemical and Mechanical Environmental Cues** *Cancer Res* **79**:1899–912
<https://doi.org/10.1158/0008-5472.Can-18-2828>
37. Bergholtz H, Lien TG, Swanson DM, Frigessi A, Bathen TF, Borgen E, Børresen-Dale AL, Engebråten O, Garred Ø, Geisler J, Geitvik GA, Hartmann-Johnsen OJ, Hofvind S, Kristensen VN, Langerød A, Lingjærde OC, Mælandsmo GM, Naume B, Russnes H, Sauer T, Schlichting E, Skjerven HK, Daidone MG, Tost J, Wärnberg F, Sørli T (2020) **Oslo Breast Cancer Research C. Contrasting DCIS and invasive breast cancer by subtype suggests basal-like DCIS as distinct lesions** *npj Breast Cancer* **6**
<https://doi.org/10.1038/s41523-020-0167-x>
38. Christin JR, Wang C, Chung CY, Liu Y, Dravis C, Tang W, Oktay MH, Wahl GM, Guo W (2020) **Stem Cell Determinant SOX9 Promotes Lineage Plasticity and Progression in Basal-like Breast Cancer** *Cell Rep* **31**
<https://doi.org/10.1016/j.celrep.2020.107742>
39. DeRose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MT, Factor R, Matsen C, Milash BA, Nelson E, Neumayer L, Randall RL, Stijleman IJ, Welm BE, Welm AL (2011) **Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes** *Nat Med* **17**:1514–20
<https://doi.org/10.1038/nm.2454>
40. Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, Van Keymeulen A, Brown D, Moers V, Lemaire S, De Clercq S, Minguijon E, Balsat C, Sokolow Y, Dubois C, De Cock F, Scozzaro S, Sopena F, Lanas A, D'Haene N, Salmon I, Marine JC, Voet T, Sotiropoulou PA, Blanpain C (2018) **Identification of the tumour transition states occurring during EMT** *Nature* **556**:463–8
<https://doi.org/10.1038/s41586-018-0040-3>
41. Altschuler SJ, Wu LF (2010) **Cellular heterogeneity: do differences make a difference?** *Cell* **141**:559–63
<https://doi.org/10.1016/j.cell.2010.04.033>
42. Eferl R, Wagner EF (2003) **AP-1: a double-edged sword in tumorigenesis** *Nature Reviews Cancer* **3**:859–68
<https://doi.org/10.1038/nrc1209>

43. Bakiri L , Macho-Maschler S , Custic I , Niemiec J , Guío-Carrión A , Hasenfuss SC , Eger A , Müller M , Beug H , Wagner EF (2015) **Fra-1/AP-1 induces EMT in mammary epithelial cells by modulating Zeb1/2 and TGF β expression** *Cell Death Differ* **22**:336–50
<https://doi.org/10.1038/cdd.2014.157>
44. Liu ML , Von Lintig FC , Liyanage M , Shibata MA , Jorcyk CL , Ried T , Boss GR , Green JE (1998) **Amplification of Ki-ras and elevation of MAP kinase activity during mammary tumor progression in C3(1)/SV40 Tag transgenic mice** *Oncogene* **17**:2403–11
<https://doi.org/10.1038/sj.onc.1202456>
45. John G Albeck , Gordon B Mills , Joan S Brugge (2013) **Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals** *Molecular Cell* **49**:249–61
<https://doi.org/10.1016/j.molcel.2012.11.002>
46. Neve RM , Chin K , Fridlyand J , Yeh J , Baehner FL , Fevr T , Clark L , Bayani N , Coppe JP , Tong F , Speed T , Spellman PT , DeVries S , Lapuk A , Wang NJ , Kuo WL , Stilwell JL , Pinkel D , Albertson DG , Waldman FM , McCormick F , Dickson RB , Johnson MD , Lippman M , Ethier S , Gazdar A , Gray JW (2006) **A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes** *Cancer Cell* **10**:515–27
<https://doi.org/10.1016/j.ccr.2006.10.008>
47. Siersbæk R , Madsen JGS , Javierre BM , Nielsen R , Bagge EK , Cairns J , Wingett SW , Traynor S , Spivakov M , Fraser P , Mandrup S (2017) **Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation** *Molecular Cell* **66**:420–35
<https://doi.org/10.1016/j.molcel.2017.04.010>
48. Oft M , Peli J , Rudaz C , Schwarz H , Beug H , Reichmann E (1996) **TGF-beta1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells** *Genes Dev* **10**:2462–77
<https://doi.org/10.1101/gad.10.19.2462>
49. Derynck R , Muthusamy BP , Saeteurn KY (2014) **Signaling pathway cooperation in TGF- β -induced epithelial-mesenchymal transition** *Current Opinion in Cell Biology* **31**:56–66
<https://doi.org/10.1016/j.ceb.2014.09.001>
50. Parvani JG , Galliher-Beckley AJ , Schiemann BJ , Schiemann WP (2013) **Targeted inactivation of β 1 integrin induces β 3 integrin switching, which drives breast cancer metastasis by TGF- β** *Mol Biol Cell* **24**:3449–59
<https://doi.org/10.1091/mbc.E12-10-0776>
51. Blee AM , Huang H (2015) **ERG-Mediated Cell Invasion: A Link between Development and Tumorigenesis** *Medical Epigenetics* **3**:19–29
<https://doi.org/10.1159/000440978>
52. Huang P , Chen A , He W , Li Z , Zhang G , Liu Z , Liu G , Liu X , He S , Xiao G , Huang F , Stenvang J , Brünner N , Hong A , Wang J (2017) **BMP-2 induces EMT and breast cancer stemness through Rb and CD44** *Cell Death Discovery* **3**
<https://doi.org/10.1038/cddiscovery.2017.39>

