Integrin-TGFβ axis induces partial EMT in basal-like cells to lead collective invasion

Abdelrahman G. El-Gammal[1*], Lisa Jansen[1*], Mathijs P. Verhagen[2$], Hilverd Kremer[1$], Sarmad Soufi[1$], Mina Masoudnia[1$], Denise Westland[1], Peter Bult[3], Riccardo Fodde[4], Monique M. A. Verstegen[5], Martijn Gloerich[1], Saskia W. C. van Mil[1], Antoine A. Khalil[1#]

[1] Center for Molecular Medicine (CMM), University Medical Center Utrecht, The Netherlands

[2] Department of Discovery Oncology, GenenTech, South San Francisco, USA

[3] Department of Pathology, Radboud University Medical Center, Nijmegen, The Netherlands

[4] Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands

[5] Department of Surgery, Erasmus MC Transplant Institute, Erasmus MC – University Medical Center, Rotterdam, The Netherlands

[*, $] Equal contribution

[#] Corresponding author: **a.khalil@umcutrecht.nl**

Abstract

Collective invasion is the predominant mode of cancer cell dissemination in breast cancer and represents the initial step of metastatic spread. Basal-like leader cells drive this process by maintaining cell-cell junctions with the follower cells while extending actin-rich protrusions and remodelling the collagen I-rich peritumoral stroma. These features resemble those of individually-invading cells following epithelial-to-mesenchymal transition (EMT), however, how leader cells acquire these traits while preserving cohesion within the collective remains unclear. Using single-cell RNA sequencing of breast cancer organoids, we identified a basal-like leader cell subset enriched for cytoskeletal, ECM-remodelling, and epithelial junction genes. We show that integrin α2 (Itgα2) links collagen I engagement to mesenchymal reprogramming by inducing Inhba expression and activating TGFβ signalling. This, in turn, upregulates Vimentin while preserving epithelial junction gene expression. In parallel, Itgα2 promotes ECM degradation through a TGFβ-independent mechanism. This dual mechanism supports a model of partial EMT as a regulated state and highlights the Itgα2-TGFβ axis as a potential therapeutic target in aggressive breast cancers.

Introduction

Collective cancer cell invasion is the predominant mode of spread into the tissues surrounding most epithelial tumors[1,2]. In invasive ductal carcinoma (IDC) of the breast, collective invasion may be initiated and guided by cancer cells expressing basal cell markers including Keratin 14 (K14)[3]. A subset of these basal-like cells acquires protrusive and motile features while maintaining cell–cell junctions with follower cancer cells, thereby adopting leader cell functions. Beyond local invasion, the basal-like cells persist throughout the metastatic cascade, contributing to circulating tumor clusters and colonization of distant organs. Their presence in primary tumors has been shown to be required for metastasis *in vivo* and correlates with poor clinical outcomes, underscoring their critical role in tumor progression[4–6]. Thus, understanding how basal-like cells acquire leader cell functions and drive collective invasion would allow for developing intervention strategies to prevent tumor progression and metastasis[3,7].

Basal-like cells localize at the tumor-stroma interface, where interactions with the extracellular matrix (ECM) are critical for acquiring leader cell function. As tumors progress, the ECM undergoes substantial remodeling, including increased fibrillar Collagen I deposition and stiffening, both of which correlate with poor prognosis in breast cancer patients[8–11]. The biochemical and physical changes in the ECM are sensed by cells through ECM receptors. Integrins, a family of transmembrane receptors composed of α and β subunits, are key mediators of ECM sensing[12–14]. Through their extracellular domain, integrins bind to ECM components, while their intracellular domains connect to the cytoskeleton via focal adhesion protein complexes[15]. Actomyosin-generated forces are transmitted through integrins to the ECM. In turn, the ECM exerts counteracting forces dictated by its stiffness and organization[16]. This dynamic force exchange is central to mechanotransduction, the process by which integrins enable cells to sense and respond to ECM rigidity and tension[15–18]. A key effector of this process is Yes-associated protein (YAP), which becomes activated in response to ECM stiffening[19]. We have previously shown that YAP1 is specifically induced in basal-like cells upon interaction with Collagen I, where it drives a transcriptional program that promotes protrusive behavior and collagen fiber realignment[7]. This reciprocal interaction between basal-like cells and Collagen I establishes mechanical and transcriptional feedback loops that reinforce leader cell emergence and invasion.

The actin-rich protrusions and matrix-pulling behavior observed in leader cells are critical for their function in collective invasion and ECM remodeling. These behaviors resemble features of mesenchymal motility, where cells downregulate cell–cell junctions and acquire protrusive and ECM-degrading properties to detach and invade individually, driven by epithelial-to-mesenchymal transition (EMT)[18,20–22]. While the role of EMT in single-cell dissemination is well established[23–25], its contribution to leader cell function and collective invasion remains poorly understood. Recent work has implicated partial EMT in collective invasion and metastatic spread in triple-negative breast cancer[26], yet the underlying mechanisms driving this transition and its precise role in leader cell function remain poorly understood. Indeed, the signals that trigger these transitions and their regulation by ECM interactions remain poorly defined. Unraveling these regulatory mechanisms is essential, as it could uncover key pathways driving leader cell behavior and reveal new therapeutic targets to inhibit metastatic spread.

Here, we test the hypothesis that collagen I-mediated mechanotransduction activates mesenchymal programs in basal-like cells to promote their transition into leader cells. Single-cell RNA-sequencing of breast cancer organoids identified a collagen-specific basal-like subpopulation enriched for EMT genes, including Vimentin. Vimentin was also expressed in leader cells at the invasive front in IDC patient tissues

and was required for collective invasion in vitro. Mechanistically, we identified integrin α2 (Itgα2), part of the α2β1 receptor for fibrillar collagen I, as a key upstream regulator. Itgα2 controlled Vimentin expression and ECM degradation via TGFβ-dependent and -independent pathways, respectively. Components of this pathway were associated with poor prognosis of breast cancer patients. These findings establish an integrin-based mechanism linking ECM sensing to mesenchymal reprogramming and collective invasion.

Results

A collagen-specific basal-like cell population expresses mesenchymal and epithelial features

To determine which mesenchymal traits are specifically acquired by leader cells in collagen I environments, we expanded our previously published single-cell RNA-sequencing analysis on MMTV-PyMT breast cancer organoids. This model recapitulates key aspects of IDC, including basal-guided collective invasion and lung metastasis[3,31]. When cultured in type I collagen, basal-like cells in these organoids acquire leader cell behavior, whereas in basement membrane extract (BME), they remain non-invasive. While our prior clustering distinguished basal- and luminal-like cell states, it did not resolve functionally distinct subsets within the basal compartment[7].

In order to zoom-in on possible subpopulations of basal-like cells, we clustered the data with a higher resolution parameter. This analysis revealed 5 distinct cell clusters, including a breakdown of basal cells into two transcriptionally distinct subpopulations: a basal 1 cluster present in both BME and collagen I conditions, and a basal 2 cluster predominantly found in collagen I (Figure 1A, B). In addition to basal-like cells, we identified luminal-like populations, both cycling and non-cycling, and a distinct cluster enriched for G protein-coupled receptor kinase 4 (*Grk4*) and midline 1 (*Mid1*). Gene ontology analysis using the hallmark gene sets identified EMT as the top enriched gene set in basal 2 cells compared to basal 1 cells (Figure 1C). In addition to EMT, basal 2 cells showed enrichment for hypoxia, inflammatory response, and TNFα–NFκB signaling pathways, all previously associated with invasive front states in breast cancer[32,33]. By contrast, basal 1 cells were enriched for oxidative phosphorylation and MYC target genes indicative of a metabolically active and transcriptionally proliferative phenotype[34,35].

Since cytoskeletal protrusions are a key determinant of leader cell function, we analyzed cytoskeletal-related genes within the hallmark EMT gene set. Among EMT-associated cytoskeletal genes, vimentin (intermediate filament; Vim) emerged as the most highly upregulated gene in basal 2 cells, followed by tropomyosin 4 (Tpm4), tropomyosin 2 (Tpm2), capping actin protein (Capg), filamin A (Flna), and Tpm1 (Figure S1A). These genes encode key regulators of actin filament stabilization (Tpm4, Tpm2), barbed-end capping (Capg), and actin crosslinking (Flna). Expression distribution plots revealed a strong rightward shift in Vim expression in basal 2 cells, in contrast to the bimodal distribution observed in basal 1 and non-basal populations, which comprised a dominant Vim-low or -negative fraction and a smaller subset of intermediate expressors (Figure 1D). This pattern indicates that Vim is not only upregulated but also uniformly expressed across the collagen I–specific basal 2 cluster. Given its magnitude of induction, uniform expression, and established role as a canonical EMT effector, vimentin served as a representative mesenchymal gene to interrogate how ECM-driven reprogramming enables leader cell function during collective invasion.

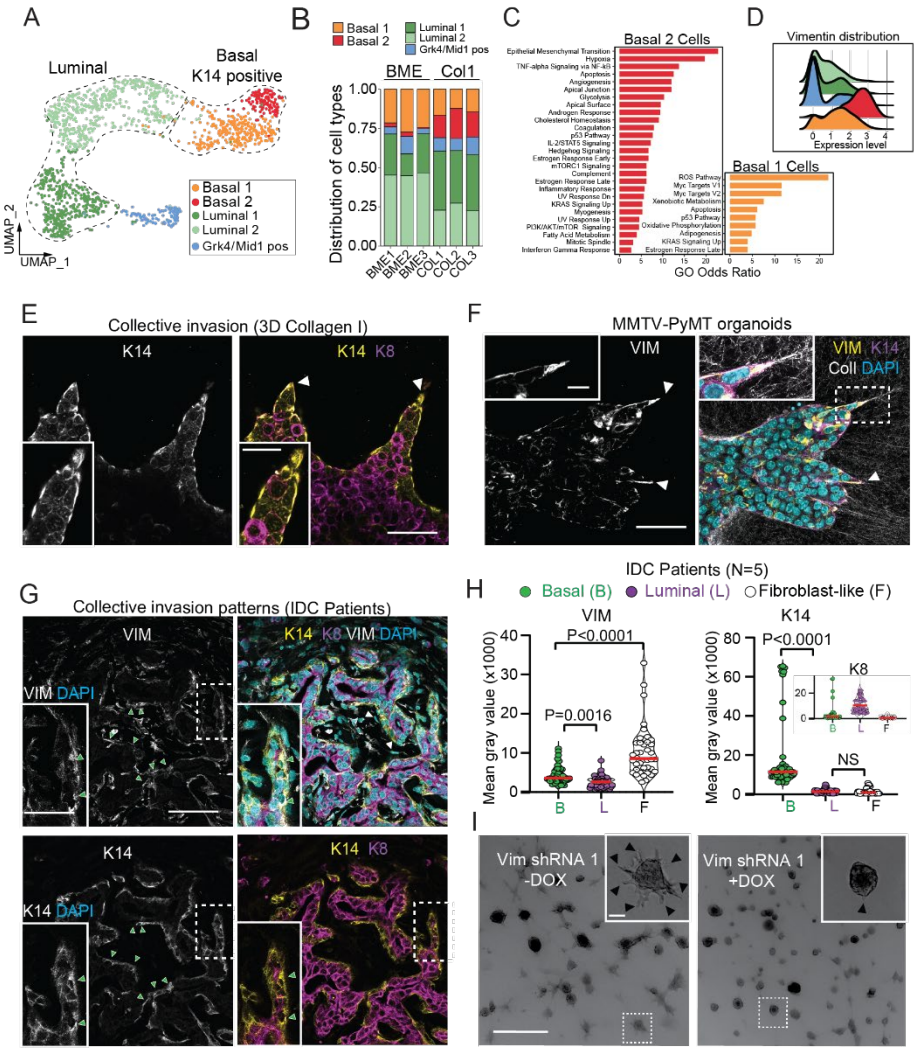


Figure 1: Basal cell heterogeneity links partial EMT and Vimentin expression to leader cell function. (A) UMAP plots based on transcriptional profiling from single-cell mRNA sequencing of MMTV-PyMT organoids cultured in 3D BME or Collagen I (Khalil A. et al., 2024). **(A)** Distinct epithelial subpopulations were identified, including two K14-positive clusters (Basal 1 and Basal 2), Grk4/Mid1-positive cells, Luminal 1, and Luminal Cycling (Luminal 2). **(B)** Distribution of these epithelial subpopulations in organoids cultured in 3D BME (799 cells from 3 × 386-well plates) or Collagen I (627 cells from 3 × 386-well plates). **(C)** Gene set enrichment analysis (GSEA) of Basal 1 and Basal 2 clusters showing significantly enriched pathways (filtered for NES > 0.5 and P < 0.01). **(D)** Ridgeline density plot of Vim expression across the epithelial subpopulations. **(E)** Confocal imaging of K14 and K8 in MMTV-PyMT embedded in 3D Collagen I for 3 days. White arrowheads depict protrusive basal-like leader cells (K14-high, K8-low). **(F)** K14 and Vimentin immunostaining in MMTV-PyMT organoids in 3D Collagen I (reflection), arrowheads depict vimentin-rich leader cells. **(G)** Confocal imaging (maximum projection) of K14, K8, and Vimentin from a whole tissue section of a K14-positive invasive ductal carcinoma (IDC) patient (representative image of N=5 patients) with zoom in on protrusive leader cells expressing K14 and Vimentin. Arrowheads: green: basal-like cells; white: fibroblast-like cells. **(H)** Mean-gray values of K8, K14, and Vimentin in the different cell types basal-like (green, 45 cells), luminal-like (magenta, 44 cells) and fibroblast-like (white, 40 cells) from 5 IDC patients. Zoom in depicting the K8 signal in the 3 populations. **(I)** Vimentin shRNA (−/+ Dox) in MMTV-PyMT organoids (Day 3, Collagen I). Arrowheads: protrusive leader cells. Scale bars: 50 µm (E, F), 10 µm (F, zoom in), 100 µm (G), 25 µm (G, zoom in), 1000 µm (I), 10 µm (I, zoom in). P values, two-sided Kruskal-Wallis test (Dunn's multiple comparison) (H).