53. González-González L , Alonso J. (2018) **Periostin: A Matricellular Protein With Multiple Functions in Cancer Development and Progression** *Front Oncol* **8**
<https://doi.org/10.3389/fonc.2018.00225>
54. Ma B , Ran R , Liao H-Y , Zhang H-H (2021) **The paradoxical role of matrix metalloproteinase-11 in cancer** *Biomedicine & Pharmacotherapy* **141**
<https://doi.org/10.1016/j.biopha.2021.111899>
55. Wu S , Chen M , Huang J , Zhang F , Lv Z , Jia Y , Cui Y-Z , Sun L-Z , Wang Y , Tang Y , Verhoeft KR , Li Y , Qin Y , Lin X , Guan X-Y , Lam K-O (2021) **ORAI2 Promotes Gastric Cancer Tumorigenicity and Metastasis through PI3K/Akt Signaling and MAPK-Dependent Focal Adhesion Disassembly** *Cancer Research* **81**
<https://doi.org/10.1158/0008-5472.CAN-20-0049>
56. Holland SJ , Powell MJ , Franci C , Chan EW , Frieria AM , Atchison RE , McLaughlin J , Swift SE , Pali ES , Yam G , Wong S , Lasaga J , Shen MR , Yu S , Xu W , Hitoshi Y , Bogenberger J , Nör JE , Payan DG , Lorens JB (2005) **Multiple roles for the receptor tyrosine kinase axl in tumor formation** *Cancer Res* **65**:9294–303
<https://doi.org/10.1158/0008-5472.Can-05-0993>
57. Zecchini S , Bombardelli L , Decio A , Bianchi M , Mazzarol G , Sanguineti F , Aletti G , Maddaluno L , Berezin V , Bock E , Casadio C , Viale G , Colombo N , Giavazzi R , Cavallaro U (2011) **The adhesion molecule NCAM promotes ovarian cancer progression via FGFR signalling** *EMBO Molecular Medicine* **3**:480–94
<https://doi.org/10.1002/emmm.201100152>
58. Black SA , Nelson AC , Gurule NJ , Futscher BW , Lyons TR (2016) **Semaphorin 7a exerts pleiotropic effects to promote breast tumor progression** *Oncogene* **35**:5170–8
<https://doi.org/10.1038/onc.2016.49>
59. Adorno-Cruz V , Hoffmann AD , Liu X , Dashzeveg NK , Taftaf R , Wray B , Keri RA , Liu H (2021) **ITGA2 promotes expression of ACLY and CCND1 in enhancing breast cancer stemness and metastasis** *Genes & Diseases* **8**:493–508
<https://doi.org/10.1016/j.gendis.2020.01.015>
60. Sayan AE , Stanford R , Vickery R , Grigorenko E , Diesch J , Kulbicki K , Edwards R , Pal R , Greaves P , Jariel-Encontre I , Piechaczyk M , Kriajevska M , Mellon JK , Dhillon AS , Tulchinsky E (2012) **Fra-1 controls motility of bladder cancer cells via transcriptional upregulation of the receptor tyrosine kinase AXL** *Oncogene* **31**:1493–503
<https://doi.org/10.1038/onc.2011.336>
61. Cancer Genome Atlas Research N (2015) **The Molecular Taxonomy of Primary Prostate Cancer** *Cell* **163**:1011–25
<https://doi.org/10.1016/j.cell.2015.10.025>
62. Chen F , Zhang Y , Parra E , Rodriguez J , Behrens C , Akbani R , Lu Y , Kurie JM , Gibbons DL , Mills GB , Wistuba II , Creighton CJ (2017) **Multiplatform-based molecular subtypes of non-small-cell lung cancer** *Oncogene* **36**:1384–93
<https://doi.org/10.1038/onc.2016.303>

63. McConkey DJ , Choi W (2018) **Molecular Subtypes of Bladder Cancer** *Current Oncology Reports* **20**
<https://doi.org/10.1007/s11912-018-0727-5>
64. Collisson EA , Bailey P , Chang DK , Biankin AV (2019) **Molecular subtypes of pancreatic cancer** *Nature Reviews Gastroenterology & Hepatology* **16**:207–20
<https://doi.org/10.1038/s41575-019-0109-y>
65. Katsuno Y , Derynck R (2021) **Epithelial plasticity, epithelial-mesenchymal transition, and the TGF- β family** *Developmental Cell* **56**:726–46
<https://doi.org/10.1016/j.devcel.2021.02.028>
66. van Dijk D , Sharma R , Nainys J , Yim K , Kathail P , Carr AJ , Burdziak C , Moon KR , Chaffer CL , Pattabiraman D , Bieri B , Mazutis L , Wolf G , Krishnaswamy S , Pe'er D (2018) **Recovering Gene Interactions from Single-Cell Data Using Data Diffusion** *Cell* **174**:716–29
<https://doi.org/10.1016/j.cell.2018.05.061>
67. Lambert AW , Weinberg RA (2021) **Linking EMT programmes to normal and neoplastic epithelial stem cells** *Nat Rev Cancer* **21**:325–38
<https://doi.org/10.1038/s41568-021-00332-6>
68. Chakrabarti R , Hwang J , Andres Blanco M , Wei Y , Lukačičin M , Romano RA , Smalley K , Liu S , Yang Q , Ibrahim T , Mercatali L , Amadori D , Haffty BG , Sinha S , Kang Y (2012) **Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2** *Nat Cell Biol* **14**:1212–22
<https://doi.org/10.1038/ncb2607>
69. Lehmann W , Mossmann D , Kleemann J , Mock K , Meisinger C , Brummer T , Herr R , Brabletz S , Stemmler MP , Brabletz T (2016) **ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types** *Nature Communications* **7**
<https://doi.org/10.1038/ncomms10498>
70. Doehn U , Hauge C , Frank SR , Jensen CJ , Duda K , Nielsen JV , Cohen MS , Johansen JV , Winther BR , Lund LR , Winther O , Taunton J , Hansen SH , Frödin M (2009) **RSK Is a Principal Effector of the RAS-ERK Pathway for Eliciting a Coordinate Promotile/Invasive Gene Program and Phenotype in Epithelial Cells** *Molecular Cell* **35**:511–22
<https://doi.org/10.1016/j.molcel.2009.08.002>
71. Shin S , Dimitri CA , Yoon S-O , Dowdle W , Blenis J (2010) **ERK2 but Not ERK1 Induces Epithelial-to-Mesenchymal Transformation via DEF Motif-Dependent Signaling Events** *Molecular Cell* **38**:114–27
<https://doi.org/10.1016/j.molcel.2010.02.020>
72. Lee MK , Pardoux C , Hall MC , Lee PS , Warburton D , Qing J , Smith SM , Derynck R (2007) **TGF- β activates Erk MAP kinase signalling through direct phosphorylation of ShcA** *The EMBO Journal* **26**:3957–67
<https://doi.org/10.1038/sj.emboj.7601818>

73. Zhao Y , Ma J , Fan Y , Wang Z , Tian R , Ji W , Zhang F , Niu R (2018) **TGF- β transactivates EGFR and facilitates breast cancer migration and invasion through canonical Smad3 and ERK/Sp1 signaling pathways** *Mol Oncol* **12**:305–21
<https://doi.org/10.1002/1878-0261.12162>
74. Blanchette F , Rivard N , Rudd P , Grondin F , Attisano L , Dubois CM (2001) **Cross-talk between the p42/p44 MAP Kinase and Smad Pathways in Transforming Growth Factor β -induced Furin Gene Transactivation** * *Journal of Biological Chemistry* **276**:33986–94
<https://doi.org/10.1074/jbc.M100093200>
75. Hough C , Radu M , Doré JJE (2012) **TGF-Beta Induced Erk Phosphorylation of Smad Linker Region Regulates Smad Signaling** *PLOS ONE* **7**
<https://doi.org/10.1371/journal.pone.0042513>
76. Pearson GW , Hunter T (2007) **Real-time imaging reveals that noninvasive mammary epithelial acini can contain motile cells** *J Cell Biol* **179**:1555–67
<https://doi.org/10.1083/jcb.200706099>
77. Dang TT , Precht AM , Pearson GW (2011) **Breast cancer subtype-specific interactions with the microenvironment dictate mechanisms of invasion** *Cancer Res* **71**:6857–66
<https://doi.org/10.1158/0008-5472.Can-11-1818>
78. Wrenn ED , Yamamoto A , Moore BM , Huang Y , McBirney M , Thomas AJ , Greenwood E , Rabena YF , Rahbar H , Partridge SC , Cheung KJ (2020) **Regulation of Collective Metastasis by Nanolumenal Signaling** *Cell* **183**:395–410
<https://doi.org/10.1016/j.cell.2020.08.045>
79. Massagué J (2012) **TGF β signalling in context** *Nat Rev Mol Cell Biol* **13**:616–30
<https://doi.org/10.1038/nrm3434>
80. David CJ , Massagué J (2018) **Contextual determinants of TGF β action in development, immunity and cancer** *Nat Rev Mol Cell Biol* **19**:419–35
<https://doi.org/10.1038/s41580-018-0007-0>
81. Clement JH , Raida M , Sanger J , Bicknell R , Liu J , Naumann A , Geyer A , Waldau A , Hortschansky P , Schmidt A , Hoffken K , Wolf S , Harris AL (2005) **Bone morphogenetic protein 2 (BMP-2) induces in vitro invasion and in vivo hormone independent growth of breast carcinoma cells** *Int J Oncol* **27**:401–7
82. Liberzon A , Birger C , Thorvaldsdottir H , Ghandi M , Mesirov JP , Tamayo P (2015) **The Molecular Signatures Database (MSigDB) hallmark gene set collection** *Cell Syst* **1**:417–25
<https://doi.org/10.1016/j.cels.2015.12.004>
83. Huang P , Chen A , He W , Li Z , Zhang G , Liu Z , Liu G , Liu X , He S , Xiao G , Huang F , Stenvang J , Brunner N , Hong A , Wang J (2017) **BMP-2 induces EMT and breast cancer stemness through Rb and CD44** *Cell Death Discov* **3**
<https://doi.org/10.1038/cddiscovery.2017.39>