To determine whether this mesenchymal program was accompanied by changes in cell–cell junction composition, we analyzed a curated set of genes encoding components of adherens, tight, desmosomal, and gap junctions. Basal 2 cells (collagen I-specific) maintained expression of most junctional genes at levels comparable to basal 1 and luminal populations, including key epithelial markers such as E-cadherin (Cdh1) (Figure S1B). Notably, they upregulated epithelial cell adhesion molecule (Epcam) and downregulated tight

junction protein 3 (Tjp3). Despite these modest transcriptional shifts, the overall epithelial junctional program remains preserved.

These data suggest that basal 2 cells acquire an enhanced capacity for dynamic cytoskeletal remodeling while retaining epithelial junction gene expression, consistent with a hybrid epithelial–mesenchymal phenotype that has been proposed for leader cells[36].

Vimentin is required for leader cell emergence and collective invasion

To confirm that Vimentin is expressed in leader cells during collective invasion, we examined its protein localization in both MMTV-PyMT and *Apc*^{1572T} breast cancer organoid models[37 38]. The *Apc*^{1572T} model, which harbors an *Apc* mutation that moderately activates Wnt signaling, is molecularly distinct from PyMT-driven tumors but similarly recapitulates basal cell-guided (K14^{+}) collective invasion when cultured in collagen I (Figure 1E, S1C). Confocal imaging revealed that leader cells at the invasive front consistently expressed high levels of Vimentin in both models (Figure 1F, S1D), confirming its upregulation in basal-like leader cells.

To validate this observation in human tumors, we analyzed breast cancer tissue sections from IDC patients selected for the presence of K14^{+} invasive fronts (N = 5) using confocal microscopy (Figure 1G, S1E). As previously described[3,7], and similar to the invasion pattern observed in organoids, K14-high, K8-low basal-like cells surrounded K8-high, K14-low luminal-like cells at the tumor-stroma interface (Figure 1G, green arrowheads). In situ immunofluorescence analysis (Figure 1H) revealed that basal-like cells expressed markedly higher levels of Vimentin compared to adjacent luminal cancer cells, and lower than the surrounding fibroblast-like stromal cells (Figure 1G, S1E, white arrowheads). These findings demonstrate that Vimentin is upregulated in basal-like leader cells during collective invasion in both mouse and human tumors.

Next, we tested whether Vimentin is necessary for basal-like cells to develop into leader cells, we generated doxycycline (Dox)-inducible shRNA knockdown organoids targeting Vimentin. Dox treatment reduced Vimentin expression by >75% (shRNA1) and >65% (shRNA2) in MMTV-PyMT organoids (Figure S1F), and >95% in *Apc*^{1572T} organoids (Figure S1G). Vimentin knockdown led to a marked reduction in invasion, evidenced by the decreases in both the number of leader cells per organoid (Figure 1I, S1H-K) S1G) and the length of invasive strands (Figure S1L, M). In MMTV-PyMT organoids, Vimentin knockdown also reduced organoid body size (Figure S1N), an effect not observed in *Apc*^{1572T} organoids (Figure S1O), suggesting a model-specific contribution to tumor growth. Together, these findings demonstrate that Vimentin is essential for the emergence of basal-like leader cells and for the extension of invasive strands, establishing it as a critical mesenchymal effector of collective invasion.

Collagen I degradation is essential for leader cell emergence and collective invasion

In addition to cytoskeletal regulation, ECM degradation is a key mesenchymal trait implicated in invasive cell behavior. To determine whether basal-like cells exhibit enhanced ECM degradation potential, we analyzed single-cell RNA-sequencing data using a selected gene set encompassing major classes of ECM-degrading enzymes, including matrix metalloproteinases (MMPs), ADAM and ADAMTS proteases,

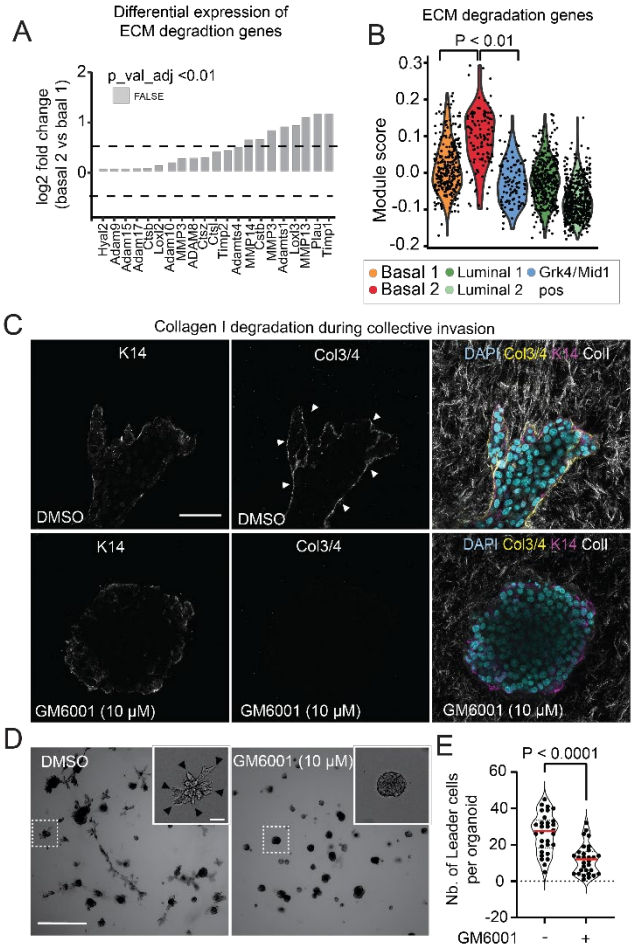


Figure 2: ECM-degrading activity in basal-like cells is required for leader cell function in 3D collagen I. (A) Log_2 fold changes in the expression of ECM-degrading and remodeling genes between Basal 2 and Basal 1 cells from scRNA-seq of MMTV-PyMT organoids. The ECM degradation gene set comprises matrix-modifying and proteolytic enzymes, including matrix metalloproteinases (MMPs), ADAM and ADAMTS family members, cathepsins, serine proteases, hyaluronidases, lysyl oxidases, and associated regulators (e.g., TIMPs, serpins). None of the observed differences reached statistical significance (adjusted P ≥ 0.01). **(B)** Module scores representing the aggregate expression of the ECM degradation gene set in individual cells. Scores reflect relative enrichment of ECM remodeling programs in basal 2 cells compared to the other subpopulations. P values, one-way ANOVA with Tukey test for multiple comparisons**. (C)** Confocal imaging of K14 and Col ¾ (degraded Collagen I) in MMTV-PyMT organoids treated with GM6001 or DMSO control for 3 days in Collagen I. Arrowheads: pericellular signal of Col ¾ adjacent to basal-like cells (K14). **(D)** Representative brightfield images of MMTV-PyMT organoids in Collagen I treated with GM6001 or DMSO control for three days. Arrowheads: invasive strands. **(E)** Number of leader cells per organoid, Median (red line), n = 30 organoids per condition from three independent experiments. Scale bars: 50 µm (C), 1000 µm (D), inserts 10 µm (D, zoom in). P values, two-sided unpaired Mann–Whitney test (E).

cathepsins (Ctsl), hyaluronidases, serine proteases (incl. PLAU, ELANE) and their regulators (TIMPs, SERPINs, cystatin). No individual gene from this ECM-degradation signature showed statistically significant enrichment in the collagen I-specific basal 2 population when analyzed separately. Among the matrix metalloproteinases, Mmp3, Mmp13, and Mmp14 were expressed across clusters, with a trend toward higher expression in basal 2 cells (Figure 2A). This lack of a dominant gene is consistent with the dispersed expression patterns typical of ECM-degrading enzymes, where functional activity often reflects contributions from multiple genes acting in parallel[39]. Indeed, gene set enrichment analysis, which captures such coordinated expression patterns, confirmed that basal 2 cells were significantly enriched for the ECM-

degradation signature compared to basal 1 and non-basal cells (Figure 2B). These findings indicate that the transcriptional program of collagen I-specific basal-like cells includes ECM-remodeling genes, providing a molecular link between matrix degradation and the leader cell state during collective invasion.

To functionally validate this transcriptional signature, we performed immunostaining for the collagen ¾ (Col¾) epitope, a fragment generated by MMP-mediated cleavage of collagen I[40–42]. Confocal microscopy of MMTV-PyMT organoids cultured in collagen I for 3 days revealed strong Col¾ accumulation at the tumor–ECM interface, including the fronts and lateral edges of K14^{+} strands (Figure 2C, arrowheads). Treatment with GM6001, a broad-spectrum MMP inhibitor, markedly reduced Col¾ signal compared to controls (Figure 2C). Pan-MMP inhibition also significantly impaired collective invasion in both MMTV-PyMT and *Apc*^{1572T} organoids, reducing leader cell numbers and strand elongation by more than 50% (Figure 2D, E, S2A-D), without affecting organoid body size (Figure S2E-F). These findings establish that collagen I degradation by basal-like leader cells is MMP-dependent and is required for both their emergence and sustained collective invasion in 3D collagen-rich environments.

TGFβ signaling promotes Vimentin expression and collective invasion.

Given the mesenchymal signature of basal-like leader cells and the established role of TGFβ signalling as a major regulator of EMT, we next investigated whether TGFβ signalling contributes to the gene expression program in these cells. TGFβ signaling, a major inducer of EMT, is initiated by ligand binding to type II receptors, which activate type I receptors, including ALK4, ALK5, and ALK7, thereby triggering downstream SMAD-dependent transcriptional responses[43,44]. Gene set enrichment analysis (KEGG) indicated elevated TGFβ pathway activity in the collagen I-specific basal 2 population (Figure S3A). We further analyzed the expression of a selected set of canonical TGFβ target genes, including *CTGF*, *PAI1*, TGFBi, and *Smad7*, and found this gene set to be also strongly upregulated in basal 2 cells compared to basal 1 subset and the other populations (Figure 3A), suggesting active TGFβ signaling in leader cells within collagen-rich environments. To test whether TGFβ regulates mesenchymal gene expression, we treated organoids cultured in collagen I with SB431542 (5 µM), a selective inhibitor of TGFβ type I receptor kinases (ALK4, ALK5 and ALK7)[45,46] and analyzed gene expression by qPCR. SB431542 treatment significantly reduced the expression of canonical TGFβ target genes, including *CTGF* (2-fold; SD = 0.13) and *Smad7* (6.1-fold), confirming effective pathway inhibition. In MMTV-PyMT organoids, the expression of mesenchymal markers *Vimentin* (2.2-fold) and *Slug* (2.3-fold) was also reduced, while *Snail1* mRNA levels remained unchanged. Furthermore, *Cdh1* (E-cadherin) expression was unaffected (Figure 3B), indicating that TGFβ promotes mesenchymal gene expression without repressing core epithelial junction components. In *Apc*^{1572T} organoids, SB431542 similarly reduced *Vimentin* (3.6-fold) and *CTGF* (10.5-fold), while changes in *Snail1* and *Smad7* were not statistically significant (Figure S3B).

Interestingly, TGFβ inhibition did not impair ECM degradation. Immunofluorescence for the collagen ¾ (Col¾) epitope showed comparable levels in SB431542-treated and control organoids (Figure 3C), suggesting that matrix degradation is TGFβ-independent. These findings were corroborated using A83-01 (0.5 µM), a more potent ALK inhibitor (Figure S3C, D).

Functionally, both SB431542 and A83-01 significantly reduced leader cell formation and invasive strand length in MMTV-PyMT (Figure 3C, D; S3E) and *Apc*^{1572T} (Figure 3E) organoids. To test whether this

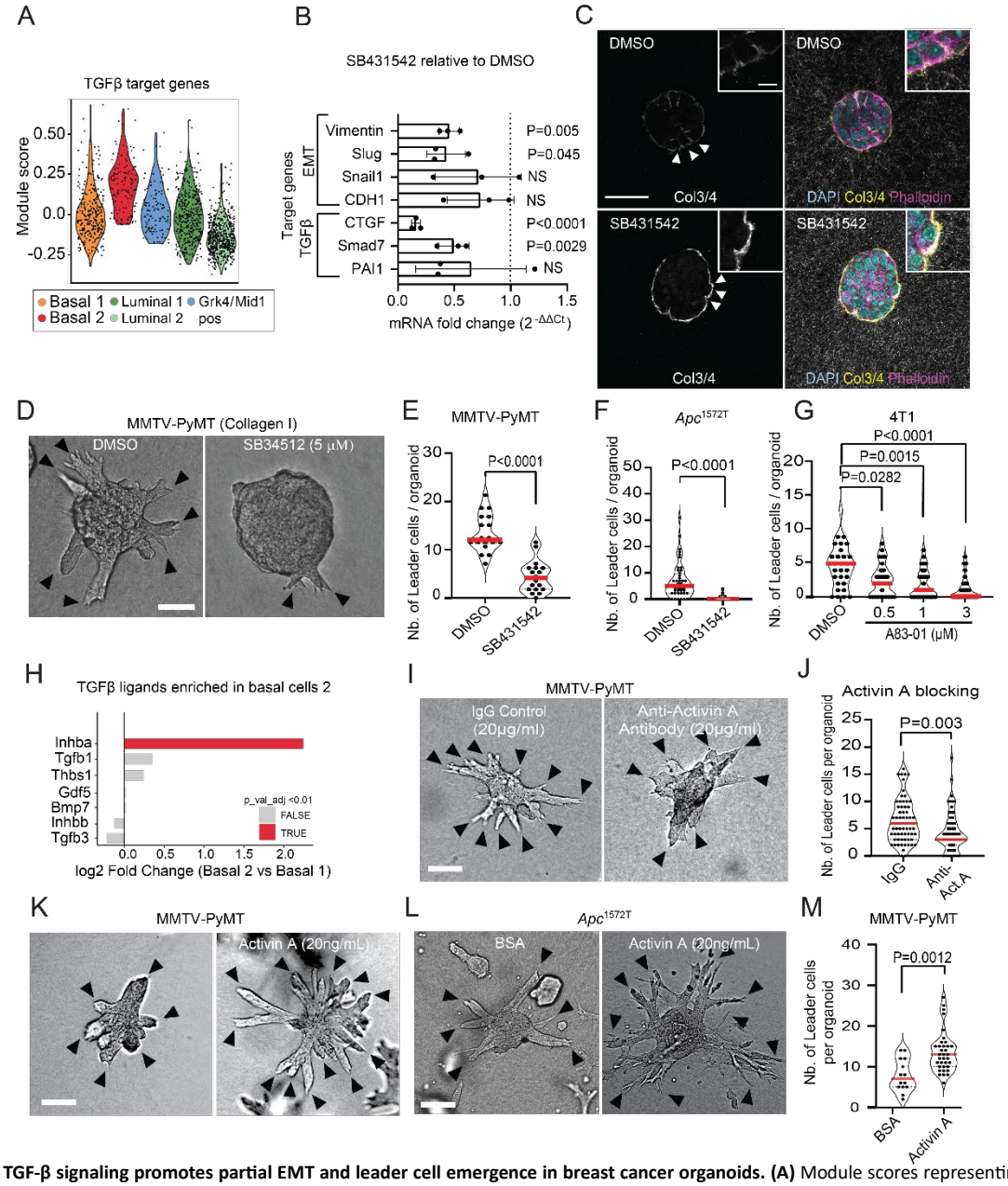


Figure 3: TGF-β signaling promotes partial EMT and leader cell emergence in breast cancer organoids. (A) Module scores representing the aggregate expression of TGF-β target genes in individual cells, reflecting relative pathway enrichment in basal 2 cells compared to the other subpopulations. P values: one-way ANOVA with Tukey's test for multiple comparisons. **(B)** qPCR analysis of classical TGF-β and EMT target genes in MMTV-PyMT organoids cultured in 3D Collagen I for three days. Values represent mean normalized mRNA expression (relative to housekeeping genes) in organoids treated with 5 µM SB431542, shown relative to DMSO-treated controls (dashed line). Data represent mean ± SD from three independent experiments. **(C)** Confocal imaging of Col ¾ (degraded Collagen I) and F-actin in MMTV-PyMT organoids cultured in 3D Collagen I and treated with SB431542 (5 µM) or DMSO control for one day. Arrowheads: pericellular Col ¾ signal. Inset: organoid-Collagen I interface. **(D)** Representative brightfield images of MMTV-PyMT organoids cultured in 3D Collagen I and treated with SB431542 or DMSO control for three days. **(E–G)** Quantification of leader cells per organoid in **(E)** MMTV-PyMT organoids, **(F)** *Apc*^{1572T} organoids treated with 5 µM SB431542 or DMSO control, and **(G)** 4T1 cancer spheroids treated with a dose series of A8301 (0.5 µM, 1 µM, 3 µM). Number of organoids analyzed per condition: **(E)** 18 organoids (N = 3), **(F)** 44 organoids (N = 3), **(G)** 30 organoids (N = 3). Red lines indicate the median. **(H)** Log_2 fold change of TGFβ ligands in the Basal 2 versus Basal 1 subset, based on scRNA-seq data from MMTV-PyMT organoids. The gene set includes ligands from all major TGFβ superfamily branches, including canonical TGFβs, Activins/Inhibins, BMPs, Nodal-related ligands, and GDFs. **(I)** Brightfield images of MMTV-PyMT organoids treated with an Activin A neutralizing antibody (20 µg/ml) or IgG control (20 µg/ml) for two days in 3D Collagen I. **(J)** Quantification of leader cell numbers per organoid under Activin A neutralization or IgG control conditions. Red lines indicate the median; n = 60 organoids per condition from three independent experiments. **(K, L)** Brightfield images of **(K)** MMTV-PyMT and **(L)** *Apc*^{1572T} organoids treated with Activin A ligand (20 ng/µl) or vehicle control (0.1% BSA) for three days in 3D Collagen I. Arrowheads (D, I, K, L): leader cells. **(M)** Quantification of leader cells per organoid in MMTV-PyMT organoids. Red line indicates the median; n = 14–36 organoids per condition from three independent experiments. Scale bars: 50 µm (C), 10 µm (C, zoom-in), 100 µm (D, I, K, L). P values: two-sided unpaired Mann–Whitney test (A, E, F, J, M); two-sided Kruskal-Wallis test with Dunn's multiple comparisons (G).

requirement for TGFβ signaling extends to other breast cancer models, we examined 4T1 spheroids, a metastatic triple-negative murine cell line known to invade collectively in 3D collagen I[47]. TGFβ inhibition in 4T1 spheroids similarly impaired leader cell emergence (Figure 3F; S3F, G). While TGFβ inhibition did not affect the overall size of MMTV-PyMT organoids or 4T1 spheroids (Figure S3H-K), it led to a reduction in *Apc*^{1572T} organoid size compared to DMSO-treated controls (Figure S2I), suggesting that TGFβ may have broader roles in tumor growth in this model (Figure S3L).

Our results demonstrate that TGFβ signaling is essential for leader cell formation and the acquisition of several EMT-associated features, including Vimentin expression, but is dispensable for matrix degradation. To identify the ligand responsible for activating TGFβ signaling in basal-like cells within collagen I environments, we analyzed the expression of a curated panel of TGFβ superfamily ligands, including canonical TGFB1-3, INHBA-E, BMP2–15, GDF1-15, NODAL, LEFTY1/2, AMH, THBS1. Among these, Inhba was the most strongly upregulated gene in basal 2 cells compared to basal 1 (Figure 3G). Inhba encodes the βA subunit of the homodimer Activin A, a functional TGFβ ligand that activates SMAD-dependent transcriptional programs[48,49]. To test whether Activin A contributes functionally to collective invasion, we neutralized endogenous Activin A using a blocking antibody, which significantly reduced leader cell numbers per organoid compared to IgG-treated controls (Figure 3H, arrowheads; Figure 3I). Conversely, supplementation with exogenous Activin A (20 ng/ml) strongly increased leader cell emergence in MMTV-PyMT organoids (1.75-fold), *Apc*^{1572T} organoids (2-fold), and 4T1 spheroids (1.8-fold) relative to BSA-treated controls (Figure 3J–L; Figure S3M, N).

These findings establish *Inhba* as a collagen I-induced ligand that activates TGFβ signaling via Activin A, thereby promoting leader cell emergence and collective invasion in 3D collagen I.

Collagen I-Itgα2β1 interactions promote mesenchymal reprogramming and TGFβ signaling via Inhba

Our data demonstrate that leader cells in collagen I–rich environments depend on mesenchymal traits, including Vimentin expression and ECM degradation. To test whether these features are driven by direct interaction with the extracellular matrix, we examined the role of integrin α2 (Itgα2), the collagen-binding subunit of the Itgα2β1 heterodimer. Itgα2β1 forms the integrin complex with the highest affinity for fibrillar collagen I[50–54]. Immunostaining of IDC patient samples revealed strong Itgα2 expression in basal-like cells (K8 low) forming tumor-stroma interface, in contrast to lower levels of Itgα2 in luminal tumor cells (K8 high) and fibroblast-like cells (Figure 4A, B), suggesting that Itgα2β1 mediates tumor–stroma interactions during collective invasion. CRISPR/Cas9-mediated knockout of Itgα2 in MMTV-PyMT organoids (Figure S4A, B) abolished collective invasion in collagen I (Figure 4C, D; S4D), preventing basal-like cells from protruding and forming invasive strands (Figure S4C). Loss of Itgα2 also markedly reduced Vimentin expression (Figure 4E) and impaired collagen degradation (Figure 4F, S4E), demonstrating that collagen I–Itgα2β1 engagement is required to activate mesenchymal features in basal-like leader cells and to sustain collective invasion.

Given the critical roles of both Itgα2 and TGFβ signaling in mesenchymal leader cell identity, we next asked whether these pathways intersect. qPCR analysis revealed that Itgα2 knockout significantly reduced the expression of TGFβ target genes, including CTGF, PAI1, and Smad7, in collagen I (Figure 4E), indicating that Itgα2 promotes TGFβ pathway activation in a collagen-dependent manner. We then tested whether Inhba, the gene encoding the Activin A subunit, functions downstream of Itgα2. Inhba was significantly downregulated in Itgα2-deficient organoids at both transcript and protein levels (Figure 4E, S4F, G),