84. Cheng YX, Xiao L, Yang YL, Liu XD, Zhou XR, Bu ZF, Cao PC, Wang DK (2021) **Collagen type VIII alpha 2 chain (COL8A2), an important component of the basement membrane of the corneal endothelium, facilitates the malignant development of glioblastoma cells via inducing EMT** *J Bioenerg Biomembr* **53**:49–59
<https://doi.org/10.1007/s10863-020-09865-1>
85. Zhao J, Ou B, Han D, Wang P, Zong Y, Zhu C, Liu D, Zheng M, Sun J, Feng H, Lu A (2017) **Tumor-derived CXCL5 promotes human colorectal cancer metastasis through activation of the ERK/Elk-1/Snail and AKT/GSK3 β / β -catenin pathways** *Mol Cancer* **16**
<https://doi.org/10.1186/s12943-017-0629-4>
86. Susek KH, Karvouni M, Alici E, Lundqvist A (2018) **The Role of CXC Chemokine Receptors 1-4 on Immune Cells in the Tumor Microenvironment** *Front Immunol* **9**
<https://doi.org/10.3389/fimmu.2018.02159>
87. Roberts PJ, Usary JE, Darr DB, Dillon PM, Pfefferle AD, Whittle MC, Duncan JS, Johnson SM, Combest AJ, Jin J, Zamboni WC, Johnson GL, Perou CM, Sharpless NE (2012) **Combined PI3K/mTOR and MEK inhibition provides broad antitumor activity in faithful murine cancer models** *Clin Cancer Res* **18**:5290–303
<https://doi.org/10.1158/1078-0432.CCR-12-0563>
88. Usary J, Zhao W, Darr D, Roberts PJ, Liu M, Balletta L, Karginova O, Jordan J, Combest A, Bridges A, Prat A, Cheang MC, Herschkowitz JI, Rosen JM, Zamboni W, Sharpless NE, Perou CM (2013) **Predicting drug responsiveness in human cancers using genetically engineered mice** *Clin Cancer Res* **19**:4889–99
<https://doi.org/10.1158/1078-0432.CCR-13-0522>
89. Reguart N, Remon J (2015) **Common EGFR-mutated subgroups (Del19/L858R) in advanced non-small-cell lung cancer: chasing better outcomes with tyrosine kinase inhibitors** *Future Oncology* **11**:1245–57
<https://doi.org/10.2217/fon.15.15>
90. Li S, Liu M, Do MH, Chou C, Stamatiades EG, Nixon BG, Shi W, Zhang X, Li P, Gao S, Capistrano KJ, Xu H, Cheung N-KV, Li MO (2020) **Cancer immunotherapy via targeted TGF- β signalling blockade in TH cells** *Nature* **587**:121–5
<https://doi.org/10.1038/s41586-020-2850-3>
91. Kanda T, Sullivan KF, Wahl GM (1998) **Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells** *Curr Biol* **8**:377–85
[https://doi.org/10.1016/s0960-9822\(98\)70156-3](https://doi.org/10.1016/s0960-9822(98)70156-3)
92. Benjamini Y, Hochberg Y (1995) **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing** *Journal of the Royal Statistical Society: Series B (Methodological)* **57**:289–300
<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
93. Love MI, Huber W, Anders S (2014) **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2** *Genome Biology* **15**
<https://doi.org/10.1186/s13059-014-0550-8>

94. Rauch A , Haakonsson AK , Madsen JGS , Larsen M , Forss I , Madsen MR , Van Hauwaert EL , Wiwie C , Jespersen NZ , Tencerova M , Nielsen R , Larsen BD , Röttger R , Baumbach J , Scheele C , Kassem M , Mandrup S (2019) **Osteogenesis depends on commissioning of a network of stem cell transcription factors that act as repressors of adipogenesis** *Nature Genetics* **51**:716–27
<https://doi.org/10.1038/s41588-019-0359-1>
95. Yu G , Wang L-G , Han Y , He Q-Y (2012) **clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters** *OMICS: A Journal of Integrative Biology* **16**:284–7
<https://doi.org/10.1089/omi.2011.0118>
96. Fernandez AI , Geng X , Chaldekias K , Harris B , Duttargi A , Berry VL , Berry DL , Mahajan A , Cavalli LR , Györfy B , Tan M , Riggins RB (2020) **The orphan nuclear receptor estrogen-related receptor beta (ERR β) in triple-negative breast cancer** *Breast Cancer Research and Treatment* **179**:585–604
<https://doi.org/10.1007/s10549-019-05485-5>
97. Cerami E , Gao J , Dogrusoz U , Gross BE , Sumer SO , Aksoy BA , Jacobsen A , Byrne CJ , Heuer ML , Larsson E , Antipin Y , Reva B , Goldberg AP , Sander C , Schultz N (2012) **The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data** *Cancer Discov* **2**:401–4
<https://doi.org/10.1158/2159-8290.CD-12-0095>
98. Liu J , Lichtenberg T , Hoadley KA , Poisson LM , Lazar AJ , Cherniack AD , Kovatich AJ , Benz CC , Levine DA , Lee AV , Omberg L , Wolf DM , Shriver CD , Thorsson V , Hu H (2018) **An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics** *Cell* **173**:400–16
<https://doi.org/10.1016/j.cell.2018.02.052>

Author information

Apsra Nasir

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0003-3644-911X](https://orcid.org/0000-0003-3644-911X)

Sharon Camacho

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0003-2390-0880](https://orcid.org/0000-0003-2390-0880)

Alec T. McIntosh

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0002-7764-3608](https://orcid.org/0000-0002-7764-3608)

Garrett T. Graham

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0002-1097-0962](https://orcid.org/0000-0002-1097-0962)

Raneen Rahhal

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0002-6428-1118](https://orcid.org/0000-0002-6428-1118)

Molly E. Huysman

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA

Fahda Alsharief

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0009-0006-5799-9144](https://orcid.org/0009-0006-5799-9144)

Anna T. Riegel

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA
ORCID iD: [0000-0001-6585-8612](https://orcid.org/0000-0001-6585-8612)

Gray W. Pearson

Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University, Washington, DC 20057, USA

For correspondence: gp507@georgetown.edu

ORCID iD: [0000-0001-6961-0227](https://orcid.org/0000-0001-6961-0227)

Editors

Reviewing Editor

Lynne-Marie Postovit

University of Alberta, Canada

Senior Editor

Erica Golemis

Fox Chase Cancer Center, United States of America

Reviewer #1 (Public Review):

The study investigates the nature of "trailblazer" cells in distinct tumor models, including luminal B (MMTV/PyMT) and triple negative (TNBC) tumors (C3-TAg). The authors note that the trailblazer phenotypes in the TNBC model are more complex relative to the Luminal B model and represent distinct EMT programs associated with the expression of distinct EMT-TFs (Zeb1, Zeb2 and Fra-1). They demonstrated that of numerous EMT-TFs, Zeb1 and Fra-1 were required for increased cancer cell migration and invasion. They reveal that TGF-beta and EGF-mediated signaling are required for the diverse EMT states that are required for trailblazer cell activity and increased cell migration/invasion. TGF-beta signaling engaged Zeb 1 and Zeb2 while EGF signaling activated Fra-1. Indeed, inhibitors of either TGF-beta or EGF signaling could impair cell migration/invasion. While both pathways contributed to trailblazer phenotypes, EGF signaling was shown to interfere with certain TGF-beta induced

transcriptional response, including the expression of genes encoding extracellular matrix proteins.