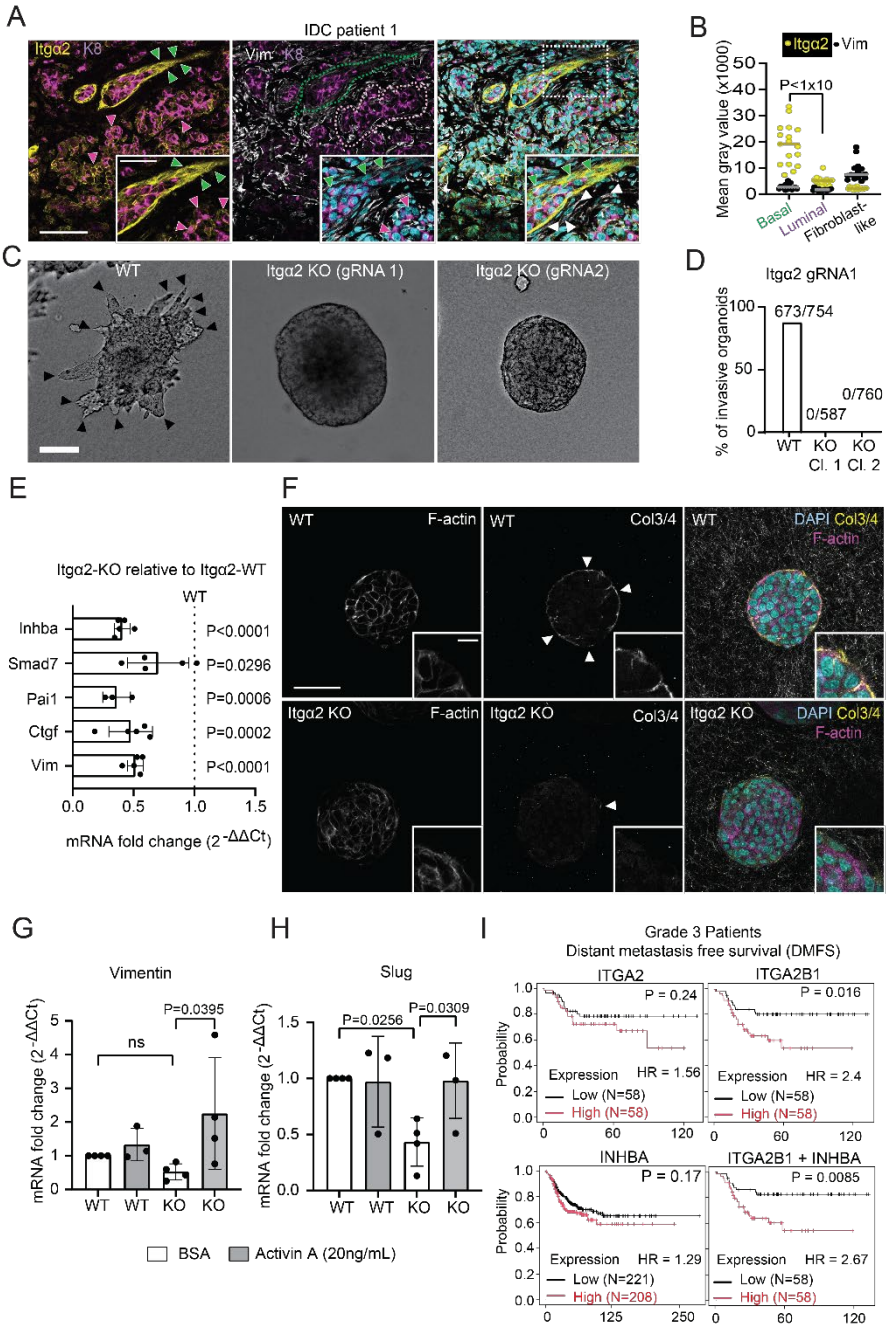


Figure 4: Integrin α2 promotes mesenchymal traits in basal-like cells through TGFβ-dependent and -independent pathways. (A) Confocal imaging of ITGα2, Vimentin, and K8 in an IDC patient tissue section. Green contour: cluster of cancer cells with protrusive morphology; green arrowheads: basal-like cells (K8-low) at the tumor–stroma interface with high ITGα2 expression. White contour: cluster of cancer cells lacking basal-like cells (K8-high), with low ITGα2 expression; magenta arrowheads. White arrowheads: fibroblast-like cells (elongated, spindle-shaped). **(B)** Quantification of mean gray values for ITGα2 and Vimentin in basal-like (n = 33), luminal-like (n = 32), and fibroblast-like (n = 32) cells from one IDC patient tissue section. **(C)** Representative brightfield images of MMTV-PyMT organoids (ITGα2-WT or ITGα2-KO, gRNA1 and gRNA2) cultured in 3D Collagen I. Black arrowheads: invasive strands. **(D)** Percentage of organoids exhibiting one or more invasive strands in ITGα2-WT and ITGα2-KO (clones 1 and 2 from gRNA1) MMTV-PyMT organoids. **(E)** qPCR analysis of classical TGF-β and EMT target genes in ITGα2-WT and ITGα2-KO MMTV-PyMT organoids cultured in 3D Collagen I for three days. Values represent mean normalized mRNA expression (relative to housekeeping genes), shown for KO organoids relative to WT controls (dashed line). Data are presented as mean ± SD from three independent experiments. **(F)** Confocal imaging of Col ¾ and F-actin in ITGα2-WT and ITGα2-KO MMTV-PyMT organoids after one day in 3D Collagen I. **(G, H)** qPCR analysis of Vimentin and Slug mRNA expression in ITGα2-KO versus ITGα2-WT MMTV-PyMT organoids treated with Activin A (20 ng/µl) or vehicle control (0.1% BSA) for three days. Bar graphs represent mean normalized expression values ± SD from four independent experiments. **(I)** Kaplan–Meier analysis correlating high vs. low mRNA expression of INHBA, ITGA2, ITGB1, and their combinations (ITGA2 + ITGB1, or INHBA + ITGA2 + ITGB1) with distant metastasis-free survival (DMFS) in patients with grade 3 breast cancer. Scale bars: 100 µm (A, C), 50 µm (A, zoom-in), 50 µm (F), 10 µm (F, zoom-in). P values: two-sided unpaired Mann–Whitney test (E), two-sided Kruskal-Wallis test with Dunn's multiple comparisons (G, H), Log-rank test (I).

suggesting that collagen I-Itgα2 signaling enhances TGFβ activation via Inhba upregulation. To directly test this link, we treated Itgα2-deficient organoids embedded in 3D Collagen I with exogenous Activin A.

Activin A restored the expression of the TGFβ target gene CTGF and the mesenchymal genes Vimentin and Slug, even in the absence of Itgα2 (Figure 4G, H; S4H). These findings demonstrate that Itgα2-induced TGFβ signaling and EMT is, at least in part, mediated through Inhba upregulation.

To assess the clinical relevance of the Itgα2β1-Inhba signaling axis in breast cancer progression, we analyzed associations between gene expression and distant metastasis-free survival (DMFS) using the Györffy cohort of grade 3 breast cancer patients[55]. Individually, ITGA2 and ITGB1 expression levels showed no significant association with outcome (Figure 4I, S4I). However, their combined high expression, suggestive of increased integrin α2β1 receptor activity, was significantly associated with reduced metastasis-free survival (hazard ratio = 2.4, $P = 0.016$). Similarly, INHBA expression alone did not associate with DMFS, but when co-expressed with high ITGA2B1, the association with poor outcome was further strengthened (hazard ratio = 2.67, $P = 0.0085$) (Figure 4I). These data suggest that concurrent high expression of ITGA2, ITGB1, and INHBA defines a patient subgroup with markedly worse metastatic outcomes, highlighting the clinical relevance of the Itgα2β1-Inhba signaling axis.

Summarized, our findings demonstrate that collagen I engagement via the Itgα2β1 receptor induces upregulation of the TGFβ ligand Inhba, thereby activating the TGFβ pathway and promoting mesenchymal reprogramming in leader cells. This integrin- autocrine/paracrine cascade links ECM sensing to collective invasion, establishing collagen I–integrin signaling as a central driver of metastatic progression and a potential therapeutic target in aggressive breast cancer.

Discussion

Our study identifies a Collagen I-Itgα2-driven program through which leader cells acquire selective mesenchymal traits required for collective invasion. We show that integrin α2 (Itgα2) links collagen I sensing to Vimentin expression and matrix degradation, establishing a feedback loop that sustains leader cell function. This program enables basal-like cells to invade while maintaining epithelial cohesion, representing a hybrid epithelial–mesenchymal state distinct from full EMT. These findings reveal how ECM engagement and autocrine/paracrine signaling converge to generate a progressively evolving microenvironment that supports and drives collective invasion in aggressive breast cancers.

Using transcriptional profiling and functional assays in breast cancer organoids, we identify a collagen I-specific subset of basal-like cells that activate a mesenchymal gene program, including Vimentin and ECM-degrading enzymes, while retaining epithelial junction components. We define this subset as leader cells based on four converging lines of evidence: (i) they emerge selectively in collagen I, a matrix permissive for invasion; (ii) their molecular features are enriched in front-localized cells within invasive strands in organoids and patient tumors; (iii) they selectively upregulate invasion-enabling genes (e.g., *Vim*, *Tpm2*, *Flna*, ECM degradation genes); and (iv) perturbing key effectors upregulated in these cells, including Vimentin and MMP-mediated degradation, impairs leader cell formation and collective invasion.

The selective mesenchymal program equips basal-like cell with the functional capacities required for them to initiate and lead collective invasion. Vimentin, uniformly upregulated in the collagen I-specific subset, reinforces actin-rich protrusions, may enhance traction force generation, and protects against mechanical

stress during migration[27–29]. Vimentin positive leader cells were recapitulated in IDC patient tissue, where positive basal-like cells (high K14-lowK8) at the tumor-stroma interface exhibit elongated, protrusive-like morphologies and upregulate Vimentin. Despite the limitations of 2D histological analysis, these features align with the invasive front-localized phenotype observed in 3D collagen I organoids, reinforcing the relevance of basal-like leader cells in human breast tumors[56,57]. In addition to Vimentin, cytoskeletal regulators such as *Tpm2*, *Tpm4*, *Capg*, and *Flna* were enriched in the leader basal-like cell subset. These genes are known to promote actin filament stability, contractility, and mechanotransduction, potentially supporting front–rear polarity and force transmission at the leading edge[58–60]. In total, this cytoskeletal molecular repertoire is essential for directional migration through mechanically resistant environments. To overcome such barriers, our work shows that leader cells exhibit enrichment of an ECM-degradation gene signature, including *Timp1*, *Plau*, *MMP3*, *MMP13*, and *MMP14*, as well as cathepsins. This repertoire supports focal collagen I degradation, as confirmed by inhibition of both collagen proteolysis and invasion upon MMP blockade. Beyond proteolysis, leader cells express ECM-components and -modifying genes such as *FN1*, *TNC*, and *LOXL3*, which may contribute to broader matrix remodeling and formation of invasion-permissive paths [61–64].

Notably, despite their mesenchymal activity, leader cells retain the expression of epithelial junction genes, including *Epcam*, *Cdh1* (E-cadherin), *Tjp2*, and *Cldn12*. An in-depth analysis of the organization and function of these junctional proteins remains to be performed. However, based on transcriptomic data and the observation that leader cells maintain adhesion to follower cells, the evidence supports the notion that these junctions enable mechanical coupling and coordinated collective movement. This junctional integrity may facilitate force transmission across the invasive strand, establishing front–rear asymmetry and supporting persistent migration as a cohesive unit[65,66]. Such a hybrid epithelial–mesenchymal phenotype enables leader cells to remodel their environment without detachment and may confer survival advantages, including resistance to cell death and therapy, frequently observed at invasive tumor fronts[56,57 67–69]. In contrast, single epithelial cells undergoing EMT often lose polarity and adhesion, rendering them less effective at sustained invasion and more vulnerable to anoikis and mechanical stress[6,47]. Together, these findings define a hybrid invasion program that integrates mesenchymal dynamics with epithelial architecture, reinforcing the central role of leader cells in collective tumor progression.

While collagen I is known to promote protrusive activity in cancer cells, its role in driving mesenchymal characteristics has remained unclear especially in collective invasion. Our data demonstrate that collagen I-induced leader cell function requires Itgα2, which mediates high-affinity adhesion to fibrillar collagen I[12]. We find that Itgα2 is essential for mesenchymal reprogramming, including Vimentin expression and MMP-mediated ECM degradation. Mechanistically, Itgα2 acts as an upstream regulator of TGFβ signaling, a well-established driver of EMT[70,71]. Multiple lines of evidence support this: (i) *Inhba* is selectively upregulated in basal-like cells exposed to collagen I; (ii) *Itgα2* knockout reduces *Inhba* expression, which encodes Activin A, a TGFβ ligand required for leader cell emergence; (iii) *Itgα2* loss impairs TGFβ pathway activity, as shown by reduced *Smad7* and *Pai1* expression and downregulation of mesenchymal markers such as Slug and Vimentin; and (iv) exogenous Activin A restores TGFβ signaling and mesenchymal gene expression in *Itgα2*-deficient cells. These findings support a model in which collagen I engagement activates mesenchymal programs via Itgα2-dependent transcriptional induction of *Inhba*, integrating mechanical and transcriptional signals to sustain leader cell identity.

Our data further establish Activin A as a key effector of collective invasion. While supplementation with exogenous Activin A promotes leader cell emergence across multiple basal-like breast cancer models, it

fails to rescue invasion in *Itgα2*-deficient organoids. This suggests that although Activin A can support mesenchymal traits, Itgα2 is indispensable for invasion, likely due to the dual role of α2β1 integrin in collagen adhesion and signal transduction. The clinical relevance of this Integrin-INHBA axis is supported by our meta-analysis showing that co-expression of *INHBA*, *ITGA2*, and *ITGB1* is associated with reduced distant metastasis-free survival in grade 3 breast cancer patients. The transcriptional regulation of *INHBA* by Itgα2 remains to be fully defined, but may involve YAP, a mechanoresponsive transcriptional co-activator active in leader cells and previously shown to control *INHBA* expression[72].

Beyond Itgα2 and Activin A, other integrin-dependent mechanisms likely contribute to ECM-mediated TGFβ activation and EMT[73,74]. Latent TGFβ ligands stored in the ECM require mechanical activation, a process mediated by integrins such as αvβ6 and αvβ8 that apply tension to the latency-associated peptide (LAP) to release active TGFβ[73]. Traction forces generated by leader cells may further promote this process, linking ECM remodeling to the reinforcement of mesenchymal traits. Notably, while ECM degradation is independent of TGFβ signaling, it remains dependent on Itgα2, suggesting alternative regulatory axes. Whether Itgα2 modulates MMP activity at the transcriptional or post-transcriptional level remains to be determined[75].

Our findings demonstrate that basal-like leader cells inhabit a dynamic niche where ECM engagement and autocrine/paracrine signaling cooperate to sustain mesenchymal traits and collective invasion. Rather than acting independently, cytoskeletal remodeling, ECM degradation, and junctional cohesion form an integrated network of feedback loops that reinforce leader cell function. This self-sustaining invasion program continuously remodels the tumor-stroma interface and enhances cellular plasticity and metastatic potential. Targeting the Itgα2-TGFβ signaling axis may therefore represent a therapeutic opportunity to disrupt leader cell–driven invasion in aggressive epithelial cancers.

Materials & Methods

Antibodies and reagents

The following antibodies were used: rabbit anti-human Vimentin (Cat#ab92547, Abcam), chicken anti-human Vimentin (Cat#PA1-16759, Invitrogen) rabbit anti-human Keratin 14 (Cat# 905301, Biolegend), rat anti-mouse Keratin 8 (Cat# 531826, DSHB), rabbit anti-rat Collagen I cleavage site (Col ¾, Cat#0217-025, immunoGlobe), rabbit anti-human integrin α2 (Cat#ab181548, Abcam), rabbit anti-human Inhba (Cat#10651-1-AP, Proteintech), Mouse IgG1 isotype control (MAB002, R&D Systems), anti-Activin A antibody (Cat#AF338, R&D Systems). The following secondary antibodies were used: Alexa fluor-488/568/647-conjugated goat anti-mouse, -rat, or -rabbit antibody (Invitrogen). To visualize the nucleus and F-actin, 4′,6-diamidino-2-phenylindole (DAPI, Cat#D9564, Sigma) and Alexa-fluor-568-conjugated phalloidin (ThermoFisher) were used, respectively.

The following reagents were used: Ilomastat/GM6001 (Cat#2983, TOCRIS), SB431542 (Cat#616461, Merck), A83-01 (Cat#2939/10, Bio-Techne), Activin A (Cat#AF338, R&D Systems), BSA (Cat#A9647, MERCK), and Dimethyl Sulfoxide (Cat# 5.89569, MERCK).

Organoid culture

The MMTV-PyMT and *Apc*^{1572T} organoid lines were generous gifts from Jacco van Rheenen[76] and R.F.[37], respectively. The 4T1 metastatic breast cancer cell line was used (CRL-2539 ATCC). All cells (organoids and 4T1) were cultured in 20-25µl drops of Cultrex Reduced Growth Factor Basement Membrane Extract (BME) (Cat# 3533-005-02, RGF type 2 Path Clear, Bio-Techne) as previously described[77]. Similar to the organoids (MMTV-PyMT, *Apc*^{1572T}), the 4T1 cells grew in BME into rounded multicellular structures (spheroids). The growth medium of MMTV-PyMT organoid line and 4T1 spheroids was composed of Dulbecco's Modified Eagle Medium F12 (Cat# 12634010, Gibco), supplemented with 10uM HEPES (Cat#15630080, Gibco), 10 U/ml Pen-Strep (Cat#17-602E, Lonza), 1% GlutaMAX 100X (Cat# 35050061, Life Technologies), 125uM N-Acetyl-L-Cysteine (Cat#A9165, Sigma), 0.01% Primocin (Cat#ant-pm-2, Bio-Connect), 1x B27 (Cat#17504044, Life Technologies), and 2.5 nM human FGF2 (42,5 ng/ml)(Cat# 100-18B, PeproTech). *Apc*^{1572T} organoids were grown in Dulbecco's Modified Eagle Medium F12 supplemented with 10 U/ml Pen-Strep, 1x B27, 0.01% Primocin, 1.16 nM FGF2 (20ng/ml), and 3.2nM EGF (20ng/ml, Cat#AF-100-15, PeproTech).

Cultures were maintained at 37°C, 5% CO2, and humidified atmosphere. They were passaged approximately twice a week by dissociating into single cells using trypsin-EDTA (TrypLE Express, Cat# 12605-010) for a maximum duration of 5 min at 37°C. The organoids and spheroids were further mechanically dissociated into single cells by repeated pipetting. The single cells were washed with DMEM/F12 (+HEPES, GlutaMAX and pen-strep) followed by centrifugation (260G, 5min. 4°C). Cell pellets were suspended in BME. After polymerization, the gels were supplemented with culture media.

Collagen I Invasion Assays

Confluent BME cultures containing medium-sized breast cancer organoids (MMTV-PyMT and *Apc*^{1572T}) and spheroids (4T1) were collected by resuspending with cold media (DMEM/F12 supplemented with HEPES, Pen-Strep, and 1% GlutaMax). Organoids were spun down for 4 minutes at 260 G (4°C), and supernatant was discarded. To remove remaining BME, the pellet was incubated with Dispase 1mg/ml (Cat#17105-041, Life Technology) for 15-minute incubation at 37°C water bath. This was followed with 4 washing rounds. Next, the organoids were embedded in 40 µl drops of Collagen I (2 mg/ml) or BME gels in 24 or 48 -well plates. Non-pepsinized rat-tail Collagen (Cat#354236, Corning) was prepared at a final concentration of 2 mg/ml using 10X phosphate buffer saline (PBS) (Gibco), NaOH (final pH of 7.5) and dH2O according to the manufacturer's instructions. Collagen I was pre-polymerized for 2 hours on ice, then organoids were added to the Collagen mix followed by 5 min. incubation at room temperature and plating on pre-warmed plates (Costar) and incubated for 30 minutes at 37°C to polymerize[77]. Subsequently, medium was added.

Single Cell Sequencing Analysis

Previously published data from GEO (GSE197821) was analyzed in R using Seurat V4[78]. Following data integration (see details in Khalil et al. 2024[7]), cells were clustered into 5 subgroups using SNN with resolution parameter 0.3. Gene ontology analysis was performed using the EnrichR package[79] based on the hallmark gene sets from the molecular signature database[80]. For this analysis, differentially expressed genes were identified by comparing the distinct basal populations with FindMarkers, and filtering genes

using cutoffs for pvalue (<0.01) and log2FC (>0.5). Other signatures were evaluated using the AddModuleScore function, and visualized with a violin plot across the distinct cell populations or log2FC values for a two-group comparison were directly visualized as bar plots.

Brightfield Microscopy

The breast cancer organoids and spheroids were imaged using the EVOS M5000 microscope using the 2x and 10x objective (2x objective NA= 0.08, 10x objective, NA=0.4). For invasion quantification, a region of interest (ROI) was manually drawn around the organoid body, and invasive strands were traced by drawing a line from the organoid border until the tip of the strand using ImageJ (ImageJ; 1.40v; National Institutes of Health). The organoid body area (in μm^2), number of pointed tips (leader cells) and strand length (in µm) were measured[47].

Immunofluorescence and confocal imaging

Organoids and spheroids cultured in 3D Collagen I were fixed using 4% paraformaldehyde (Sigma-Aldrich) (15-minute incubation at RT), followed by three washing rounds using 1X PBS (Cat#D8537, Sigma-Aldrich). After washing, the collected gels were transferred into 1.5 mL Eppendorf tubes and blocked using 10% normal goat serum (Gibco) with 0.3% Triton-X (Sigma-Aldrich) in 1X PBS and incubated for 1 hour while shaking at RT. This was followed with primary antibody incubation in 0.1% Bovine Serum Albumin (BSA; Sigma Life Sciences) with 0.3% Triton in 1X PBS. Samples were incubated with the primary solution overnight at 4 °C with shaking. Dilutions of primary antibodies used in this study include: K14 antibody (1:400), K8 antibody (1:50), vimentin antibody (1:200), Collagen ¾ antibody (5 µg/mL), Itgα2 antibody (1:400), Inhba antibody (1:70). After removing the primary antibody solution, the samples were washed four times 15 minutes in 1X PBS. After washing, the samples were incubated with the secondary antibody solution containing Alexa-fluor 488/568/647 conjugated secondary antibodies together with DAPI (5µg/ml) and Phalloidin (1:500) overnight at 4 °C with shaking. The next day, the gels were thoroughly washed by four cycles of 10 minutes 1X PBS. For Collagen ¾, similar staining steps were performed but the blocking and antibody buffers did not contain Triton. The samples were imaged on glass bottom WillCo Wells (WillCo Well BV) using confocal microscopy (LSM 880 Zeiss) with a 40X magnification objective (NA=1.1). To analyze the effect of Itgα2 KO on Inhba expression, a single confocal z-slice from the mid-plane of WT and KO organoids was used. Using Fiji, contours were manually drawn around the cytosol and nucleus (based on Phalloidin and DAPI signals) of basal-like cells (low K8) that are located at the organoid-ECM interface. The mean gray value of Inhba intensity was measured on 8-bit images, with background correction performed by subtracting the intensities of cell-free regions. The collagen ¾ signal quantification was performed using an ImageJ macro that starts by selecting middle z-layer from the acquired confocal stack. The script then applies a threshold (adjusted manually) to the F-actin signal to create a binary mask, selecting the organoid while excluding the surrounding ECM and background. Using this mask, the macro generates an outline (line width = 15) forming the border of organoid and defining the region of interest. Collagen ¾ and Collagen I reflection signal are then measured within the outlined regions and the relative ratio of Collagen ¾ to Collagen I reflection signal is calculated.

Immunofluorescent staining of Patient IDC sample

Immunofluorescence was performed on 8 µm whole invasive breast cancer tissue sections diagnosed as high grade IDC (Table 3). The use of material was approved by the Tissue Science Committee of the University Medical Center Utrecht and Department of the Pathology of the Radbouumc, and informed consent was obtained from the participants. Tissues were deparaffinized with Xylene (2X, 2 minutes) followed by hydrating the tissues using a sequential descending percentage of ethanol (100%, 96%, 70%, and 50%; 2X, 2 minutes each). Antigen retrieval was performed using Tris-EDTA buffer (15 min, 95–100 °C) followed by immunostaining as previously described[47]. Briefly, tissue sections were blocked using 5% normal goat serum in 0.05% Tris-buffered saline (TBS) with Tween™ 20 detergent (TBST) for 1 h at room temperature followed by primary antibody incubation at 4 °C (overnight) using the following antibody dilutions: Itgα2 (1:300), vimentin (1:300), and K8 (1:50). After antibody incubation, samples were washed four times (TBST, 5 min each) and incubated with secondary Alexa Fluor 488/546/647-conjugated goat anti-mouse, anti-rabbit, anti-rat antibodies and with DAPI (1 µg/mL), for 1 h at room temperature. Samples were washed four times (TBST, 5 min each) and mounted with Immu-Mount (Epredia). Imaging was performed using confocal microscopy (LSM 880 Zeiss) with a 40x objective (NA =1.1). For the analysis of Itgα2, K8, K14, and vimentin expression in tumor cells, cancer cell groups were first identified based on their K8-positive cells organized into protrusive multicellular structures. Cells were classified as luminal-like (K8-high), basal-like (K14-high), or stromal (spindle-shaped with elongated nuclei). Using Fiji, contours were manually drawn around the cytosols of selected cells using a merged channel combining all cytoskeletal markers (K8, K14, and vimentin). The mean gray value of cytoskeletal markers and Itgα2 intensity was then measured within the selected regions of interest and background-corrected by subtracting the intensity of a cell-free region.

Generation of doxycycline inducible knockdown organoid models

To knockdown vimentin, MMTV-PyMT and APC organoids expressing doxycycline-inducible shRNA against vimentin were generated. Two independent shRNA sequences were designed (Table 1). Successful incorporation of the shRNA sequences was confirmed by agarose gel electrophoresis, showing digestion products, and by DNA sequencing. Lentiviral particles were then produced by transfecting HEK293 cells with the FUTG-tetInd-shRNA vector and packaging plasmids using X-tremeGENE (Cat#6365809001, Sigma Aldrich).

Supernatant containing viral particles from HEK293 cell cultures were collected at two time points, 48 and 72 hours following transfection. To concentrate the viral particles, the supernatant was spun down by ultra-centrifugation (175000G for 2.5 hours), using the Optima LE-80K Ultracentrifuge. For organoid infection, organoids were first dissociated into small clusters of cells and single cells, then incubated with concentrated lentiviral particles for 4–6 hours. Following transduction, cells were re-embedded into BME and maintained for 48 hours. Transduced organoids were enriched by adding the selection marker blasticidin (10 µg/mL) (Cat# ANT-BL-05, Invitrogen). To induce vimentin knockdown, doxycycline (2 µg/mL, Cat# D9891, Sigma-Aldrich) was added to the organoid cultures in BME for 3 days. After pretreatment, organoids were extracted from BME and seeded into Collagen I matrices to assess for effects on invasion and knockdown efficiency (Western blot analysis and/or immunofluorescence microscopy).

Generation of Itgα2-deficient organoids

Itgα2 knockout MMTV-PyMT organoids were generated by transfection with a pX458 vector that expresses Cas9-T2A-GFP and Itgα2 sgRNAs targeting the first coding exon of Itgα2 (sgRNA1 gAATTGTCTGGCGTATAATGT, sgRNA2 cGCGTATAATGTTGGCCTCCC). To transfect the organoids, 6 µg of the pX458 vector was mixed with 300 µL Optimem and 18 µL TransIT transfection reagent (Cat# MIR 2300, Mirus) and incubated for 15 minutes at room temperature to allow complex formation. The transfection mixture was then added to single-cell suspension in culture medium (5% FCS) and plated on 3 wells of a 6-well plate for 24 hours. Next, the culture medium was replaced with the FCS-free organoid culture media. The following day, GFP-positive cells were sorted using a FACS Aria II flow cytometer and were embedded in BME and maintained in culture until the single cells developed into multicellular organoids. Around 10 single organoids were picked from the BME gel using a pipette under a brightfield microscope. These selected organoids were trypsinized and expanded separately in BME gels to establish monoclonal lines. To confirm Itgα2 knockout, western blot analysis DNA sequencing were performed. For sequencing, genomic DNA was isolated from organoid pellets using the QIAamp DNA Micro Kit (Cat# 56304, Qiagen). A 999 bp region around the Cas9 target site was amplified in a 50 µL reaction containing 0.5 µLQ5 Polymerase (Cat# NEB M0491L, Bioke), 10 µL 5x Q5 Buffer (Bioke), 1 µL isolated gDNA, 500 nM forward primer Itgα2: TCCAGAGCCTTCCCTTACAA, 500 nM reverse primer Itgα2: AGCCCTGCAGAAAGACGATA, 10 µL Q5 high GC buffer (Bioke), 200 µLM dNTP's (Roche) and 17.5 µL H2O. The PCR program was as follows: initial denaturation at 96°C for 2 min, denaturing at 95°C for 30 seconds, annealing at 57.9°C for 30 seconds extension at 72°C for 45 seconds, and final extension at 72°C for 7 minutes. The denaturing, annealing, and extension steps were repeated 32 times. The PCR products at around 999 bp were isolated from TAE agarose gels using the QIAquick Gel Extraction Kit (Cat# 28704, Qiagen) and sent for sequencing (Macrogen) using the forward primer.

Western blot analysis

To isolate organoids for whole cell lysates, Collagen I gels were first digested with collagenase mixture (0.56 mg/mL; Cat# C0130, Sigma-Aldrich) for 3 minutes at 37°C with gentle shaking[77]. Following digestion, organoids were washed with PBS to remove residual collagenase.