One concern was the heavy reliance of the C3-Tag as the sole TNBC model in which the distinct trailblazer phenotypes were described. The data in Fig. 3 of the submission reveals that the phenotypes observed in the C3-Tag model could be recapitulated in a TNBC patient-derived xenograft model (PDX). Using this PDX, the authors were able to show vimentin expression in lung metastatic TNBC cells that were intravascular, those that had extravasated and clusters of cancer cells fully within the lung parenchyma. This was an important addition to the manuscript. The additional experiments to investigate the role of Zeb1 and Zeb1 more fully, beyond the focus on Fra-1 in the initial submission was an additional strength of the new submission. Additional clarifications to the discussion also clarified the concepts articulated in the study. The study employs multiple breast cancer models, utilizes numerous in vitro and in vivo assessments of the trailblazer phenotypes, and the experimental design is rigorous and the interpretation of the data is sound. The manuscript will be of general interest to the research community.

Reviewer #2 (Public Review):

This represents an important study that demonstrates a high degree of heterogeneity within trailblazer cells in clusters that participate in collective migration. Solid methods highlight this heterogeneity and show that in TNBC cancers, trailblazer cells are defined by vimentin (and not Keratin 14) and are dependent on both TGFbeta and EGFR signaling. Additional, single cell studies would further support this work.

Strengths:

The paper highlights that collective migration, and the nature of trailblazer cells can be highly heterogeneous. This is important as it suggests that the ability to move between states may supersede a singular phenotype.

The paper uses animal models and organoids and in several areas attempts to correlate findings to human tissues.

The experiments are logically described.

Reviewer #3 (Public Review):

Cancer is a disease of many faces and in particular, the ability of cancers cells to change their phenotypes and cell behaviors - cancer cell plasticity - is a major contributor to cancer lethality and therapeutic challenge of treating this disease. In this study, Nasir, Pearson et al., investigate tumor cell plasticity through the lens of invasive heterogeneity, and in particular in models of triple-negative breast cancer (TNBC), a subtype of breast cancer with particularly poor clinical prognosis and more limited treatment modalities. Using organoid models in a variety of matrix systems, microscopy, and signaling pathway inhibitors, they find that invading TNBC breast tumors, primarily in the C31-Tag genetically engineered mouse model of TNBC, are composed of heterogeneous invasive/"trailblazer" type tumor cells that in many cases express vimentin, a classical intermediate filament marker of epithelial-mesenchymal transition, and reduced keratin-14, another filament marker of basal epithelial cells associated with collective invasion in different breast cancer models. Supportive genetic and pharmacologic evidence is provided that generation of these cells is TGF-beta signaling pathway driven, likely in vivo from the surrounding tumor microenvironment, in accord with published studies in this space. Another important aspect

of this study is the good transcriptional evidence for multiple migratory states showing differing degrees of partial overlap with canonical EMT programs, dependent on TGF-beta, and suggestive but at present incomplete understanding of a parallel program involving Egfr/Fra-1 mediated effects on invasion. When taken in context with other recent studies (Grasset et al. Science Translational Medicine 2022), these data are broadly supportive of concept of targeting vimentin-dependent invasion programs in TNBC tumors.

The core conclusions of this paper are generally supported by the data, but there are some conceptual and technical considerations that should be taken into account when interpreting this study. Specific comments:

1. The contribution of the different vimentin-positive trailblazer cells to distant metastasis was not directly confirmed in vivo in this study. Given the limited proliferative potential of many fully EMT'd cells and in light of recent studies indicating that invasion can be uncoupled from metastatic potential, it seems important to directly test whether the different C31-tag isolates, varying in invasive potential in this study, produce metastases and if so do metastases abundance correlate with the invasive potential in 3D culture. The collection of lungs at 34 days post injection described in methods is too short to evaluate metastatic frequency.
2. The invasion of cancer cells is dependent on 3D matrix composition. In other studies, collective cancer invasion is performed in exclusively collagen type 1 gels or in other instances entirely in 3D reconstituted basement membrane gel, e.g. lung cancer invasion studies. In this study, the authors use a mixture composed of both matrices. Given the invasion suppressive effects of matrigel, particularly for epithelial type cells, further studies would be important to determine whether the invasion phenotypes seen in this study are generalizable across matrix environments.
3. TGF-beta is well known to induce EMT. Although this study identifies potential transcriptional mediators of the invasion/trailblazer program, is this program reversible?

Author Response:

Reviewer #1 (Public Review):

The study investigates the nature of "trailblazer" cells in distinct tumor models, including luminal B (MMTV/PyMT) and triple negative (TNBC) tumors (C3-TAG). The authors note that the trail-blazer phenotypes in the TNBC model are more complex relative to the Luminal B model and represent distinct EMT programs associated with the expression of distinct EMT-TFs (Zeb1, Zeb2 and Fra-1). They demonstrated that of numerous EMT-TFs, Zeb1 and Fra-1 were required for increased cancer cell migration and invasion. They reveal that TGF-beta and EGF-mediated signaling are required for the diverse EMT states that are required for trailblazer cell activity and increased cell migration/invasion. TGF-beta signaling engaged Zeb 1 and Zeb2 while EGF sig-naling activated Fra-1. Indeed, inhibitors of either TGF-beta or EGF signaling could impair cell migration/invasion. While both pathways contributed to trailblazer phenotypes, EGF signaling was shown to interfere with certain TGF-beta induced transcriptional response, including the ex-pression of genes encoding extracellular matrix proteins.