After washing with PBS, organoids were lysed in SDS sample buffer (62.5 mM Tris HCl, 2% SDS, 10% glycerol, 50 mM DTT and bromophenol blue). Lysates were sonicated for 5 cycles at 4°C (Bioruptor) followed by 5-minute boiling at 95°C. Samples were loaded onto a 10% acrylamide gel and separated by SDS-PAGE in Tris-glycine running buffer at 100V for ~2 hours. Proteins were then transferred onto a PVDF membrane (Cat# IPFL00010, Millipore) using wet transfer at 0.35 A for 90 minutes. To reduce non-specific antibody binding, membranes were blocked for 1 hour at room temperature in 5% non-fat dry milk dissolved in 1X TBS. Blots were incubated overnight at 4°C with primary antibodies in antibody solution containing 0.1% Tween-20 and 5% BSA in TBS. After extensive washing (4X 10 minutes, 0.1% Tween-20 in TBS), membranes were incubated with secondary antibodies for 1 hour at room temperature followed by a second round of washing. The following dilutions of primary and secondary antibodies were used integrin α2 (Cat#ab181548, Abcam; 1:5000), Vimentin (Cat#ab92547, Abcam; 1:2500) and GAPDH (Cat#MAB374; 1:5000). Secondary antibodies used were Odyssey goat anti-mouse 800 and goat anti-rabbit 680 (1:2500). Fluorescent signal was measured and visualized using the Amersham Typhoon scanner. To calculate knockdown efficiency using western blot, mean gray values were measured using FIJI software for both the

protein of interest and the housekeeping gene (GAPDH). The efficiency was determined using the following formula:

$(Ratio^{control} - Ratio^{Doxy\ treated}) / Ratio^{control} X\ 100\%$

The ratio was calculated by dividing the mean grey values of the protein of interest (band) by the corresponding mean gray value of the GAPDH band.

Real-Time PCR (qPCR)

For RNA isolation from organoids embedded in Collagen I or BME, gels were collected and lysed by mechanical disruption in 500 µL TRIzol™ reagent (Thermo Fisher) per 40 µL gel. RNA was extracted using the NucleoSpin® RNA kit (Cat#740955.50, Macherey-Nagel) and subsequently converted into cDNA using the iScript™ cDNA Synthesis Kit (Cat#1708891, Bio-Rad) (500-1000 ng in 20 µL).

For qPCR, the reaction mixture contained cDNA, 0.4 µM forward and reverse primers, and 1X FastStart Universal SYBR Green Master mix (Cat#4913850001, Roche) to a final volume of 15 µL.

qPCR primers (Table 2) were designed to generate products between 50-150 bp and were confirmed to specifically bind the mouse gene of interest using the NCBI n-BLAST tool. Primers were validated through cDNA titration and were approved when amplification efficiency ranged between 80-120% and when a single distinct melting peak was observed during the qPCR melting curve analysis

The qPCR was performed using a thermal cycler (CFX384, Bio-Rad) using the following program: 95 °C for 10 min, followed by a 40x repeated cycle of 95 °C for 10 s (denaturation), 55 °C for 10 s (annealing) and 72 °C for 30 s (extension). A melt curve analysis was conducted with a program of 95 °C for 10 s, followed by incremental increases of 0.5 °C from 65 °C to 95 °C to confirm primer specificity and detect potential non-specific amplification. mRNA expression levels were normalized across samples using the average expression of four housekeeping genes (Cyca, Tuba1, HNRNPA, and PBGD). Data was analyzed using the CFX Maestro software, and relative gene expression changes were calculated using the ΔΔCt method.

Statistical Analysis

For statistical analysis Graph Prism (version 8.0) was used. The two-tailed unpaired Mann–Whitney (two groups) or Kruskal–Wallis test (more than two groups) with Dunns' correction were used for statistical analysis. P values less than 0.05 were considered significant.

Tables

Table 1. Sequences of shRNA constructs (1 and 2) targeting vimentin.

shRNA	Sequence Fwd primer (5' à 3')	Sequence Rev primer (5' à 3')
Vimentin 1	tgatagagatACCGGGTGGAATCCTTGCAGGAAGAACTCGAGTTCTTCCTGCAAGGATTCCACTTTTTtcgagacgcgtccta	taggacgcgtctcgaAAAAAGTGGAATCCTTGCAGGAAGAACTCGAGTTCTTCCTGCAAGGATTCCACCCGGTatctctatca
Vimentin 2	tgatagagatACCGGGCTGTTTAGCAGACTTCATAACTCGAGTTATGAAGTCTGCTAAACAGCTTTTTtcgagacgcgtccta	taggacgcgtctcgaAAAAAGCTGTTTAGCAGACTTCATAACTCGAGTTATGAAGTCTGCTAAACAGCCCGGTatctctatca

Table 2. Sequences of qPCR primers.

Gene (mouse)	Sequence Fwd primer (5' à 3')	Sequence Rev primer (5' à 3')
Cyca	ACTGAATGGCTGGATGGCAA	CAAAACGCTCCATGGCTTCC
Tuba1	AGGATTATGAGGAGGTTGGT	ATAAACATCCCTGTGGAAGC
HNRNPA	TGACAGCTATAACAACGGAG	AAAGTTTCCTCCCTTCATCG
PBGD	GCCTACCATACTACCTCCTGGCT	AAGACAACAGCATCACAAGGGTT
Snail1	GCGTGTGTGGAGTTCACCTT	GGTACCAGGAGAGAGTCCCA
Smad7	GGGCTTTCAGATTCCCAACTTC	CCATTCCCCTGAGGTAGATCAT
Pai1	CAGCATGTTCATCGCTGCAC	CCAGAGAGAACTTAGGCAGGA
Slug	CCAAGAAGCCCAACTACAGC	TGGTATGACAGGTATAGGGTAACT
CTGF	TGCGAAGCTGACCTGGAGGAAA	CCGCAGAACTTAGCCCTGTATG
PTHRP	GAAGCGAAGGACTCGGTCTG	TCAATGCGTCCTTAAGCTGG
CDH1	CCAAGCACGTATCAGGGTCA	GTGTACACAGCTTTCCACGC
Yap1	CCTCGTTTTGCCATGAACCAG	GCTGCTGCTGGTTGGAACTG
ANKRD1	GCTGGTAACAGGCAAAAAGAAC	CCTCTCGCAGTTTCTCGCT
Cyr61	GAAAGGGATCTGCAGAGCTCAG	GCACTGGTGTTTACAGTTGG
Areg	GGTGGTGACATGCAATTGTC	GCAATCTTGGATAGGTCCTTG
Vimentin	GCACGATGAAGAGATCCAGGA	CATACTGCTGGCGCACATCAC
PLS3	AGGAATGAAGCACTGGCAGCT	TGATGTCAGCGCTGAAGTTGT
ITGA2	GGAGGCTTTCTAACCTGTGGTC	CCTGAACTGCAGGTGAAAAGC
INHBA	GGAGATAGAGGACGACATTGGC	ACGCTCCACTACTGACAGGTCA
INHBB	TTGCAGAGACAGATGGCCTCG	TCCAGGACATAGGGGAGCAGT
Jag1	AGTGACCTGTGATGACCACT	CGGCAGATAGCTTTGTTGCA

Table 3. Patients' information.

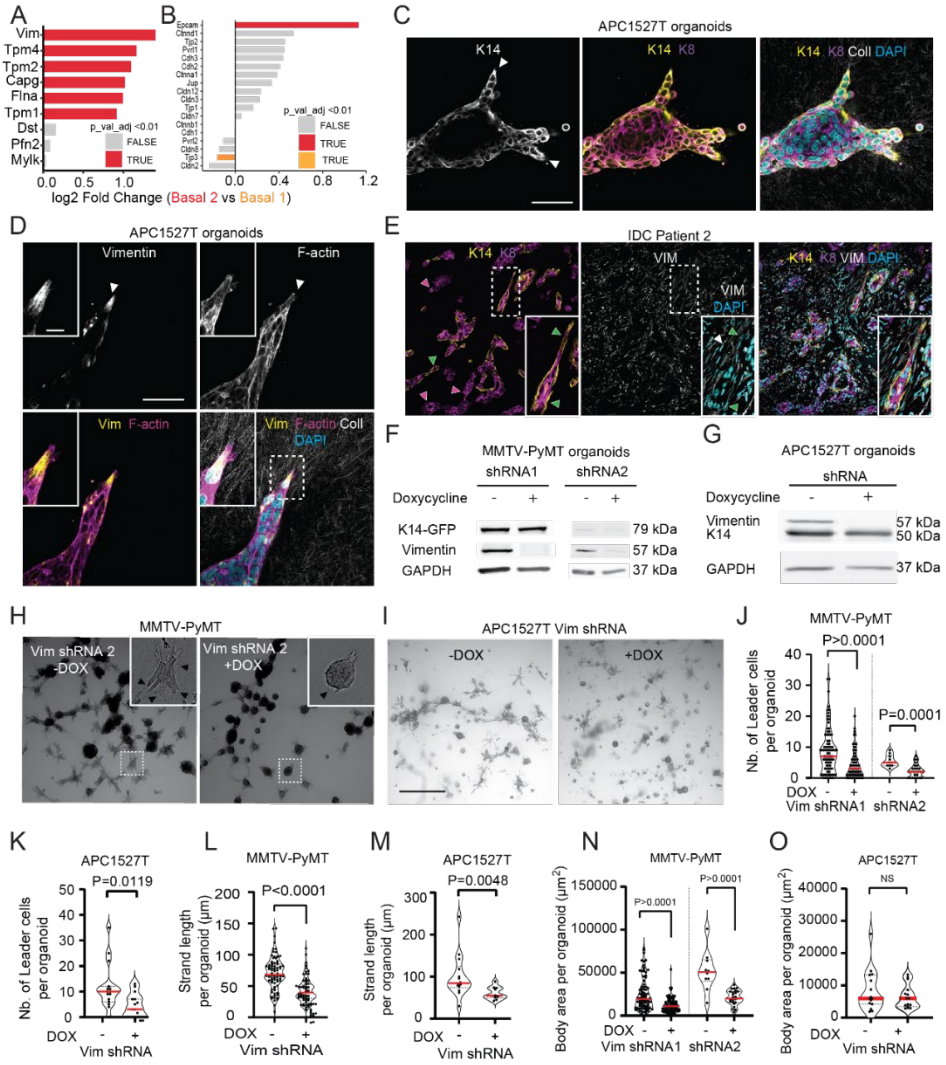
Patients	Diagnosis	Subtype	Age	Grade
14	IDC-NST	ER (-) / PR (-) / HER2 (+)	74	3
25	IDC-NST	ER (+) / PR (+) / HER2 (-)	47	2
19	IDC-NST	ER (+) / PR (+) / HER2 (-)	37	3
45	IDC-NST	ND	42	3
5	IDC-NST	ER (+) / PR (-) / HER2 (-)	39	3

Age: age at primary diagnosis

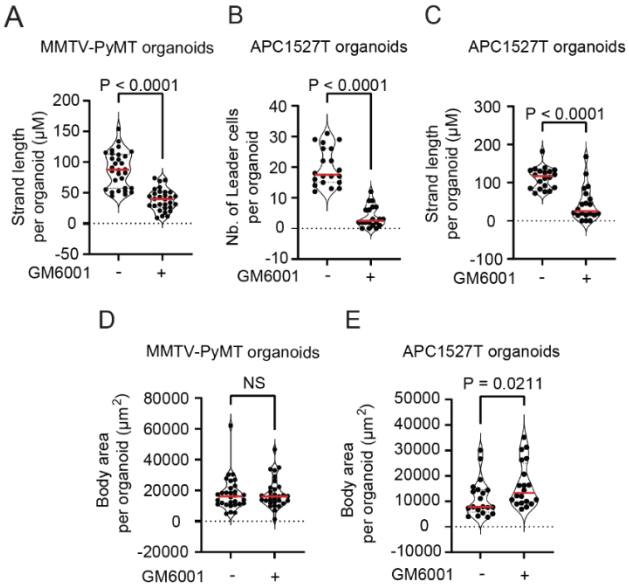
IDC-NST: Invasive ductal carcinomaNST: no special type

ND: Not Determined

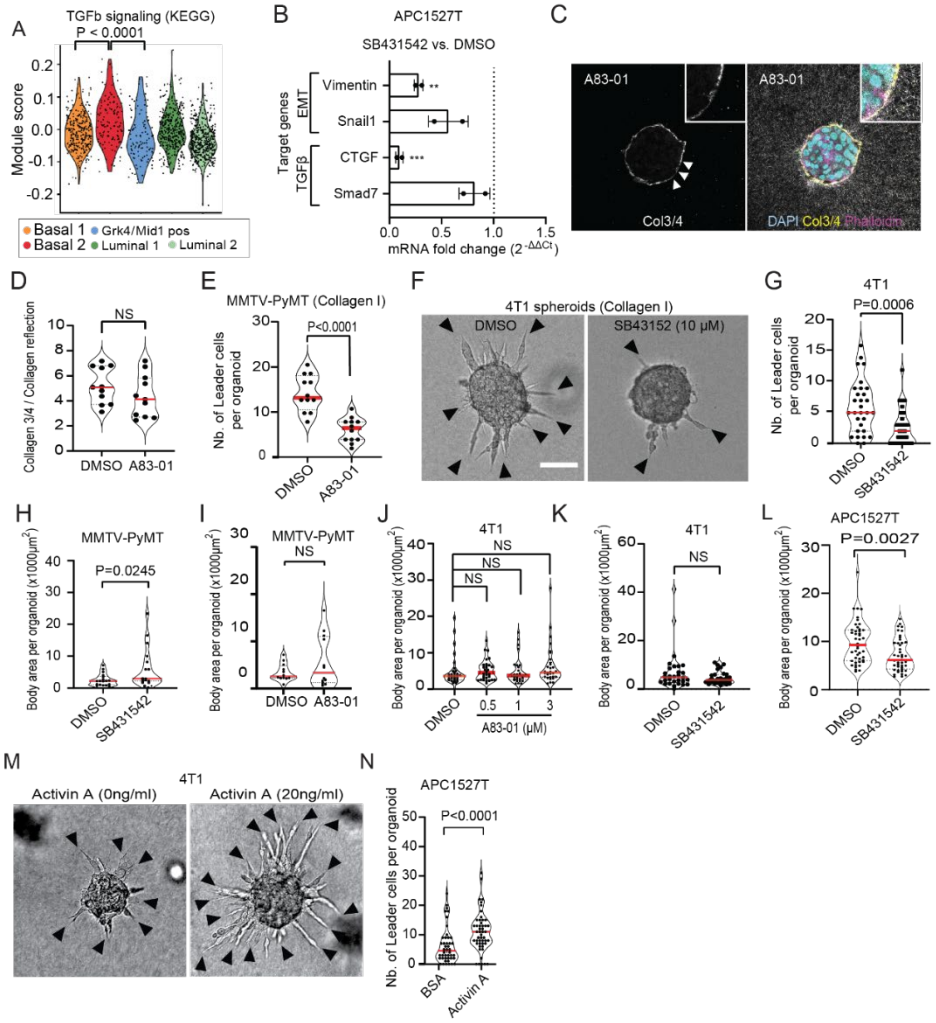
Supplementary Figures:



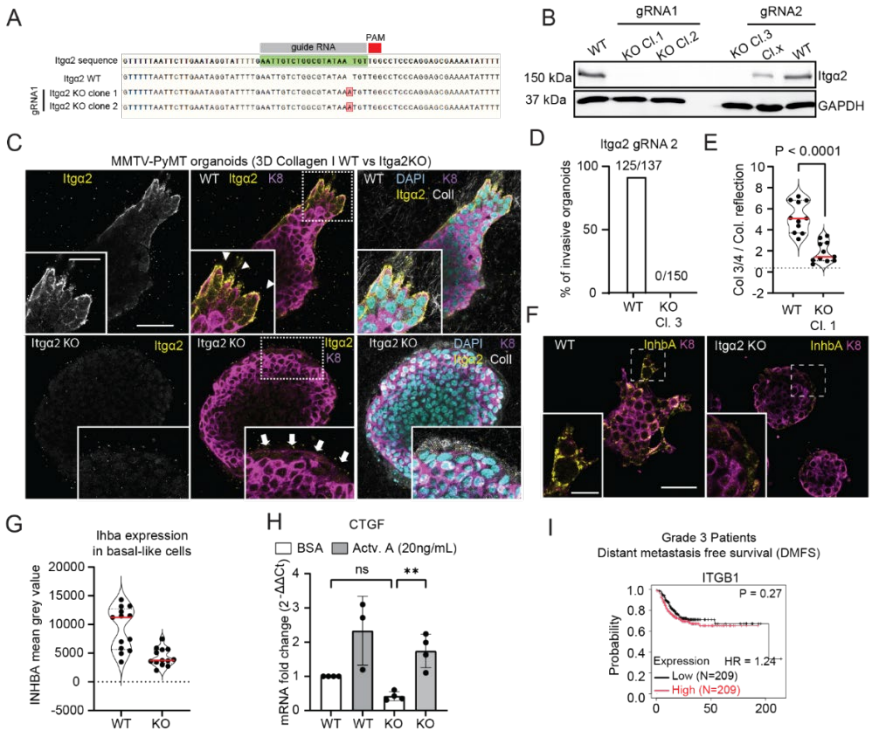
Supplementary Figure 1. (A, B) Log_2 fold changes in gene expression between Basal 2 and Basal 1 subsets, based on scRNA-seq data from MMTV-PyMT organoids**. (A)** Genes belonging to the Hallmarks of Cancer gene set that encode cytoskeletal components. (B) Genes within a curated list of known cell–cell junction components, including adherens, tight, gap, and desmosomal junction proteins. **(C, D)** Confocal imaging of *Apc*^{1572T} organoids grown in 3D Collagen I (reflection) for three days. **(C)** K14 and K8; white arrowheads indicate basal-like leader cells (K14-high, K8-low) driving collective invasion as strands. **(D)** Vimentin expression in protrusive leader cells (white arrowheads). **(E)** Immunostaining of K14, K8, and Vimentin in a whole tissue section from a K14-positive IDC patient, with zoom-in on protrusive cell co-expressing K14 and Vimentin. Arrowheads: green, Vimentin; magenta, basal-like cells. **(F, G)** Western blot analysis of Vimentin, K14, and GAPDH from whole-cell lysates of **(F)** MMTV-PyMT and **(G)** APC1527T organoids expressing doxycycline-inducible Vimentin shRNAs cultured in Collagen I for three days with or without doxycycline (DOX). **(H, I)** Brightfield imaging of control and organoids with Vimentin knockdown (Day 3, Collagen I). Arrowheads (H): leader cells. **(J-O)** Quantification of (J, K) number of leader cells per organoid, (L, M) average length of invasive strands, and (N, O) organoid surface area in control and Vimentin knockdown organoids: (J, L, N) MMTV-PyMT: n = 87 organoids per condition (shRNA1), and n = 29 (−DOX), n = 44 (+DOX) for shRNA2, from three (shRNA1) and two (shRNA2) independent experiments. (K, M, O) *Apc*^{1572T}: n = 14 (−DOX), n = 10 (+DOX) organoids (shRNA1) from one experiment. Red lines, median. Scale bars: 50 µm (C, D), 10 µm (D, zoom-in), 100 µm (E), 25 µm (E, zoom-in), 1000 µm (H, I), 10 µm (H, zoom-in). P values**:** two-sided unpaired Mann–Whitney test (K, L, M, O); two-sided Kruskal-Wallis test with Dunn's multiple comparisons (J, N).



Supplementary Figure 2. (A, D) Average length of strands per organoid in (A) MMTV-PyMT or (D) *Apc*^{1572T} organoids cultured for 3 days in Collagen I and treated with GM6001 or DMSO control. (B) Representative brightfield images of *Apc*^{1572T} organoids in Collagen I treated with GM6001 or DMSO control for three days. Arrowheads: invasive strands. (C) Number of leader cells per *Apc*^{1572T} organoids cultured in 3D Collagen I for 3 days and treated with GM6001 or DMSO control. (E, F) Surface area of organoid body after GM6001 treatment in (E) MMTV-PyMT or (F) *Apc*^{1572T} organoids cultured for 3 days in Collagen I. (A, CF) Median (red line), n = 30 organoids per condition from three independent experiments. Scale bars: 1000 µm (B), inserts 10 µm (B, zoom in). P values, two-sided unpaired Mann–Whitney test (A-E).



Supplemetary Figure 3. (A) Module scores representing active TGF-β signaling based on the KEGG gene set reflecting relative pathway enrichment in basal 2 compared to the other subpopulations identified by single-cell RNA sequencing. P values: one-way ANOVA with Tukey's test for multiple comparisons. **(B)** qPCR analysis of classical TGF-β and EMT target gene expression in APC1527T organoids cultured in 3D Collagen I for three days. Values represent mean normalized mRNA expression (relative to housekeeping genes) in organoids treated with 5 µM SB431542, shown relative to DMSO-treated controls (dashed line). Data represent mean ± SD from two independent experiments. **(C)** Confocal imaging of Col ¾ and F-actin in MMTV-PyMT organoids in 3D Collagen I treated with A83-01 (0.5 µM) for 1 day. Arrowheads: pericellular signal of Col ¾ (degraded Collagen I). Inset, organoid-Collagen I interface. **(D)** Mean gray value of Col ¾ relative to Collagen I reflection in MMTV-PyMT organoids after treatment with A83-01 or DMSO control. Median: red lines, from n = 11 organoids per group from two independent experiments. **(E)** Number of leader cells per organoid in DMSO and A831-01 treated organoids for 3 days. **(F)** Brightfield images of 4T1 spheroids cultured in Collagen I for two days, treated with 10 µM SB431542 or DMSO control. Black arrowheads: leader cells. **(G)** Number of leader cells per spheroids in treated (10 µM SB431542) and control conditions (DMSO). Median: red lines, from 30 organoids per condition from three independent experiments**. (H-L)** Surface area of organoid body after treatment with SB431542 or DMSO control in **(H)** MMTV-PyMT organoids, (K) 4T1 spheroids, and (L) APC1527T organoids, or treatment with A83-01 or DMSO control in (I) MMTV-PyMT organoids and (J) 4T1 spheroids by a serial dilution (0.1 µM, 1 µM, and 3 µM). Median: red lines, from (H) n =18 organoids, (E, I) n = 12 organoids, (J) n = 30 organoids, (K) n = 30 organoids, (L) n = 44 organoids per condition from three independent experiments. (M) Brightfield images of 4T1 spheroids cultured in Collagen I for two days, treated with Activin A ligand (20ng/µl) or an equal volume of 0.1% BSA in PBS as control. Black arrowheads: invasive strands. (N) Quantification of leader cells per organoid in APC1527T organoids treated with Activin A ligand or BSA control for three days in 3D Collagen I. Median: red lines, from 44 organoids per condition from three independent experiments. Scale bars: 50 µm (C), inserts 10 µm (C, Zoom in), 100 µm (F, M). P values, two-sided unpaired Mann–Whitney test (B, D, E, G, H, I, K, L, N), two-sided Kruskal-Wallis test (Dunn's multiple comparison) (J).



Supplementary Figure 4: (A) DNA sequence of the Itgα2 gene to confirm gene knockout by Crispr-Cas9 gene editing. Red regions indicate the insertion of one base pair compared to the wildtype Itgα2 sequence. **(B)** Western blot analysis showing Itgα2 and GAPDH expression from whole cell lysates of MMTV-PyMT organoids WT or KO (gRNA1 or gRNA2). **(C)** Confocal imaging of Itgα2 and K8 in MMTV-PyMT organoids with Itgα2 WT or knockout (KO) grown in Collagen I for 3 days. White arrowheads: invading strands in Itgα2 WT organoids led by basal-like cells (low K8, high Itgα2, insets), White arrows: non-invading basal-like cells at the ECM interface in Itgα2 KO organoids (low K8, low Itgα2, insets). **(D)** Percentage of invasive organoids in MMTV-PyMT organoids Itgα2-WT versus Itgα2-KO (gRNA2). **(E)** Mean gray value of Col ¾ relative to Collagen I reflection in MMTV-PyMT organoids with Itgα2-WT and Itgα2-KO. Median: red lines, from n = 11 organoids per group from two independent experiments. **(F)** Single confocal slice showing INHBA and K8 expression in Itgα2-WT and Itgα2-KO MMTV-PyMT organoids cultured in 3D Collagen I for three days. Insets show basal-like cells (K8 low) guiding invasive strands (Itgα2-WT) or remaining at the organoid rim (Itgα2 KO, non-invasive). **(G)** Mean-gray value of Inhba in basal-like cells (K8-low) located at the rim of Itgα2 WT vs. KO MMTV-PyMT organoids (3D Collagen I, day 1). Median: red lines, n =11 cells from 6 Itgα2-WT organoids, and n = 12 cells from 7 Itgα2-KO organoids from one experiment. **(H)** qPCR analysis showing relative mRNA expression of CTGF in MMTV-PyMT Itgα2-KO organoids compared to Itgα2-WT organoids treated with Activin A ligand (20 ng/µl) or vehicle control (0.1% BSA) for three days. Bar graph represents mean normalized expression values (relative to housekeeping genes) ± SD from four independent experiments. **(I)** Kaplan-Meier plot correlating ITGB1 gene expression with DMFS in grade 3 breast cancer patients. Scale bars: 50 µm (C, F), 25 µm (C, F, Zoom in). P values, two-sided unpaired Mann–Whitney test (E, G), two-sided Kruskal-Wallis test (Dunn's multiple comparison) (H), Logrank test (I).

Acknowledgments
This work was supported by grants from the Dutch Cancer Society (KWF 2020-13552, AAK), the European Union's Horizon 2020 FET Proactive program (grant agreement No. 731957, MECHANO-CONTROL, AAK), and ZonMw (Vici 9150181910029). We thank Kitty van Zwieten, Lars Kemp, Marjolein Vliem, Isabella da Silva, Livio Kleij, Mirjam van der Net and Ingrid Verlaan for the technical support, and Dr. Johan de Rooij,Dr. Fried Zwartkruis, Dr. Eric Kalkhoven and Dr. Paul Coffer for scientific discussions.

Author contributions A.G.E.G., L.J., and A.A.K. designed the study. A.G.E.G., L.J., S.S., H.K., and M.M. performed the experiments. M.P.V. performed bioinformatics analysis. A.G.E.G., L.J., S.S., H.K., M.M. and A.A.K. analyzed the data. L.J., D.W., H.K., R.F. established and provided the organoid lines. A.G.E.G. and A.A.K. wrote the manuscript. All authors read the manuscript and were given the opportunity to provide input.

Conflict of Interests

The authors declare no conflict of interests

References

1. Bronsert, P., Enderle-Ammour, K., Bader, M., Timme, S., Kuehs, M., Csanadi, A., Kayser, G., Kohler, I., Bausch, D., Hoeppner, J., et al. (2014). Cancer cell invasion and EMT marker expression: a three-dimensional study of the human cancer-host interface. J Pathol *234*, 410–422. https://doi.org/10.1002/path.4416.

2. Khalil, A.A., Ilina, O., Gritsenko, P.G., Bult, P., Span, P.N., and Friedl, P. (2017). Collective invasion in ductal and lobular breast cancer associates with distant metastasis. Clin Exp Metastasis *34*, 421–429. https://doi.org/10.1007/S10585-017-9858-6/FIGURES/3.