One concern was the heavy reliance of the C3-TAG as the sole TNBC model in which the dis-tinct trailblazer phenotypes were described. The data in Fig. 3 of the submission reveals that the phenotypes observed in the C3-TAG model could be recapitulated in a TNBC patient-derived xenograft model (PDX). Using this PDX, the authors were able to show vimentin expression in lung metastatic TNBC cells that were intravascular, those that had extravasated and clusters of

cancer cells fully within the lung parenchyma. This was an important addition to the manuscript. The additional experiments to investigate the role of Zeb1 and Zeb1 more fully, beyond the focus on Fra-1 in the initial submission was an additional strength of the new submission. Additional clarifications to the discussion also clarified the concepts articulated in the study. The study employs multiple breast cancer models, utilizes numerous in vitro and in vivo assessments of the trailblazer phenotypes, and the experimental design is rigorous and the interpretation of the data is sound. The manuscript will be of general interest to the research community.

Thank you for the supportive comments. We are glad that the revisions addressed your prior concerns.

Reviewer #2 (Public Review):

This represents an important study that demonstrates a high degree of heterogeneity within trailblazer cells in clusters that participate in collective migration. Solid methods highlight this heterogeneity and show that in TNBC cancers, trailblazer cells are defined by vimentin (and not Keratin 14) and are dependent on both TGFbeta and EGFR signaling. Additional, single cell studies would further support this work.

Thank you for the suggestion. Our current data establishes that trailblazer cells are heterogeneous using FACS, immunostaining and functional studies of fresh tumor organoids and established tumor organoid lines. In addition, our RNA-seq experiments provided deep insight into the nature of gene expression changes that corresponded with the evolution of new trailblazer states. This discovery of trailblazer cell heterogeneity was one of multiple key new discoveries in this manuscript, along with revealing a Krt14-independent invasion mechanism, the regulation of trailblazer cells by Tgfβ and Egfr signaling and a new compromise mode of signal integration. We agree that our results support further investigation of the nature and function of basal-like breast cancer heterogeneity during the progression to metastasis. However, a comprehensive implementation of scRNA-seq is mostly likely required to further unravel new aspects of heterogeneity that substantially advance upon the conclusions supported by our current data. Such an undertaking is beyond the scope of this investigation.

We agree that scRNA-seq would be confirmatory of trailblazer cell heterogeneity that has been demonstrated with multiple approaches rather than a new discovery of heterogeneity.

Strengths:

The paper highlights that collective migration, and the nature of trailblazer cells can be highly heterogeneous. This is important as it suggests that the ability to move between states may supersede a singular phenotype.

The paper uses animal models and organoids and in several areas attempts to correlate findings to human tissues.

The experiments are logically described.

Reviewer #3 (Public Review):

Cancer is a disease of many faces and in particular, the ability of cancer cells to change their phenotypes and cell behaviors - cancer cell plasticity - is a major contributor to cancer lethality and therapeutic challenge of treating this disease. In this study, Nasir, Pearson et al., investigate tumor cell plasticity through the lens of invasive heterogeneity, and in particular in models of triple-negative breast cancer (TNBC), a subtype of breast cancer with particularly poor clinical prognosis and more limited treatment modalities. Using organoid models in a variety of matrix systems, microscopy, and signaling pathway inhibitors, they find that invading TNBC breast tu-

mors, primarily in the C31-Tag genetically engineered mouse model of TNBC, are composed of heterogeneous invasive/"trailblazer" type tumor cells that in many cases express vimentin, a classical intermediate filament marker of epithelial-mesenchymal transition, and reduced keratin-14, another filament marker of basal epithelial cells associated with collective invasion in different breast cancer models. Supportive genetic and pharmacologic evidence is provided that generation of these cells is TGF-beta signaling pathway driven, likely in vivo from the surrounding tumor microenvironment, in accord with published studies in this space. Another important aspect of this study is the good transcriptional evidence for multiple migratory states showing differing degrees of partial overlap with canonical EMT programs, dependent on TGF-beta, and suggestive but at present incomplete understanding of a parallel program involving *Egfr/Fra-1* mediated effects on invasion. When taken in context with other recent studies (Grasset et al. *Science Translational Medicine* 2022), these data are broadly supportive of concept of targeting vimentin-dependent invasion programs in TNBC tumors.

The core conclusions of this paper are generally supported by the data, but there are some conceptual and technical considerations that should be taken into account when interpreting this study. Specific comments:

1. The contribution of the different vimentin-positive trailblazer cells to distant metastasis was not directly confirmed in vivo in this study. Given the limited proliferative potential of many fully EMT'd cells and in light of recent studies indicating that invasion can be uncoupled from meta-static potential, it seems important to directly test whether the different C31-tag isolates, varying in invasive potential in this study, produce metastases and if so do metastases abundance correlate with the invasive potential in 3D culture. The collection of lungs at 34 days post injection described in methods is too short to evaluate metastatic frequency.

We agree that it is important to determine the contribution of trailblazer cells towards metastatic dissemination. In this manuscript, we show that Vimentin expressing cells in a triple negative breast cancer (TNBC) PDX model disseminate to the lungs (Figure 3F). We have also shown that Vimentin expressing SUM159 breast cancer (BC) trailblazer cells spontaneously metastasize to the lungs in previous publications (Fig. 2-figure supplement 1C) and (Westcott et al, *J Clin Invest*, 2015, 10.1172/JCI77767 and Maine et al, *Oncotarget*, 2016, 10.18632/oncotarget.7408). Notably, the depletion of genes specifically expressed in trailblazer cells reduced spontaneous metastasis without significantly impinging on primary tumor growth (Westcott et al, *J Clin Invest*, 2015, 10.1172/JCI77767 and Maine et al, *Oncotarget*, 2016, 10.18632/oncotarget.7408). Our new results in Figure 5D show that Tgfβ activates genes that define the trailblazer state in the metastatic SUM159 trailblazer cell model. Thus, features of the Tgfβ regulated trailblazer program in the C3-TAG cells is active in the SUM159 trailblazer model of spontaneous metastasis. In addition, commonly employed BC cell line metastasis models, such as MDAMB231 derivatives are highly mesenchymal (Fig. 2-figure supplement 1C) and (Kang et al, *Cell*, 2003, 10.1016/S1535-6108(03)00132-6 and Minn et al, *Nature*, 2005, 10.1038/nature03799, as examples).

It is not technically feasible to establish a correlation between the relative invasion of The C3-TAG GEMM primary tumors and spontaneous metastasis. C3-TAG GEMM primary tumors develop rapidly and the mice must be euthanized prior to the detection of metastasis. This limitation of the model is mentioned in the Results section "Trailblazer cells are specified by Vimentin expression in basal-like breast cancer patient tumors". The aggressive primary tumor growth and limited spontaneous metastasis of the the C3-TAG model has also been previously reported by others (Green et al, *Oncogene*, 2000, 10.1038/sj.onc.1203280). Surgical resection of the original primary tumor is not feasible option to allow metastases to form since additional tumors develop in multiple mammary glands.

In response to reviewer requests, we initiated the growth of orthotopic primary tumors from control or Tgf β treated 1339-org cells to address the relationship between induction of the trailblazer state and primary tumor cell dissemination. We had to euthanize the mice at day 34 (d34) because tumors within both cohorts had reached the maximum permitted diameter of 2 cm. This will be indicated in the Methods section with revised text. We detected CTCs from the mice bearing control and Tgf β treated 1339-org cell tumors. However, no micrometastases were detected, which is indicated in the text describing Figure 4—figure supplement 3A-B. Thus, performing surgical resection in new experiments would not be expected to allow the later detection of metastasis, as there did not appear to be DTCs in the lungs that could initiate colonization. In addition, we would have to resect the tumors prior to d34 to successfully and humanely remove the primary tumors, further reducing the odds of metastases developing. We will continue our work to identify an experimental balance that permits sufficient primary tumor growth to initiate spontaneous metastasis. However, the time scale of resolving this technical challenge is uncertain and we believe that our published analysis of trailblazer cell metastasis and new findings here showing the dissemination of Vimentin expressing cells in a PDX model addresses the question of whether Vimentin expressing trailblazer cells metastasize.

We agree that certain cell states induced by EMT programs can limit the proliferative potential of tumor cells. As described in the Introduction, we previously found that the induction of a trailblazer state in a subset of breast cancer cell line models triggers a collateral cost in fitness that limits the ability of trailblazer cells to initiate tumor growth (Westcott et al, Cancer Res, 2020, 10.1158/0008-5472.CAN-20-0014). The traits that distinguish trailblazer cells which are capable of tumor initiation and metastasis versus trailblazer cells with reduced fitness have begun to be delineated. Our prior report suggested that cells that were dependent on p63 for growth lost their proliferative capacity when converting to a trailblazer state (Westcott et al, Cancer Res, 2020, 10.1158/0008-5472.CAN-20-0014). C3-TAg cells are not dependent on p63 for growth, which is indicated by the vast majority of the tumor cells lacking p63 expression in primary tumors and primary tumor organoids (Westcott et al, Cancer Res, 2020, 10.1158/0008-5472.CAN-20-0014), similar to the metastatic SUM159 breast cancer cell line model. We were also able to derive clonal trailblazer cell lines that lacked detectable p63 expression from a C3-TAg tumor (Figure 2—figure supplement 1B) and grow organoids even when the limited extent of p63 expression was further reduced by Tgf β (Figure 5C). Additionally, the persistent Tgf β treated 1339-org cells, which were enriched for trailblazer cells and had reduced p63 expression, were capable of initiating primary tumor growth (Figure 4F). Together, these results indicate that C3-TAg trailblazer cells are capable of initiating metastatic colonization. However, given the heterogeneity in trailblazer states that we discovered, it is possible that a subset of trailblazer cell states have reduced proliferative capacity. Our analysis approach in this manuscript would not necessarily detect these low fitness trailblazer cells if they were a relatively small fraction of the total trailblazer population. We will clarify this point in the Discussion section in the revised manuscript. Our results have begun to reveal mechanisms for the transcriptional regulation of trailblazer cell heterogeneity. We plan to continue delineating the regulatory programs conferring specific transcription state, defining approaches for the prospective isolation of distinct trailblazer subpopulations and determining trailblazer subpopulation specific biomarkers to understand the specific contribution of distinct trailblazer subpopulations towards metastasis. Given the scope of this analysis, it is not feasible to incorporate these future studies into this manuscript.

1. The invasion of cancer cells is dependent on 3D matrix composition. In other studies, collective cancer invasion is performed in exclusively collagen type 1 gels or in other instances entirely in 3D reconstituted basement membrane gel, e.g. lung cancer invasion studies. In this study, the authors use a mixture composed of both matrices. Given the

invasion suppressive effects of matrigel, particularly for epithelial type cells, further studies would be important to determine whether the invasion phenotypes seen in this study are generalizable across matrix environments.

The invasion of C3-TAg and PyMT organoids embedded in a 100% pure reconstituted basement is shown in Fig. 1–figure supplement 1G. We will emphasize that trailblazer invasion was evaluated in multiple ECM compositions with revised text and figure graphic. We also provide images for the reviewer showing that C3-TAg organoids collectively invade in a pure Collagen I ECM. Importantly, these findings are consistent with our results showing that Vimentin expressing cells are associated with basal-like mammary tumor cell invasion in the complex ECM of C3-TAg GEMM primary tumors (Figure 2G) and patient primary tumors (Figure 3D). Moreover, Vimentin expressing cells disseminated to the lungs in the TNBC PDX that we evaluated (Figure 3F).

The ECM composition selected for experiments is dictated by the experimental question(s) being addressed. It is unlikely that mammary tumor cells would only ever collectively invade through an ECM that is either pure Collagen I or pure reconstituted basement membrane (BM). Indeed, it has been proposed that mixtures of Collagen I and BM proteins best reconstitute the complexity of primary tumor ECM (Hooper et al, Methods Enzymol, 2006, 10.1016/S0076-6879(06)06049-6). In line with this observation, mixtures of Collagen I and BM proteins have been routinely used for the past 20 years to define mechanisms of 3D invasion; Xiang and Muthuswamy, Methods Enzymol, 2006, 10.1016/S0076-6879(06)06054-X; Calvo et al, Nat Cell Biol, 2013 10.1038/ncb2756; and Kato et al, eLife, 2023, 10.7554/eLife.76520, as examples).

Consistent with the known complexity of the ECM in the tumor microenvironment (TME), we detect Collagen I and Collagen IV (a key component of experimental BM) in the TME of primary breast cancer tumor models (Westcott et al, J Clin Invest, 2015, 10.1172/JCI77767). Importantly, we have found that a mixture of collagen I and experimentally derived BM proteins reliably reveals breast cancer trailblazer cell invasion mechanisms that promote the malignant progression and metastasis of primary tumors and whose expression correlates with poor patient outcome (Westcott et al, J Clin Invest, 2015, 10.1172/JCI77767 and Westcott et al, Cancer Res, 2020, 10.1158/0008-5472.CAN-20-0014, as examples). Notably, the relative differences in trailblazer and opportunist cell invasive phenotypes are not dictated by the ECM composition used in our 3D assays. We have previously tested the invasion of trailblazer and opportunist subpopulations in different ECM compositions using both spheroid vertical invasion assays (Westcott et al, J Clin Invest, 2015, 10.1172/JCI77767). Increasing collagen I concentration enhanced the relative rate of trailblazer cell invasion, with trailblazer cells always showing a significantly enhanced invasion relative to opportunist cells.

The relationship between trailblazer and opportunist cells that we have detected in primary tumors is recapitulated when using mixtures of Collagen I and BM proteins in our past publications and in this manuscript. The clonal opportunist cell lines derived from a C3-TAg tumor expressed high levels of the transcription factor p63 (Figure 2–figure supplement 1A-B). We previously showed that p63 restricts induction of a trailblazer state in human breast cancer trailblazer cell lines (Westcott et al, Cancer Res, 2020, 10.1158/0008-5472.CAN-20-0014). Notably, we showed that p63 expressing C3-TAg cells were not able to initiate collective invasion in the same ECM composition used in our current manuscript. Moreover, p63 cells in primary C3-TAg tumors were noninvasive opportunist cells that were limited to trailing p63-low trailblazer cells when collectively invading in primary tumors and in organoids (Westcott et al, Cancer Res, 2020). We now show that p63 expressing opportunist cell lines are limited to invading behind primary C3-TAg trailblazer cells and trailblazer cell lines in our 3D invasion assays (Figure 1B and Figure 1–figure supplement 1D-E). Together, these results indicate that the ECM employed in our 3D assays reveals the mechanistic underpinnings of both trailblazer and opportunist cell invasion in primary tumors.

With respect to lung cancer invasion, leader cells that we would classify as trailblazer cells have been isolated from 2 non-small cell lung cancer cell line spheroid models grown in pure reconstituted BM extract (Konen et al, Nat Comm, 2017, 10.1038/ncomms15078). However, it is unclear whether these cell line derived NSCLC trailblazer cells are more intrinsically invasive than non-trailblazer siblings in primary NSCLC tumors or if the traits associated cell line NSCLC trail-blazer cells are required for metastasis. These tests have never been reported to the best of our knowledge. Similarly, it is not clear whether these NSCLC cell line derived trailblazer cells reflect features of primary NSCLC primary tumor cells, as we are unaware of any such comparisons being reported. Thus, there is no reason to consider pure reconstituted BM to be an equivalent or preferred experimental option to define trailblazer cell features. Nevertheless, as we mentioned before, our discovery approach identifies trailblazer cells that are intrinsically more invasive than opportunist siblings across multiple ECM conditions, including pure reconstituted BM and, importantly, in primary tumors.

1. TGF-beta is well known to induce EMT. Although this study identifies potential transcriptional mediators of the invasion/trailblazer program, is this program reversible?

We have previously shown the breast cancer trailblazer cells can convert to an opportunist state, demonstrating that trailblazer states are reversible (Westcott et al, J Clin Invest, 2015, 10.1172/JCI77767). In this manuscript, we show that C3-TAg organoid lines derived in the Tgfr1 inhibitor A83-01 have few if any cells with a trailblazer phenotype relative to C3-TAg primary tumors, suggesting a reversion of the trailblazer state (Fig. 4C and Figure 4-figure supplement 2A-C). However, our results do not entirely rule out the possibility that only non-trailblazer cells grew to establish the organoid lines. Indeed, the problem of tracing phenotypic conversions when evaluating heterogeneous populations is a systemic challenge that extends beyond our analysis of trailblazer cells. Clearly defining the conversion rates for trailblazer cells will require multiple genetic markers to distinguish the different trailblazer states we have now identified, in addition to phenotypic and molecular analysis over multiple days, or possibly weeks. Thus, further definition of the rate of reversion of different trailblazer cells is worthy line of future investigation rather than a feasible objective of this study.