3. Cheung, K.J., Gabrielson, E., Werb, Z., and Ewald, A.J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. Cell *155*, 1639–1651. https://doi.org/10.1016/J.CELL.2013.11.029/ATTACHMENT/2A1FD816-3063-4634-A607-E4C61311C346/MMC8.PDF.

4. Cheung, K.J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J.D., Fairchild, A.N., Gorin, M.A., Verdone, J.E., Pienta, K.J., Bader, J.S., et al. (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–E863. https://doi.org/10.1073/PNAS.1508541113/SUPPL_FILE/PNAS.201508541SI.PDF.

5. Sonzogni, O., Haynes, J., Seifried, L.A., Kamel, Y.M., Huang, K., BeGora, M.D., Yeung, F.A., Robert-Tissot, C., Heng, Y.J., Yuan, X., et al. (2018). Reporters to mark and eliminate basal or luminal epithelial cells in culture and in vivo. PLoS Biol *16*, e2004049. https://doi.org/10.1371/JOURNAL.PBIO.2004049.

6. Cheung, K.J., and Ewald, A.J. (2016). A collective route to metastasis: Seeding by tumor cell clusters. Science (1979) *352*, 167–169. https://doi.org/10.1126/SCIENCE.AAF6546.

7. Khalil, A.A., Smits, D., Haughton, P.D., Koorman, T., Jansen, K.A., Verhagen, M.P., van der Net, M., van Zwieten, K., Enserink, L., Jansen, L., et al. (2024). A YAP-centered mechanotransduction loop drives collective breast cancer cell invasion. Nature Communications 2024 15:1 *15*, 1–17. https://doi.org/10.1038/s41467-024-49230-z.

8. Kai, F.B., Drain, A.P., and Weaver, V.M. (2019). The Extracellular Matrix Modulates the Metastatic Journey. Dev Cell *49*, 332–346. https://doi.org/10.1016/J.DEVCEL.2019.03.026/ASSET/4E5A940A-878F-4F5D-A25D-A94B3808AADC/MAIN.ASSETS/GR4.JPG.

9. Conklin, M.W., and Keely, P.J. (2012). Why the stroma matters in breast cancer. Cell Adh Migr *6*, 249–260. https://doi.org/10.4161/CAM.20567.

10. Conklin, M.W., Eickhoff, J.C., Riching, K.M., Pehlke, C.A., Eliceiri, K.W., Provenzano, P.P., Friedl, A., and Keely, P.J. (2011). Aligned collagen is a prognostic signature for survival in human breast carcinoma. Am J Pathol *178*, 1221–1232. https://doi.org/10.1016/J.AJPATH.2010.11.076.

11. Hamidi, H., and Ivaska, J. (2018). Every step of the way: integrins in cancer progression and metastasis. Nature Reviews Cancer 2018 18:9 *18*, 533–548. https://doi.org/10.1038/s41568-018-0038-z.

12. Jokinen, J., Dadu, E., Nykvist, P., Käpylä, J., White, D.J., Ivaska, J., Vehviläinen, P., Reunanen, H., Larjava, H., Häkkinen, L., et al. (2004). Integrin-mediated Cell Adhesion to Type I Collagen Fibrils. Journal of Biological Chemistry *279*, 31956–31963. https://doi.org/10.1074/JBC.M401409200.

13. Ross, T.D., Coon, B.G., Yun, S., Baeyens, N., Tanaka, K., Ouyang, M., and Schwartz, M.A. (2013). Integrins in mechanotransduction. Curr Opin Cell Biol *25*, 613–618. https://doi.org/10.1016/J.CEB.2013.05.006.

14. Zuidema, A., Wang, W., and Sonnenberg, A. (2020). Crosstalk between Cell Adhesion Complexes in Regulation of Mechanotransduction. Bioessays *42*. https://doi.org/10.1002/BIES.202000119.

15. Sun, Z., Guo, S.S., and Fässler, R. (2016). Integrin-mediated mechanotransduction. Journal of Cell Biology *215*, 445–456. https://doi.org/10.1083/JCB.201609037.

16. Schwartz, M.A. (2010). Integrins and Extracellular Matrix in Mechanotransduction. Cold Spring Harb Perspect Biol *2*, a005066. https://doi.org/10.1101/CSHPERSPECT.A005066.

17. Wagh, K., Ishikawa, M., Garcia, D.A., Stavreva, D.A., Upadhyaya, A., and Hager, G.L. (2021). Mechanical regulation of transcription: Recent advances. Trends Cell Biol *31*, 457. https://doi.org/10.1016/J.TCB.2021.02.008.

18. Friedl, P., and Alexander, S. (2011). Cancer invasion and the microenvironment: Plasticity and reciprocity. Cell *147*, 992–1009. https://doi.org/10.1016/J.CELL.2011.11.016/ASSET/EAF86637-C1AD-45E6-BCC2-D5F8F216CC35/MAIN.ASSETS/GR3.JPG.

19. Dupont, S., Morsut, L., Aragona, M., Enzo, E., Giulitti, S., Cordenonsi, M., Zanconato, F., Le Digabel, J., Forcato, M., Bicciato, S., et al. (2011). Role of YAP/TAZ in mechanotransduction. Nature 2011 474:7350 *474*, 179–183. https://doi.org/10.1038/nature10137.

20. Khalil, A.A., and Friedl, P. (2010). Determinants of leader cells in collective cell migration. Integrative Biology *2*, 568–574. https://doi.org/10.1039/C0IB00052C.

21. Kopanska, K.S., Alcheikh, Y., Staneva, R., Vignjevic, D., and Betz, T. (2016). Tensile Forces Originating from Cancer Spheroids Facilitate Tumor Invasion. PLoS One *11*, e0156442. https://doi.org/10.1371/JOURNAL.PONE.0156442.

22. Gjorevski, N., Piotrowski, A.S., Varner, V.D., and Nelson, C.M. (2015). Dynamic tensile forces drive collective cell migration through three-dimensional extracellular matrices. Scientific Reports 2015 5:1 *5*, 1–13. https://doi.org/10.1038/srep11458.

23. Grünert, S., Jechlinger, M., and Beug, H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nature Reviews Molecular Cell Biology 2003 4:8 *4*, 657–665. https://doi.org/10.1038/nrm1175.

24. Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J Clin Invest *119*, 1420–1428. https://doi.org/10.1172/JCI39104.

25. Yilmaz, M., and Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer and Metastasis Reviews *28*, 15–33. https://doi.org/10.1007/S10555-008-9169-0/FIGURES/3.

26. Grasset, E.M., Dunworth, M., Sharma, G., Loth, M., Tandurella, J., Cimino-Mathews, A., Gentz, M., Bracht, S., Haynes, M., Fertig, E.J., et al. (2022). Triple-negative breast cancer metastasis involves complex epithelial-mesenchymal transition dynamics and requires vimentin. Sci Transl Med *14*, 7571. https://doi.org/10.1126/SCITRANSLMED.ABN7571/SUPPL_FILE/SCITRANSLMED.ABN7571_MDAR_REPRODUCIBILITY_CHECKLIST.PDF.

27. Helfand, B.T., Mendez, M.G., Murthy, S.N.P., Shumaker, D.K., Grin, B., Mahammad, S., Aebi, U., Wedig, T., Wu, Y.I., Hahn, K.M., et al. (2011). Vimentin organization modulates the formation of lamellipodia. Mol Biol Cell *22*, 1274–1289. https://doi.org/10.1091/MBC.E10-08-0699.

28. Ho Thanh, M.T., and Patteson, A. (2023). Vimentin intermediate filaments enhances collective cell migration through 3D extracellular matrix networks . Biophys J *122*, 266a. https://doi.org/10.1016/j.bpj.2022.11.1525.

29. Patteson, A.E., Vahabikashi, A., Pogoda, K., Adam, S.A., Mandal, K., Kittisopikul, M., Sivagurunathan, S., Goldman, A., Goldman, R.D., and Janmey, P.A. (2019). Vimentin protects cells against nuclear rupture and DNA damage during migration. J Cell Biol *218*, 4079. https://doi.org/10.1083/JCB.201902046.

30. Friedl, P., and Wolf, K. (2008). Tube Travel: The Role of Proteases in Individual and Collective Cancer Cell Invasion. Cancer Res *68*, 7247–7249. https://doi.org/10.1158/0008-5472.CAN-08-0784.

31. Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol Cell Biol *12*, 954–961. https://doi.org/10.1128/MCB.12.3.954-961.1992.

32. Schito, L., and Rey-Keim, S. (2023). Hypoxia signaling and metastatic progression. Semin Cancer Biol *97*, 42–49. https://doi.org/10.1016/J.SEMCANCER.2023.11.001.

33. Kunita, A., Baeriswyl, V., Meda, C., Cabuy, E., Takeshita, K., Giraudo, E., Wicki, A., Fukayama, M., and Christofori, G. (2018). Inflammatory Cytokines Induce Podoplanin Expression at the Tumor Invasive Front. American Journal of Pathology *188*, 1276–1288. https://doi.org/10.1016/J.AJPATH.2018.01.016/ATTACHMENT/C80A1EA0-03D0-46EB-B6AD-DF4218FC6248/MMC2.DOCX.

34. Yao, C.H., Wang, R., Wang, Y., Kung, C.P., Weber, J.D., and Patti, G.J. (2019). Mitochondrial fusion supports increased oxidative phosphorylation during cell proliferation. Elife *8*, e41351. https://doi.org/10.7554/ELIFE.41351.

35. Chen, H., Liu, H., and Qing, G. (2018). Targeting oncogenic Myc as a strategy for cancer treatment. Signal Transduction and Targeted Therapy 2018 3:1 *3*, 1–7. https://doi.org/10.1038/s41392-018-0008-7.

36. Doran, B.R., Moffitt, L.R., Wilson, A.L., Stephens, A.N., and Bilandzic, M. (2024). Leader Cells: Invade and Evade—The Frontline of Cancer Progression. International Journal of Molecular Sciences 2024, Vol. 25, Page 10554 *25*, 10554. https://doi.org/10.3390/IJMS251910554.

37. Gaspar, C., Franken, P., Molenaar, L., Breukel, C., Van Der Valk, M., Smits, R., and Fodde, R. (2009). A Targeted Constitutive Mutation in the Apc Tumor Suppressor Gene Underlies Mammary But Not Intestinal Tumorigenesis. PLoS Genet *5*, e1000547. https://doi.org/10.1371/JOURNAL.PGEN.1000547.

38. Zagami, P., and Carey, L.A. (2022). Triple negative breast cancer: Pitfalls and progress. npj Breast Cancer 2022 8:1 *8*, 1–10. https://doi.org/10.1038/s41523-022-00468-0.

39. Kessenbrock, K., Plaks, V., and Werb, Z. (2010). Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. Cell *141*, 52. https://doi.org/10.1016/J.CELL.2010.03.015.

40. Rosenblum, G., van den Steen, P.E., Cohen, S.R., Bitler, A., Brand, D.D., Opdenakker, G., and Sagi, I. (2010). Direct Visualization of Protease Action on Collagen Triple Helical Structure. PLoS One *5*, e11043. https://doi.org/10.1371/JOURNAL.PONE.0011043.

41. Gross, J., Harper, E., Harris, E.D., McCroskery, P.A., Highberger, J.H., Corbett, C., and Kang, A.H. (1974). Animal collagenases: Specificity of action, and structures of the substrate cleavage site. Biochem Biophys Res Commun *61*, 605–612. https://doi.org/10.1016/0006-291X(74)91000-6.

42. Miller, E.J., Harris, E.D., Chung, E., Finch, J.E., McCroskery, P.A., and Butler, W.T. (1976). Cleavage of Type II and III collagens with mammalian collagenase: site of cleavage and primary structure at the NH2-terminal portion of the smaller fragment released from both collagens. Biochemistry *15*, 787–792. https://doi.org/10.1021/BI00649A009.

43. Drabsch, Y., and Ten Dijke, P. (2012). TGF-β signalling and its role in cancer progression and metastasis. Cancer and Metastasis Reviews *31*, 553–568. https://doi.org/10.1007/S10555-012-9375-7/TABLES/2.

44. Xu, J., Lamouille, S., and Derynck, R. (2009). TGF-beta-induced epithelial to mesenchymal transition. Cell Res *19*, 156–172. https://doi.org/10.1038/CR.2009.5.

45. Inman, G.J., Nicolás, F.J., Callahan, J.F., Harling, J.D., Gaster, L.M., Reith, A.D., Laping, N.J., and Hill, C.S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. Mol Pharmacol *62*, 65–74. https://doi.org/10.1124/MOL.62.1.65.

46. Massagué, J. (2008). TGFbeta in Cancer. Cell *134*, 215–230. https://doi.org/10.1016/J.CELL.2008.07.001.

47. Khalil, A.A., Ilina, O., Vasaturo, A., Venhuizen, J.H., Vullings, M., Venhuizen, V., Bilos, A.B., Figdor, C.G., Span, P.N., and Friedl, P. (2020). Collective invasion induced by an autocrine purinergic loop through connexin-43 hemichannels. Journal of Cell Biology *219*. https://doi.org/10.1083/JCB.201911120/152009.

48. Ling, N., Ying, S.Y., Ueno, N., Shimasaki, S., Esch, F., Hotta, M., and Guillemin, R. (1986). Pituitary FSH is released by a heterodimer of the beta-subunits from the two forms of inhibin. Nature *321*, 779–782. https://doi.org/10.1038/321779A0.

49. Vale, W., Rivier, J., Vaughan, J., Mcclintock, R., Corrigan, A., Woo, W., Karr, D., and Spiess, J. (1986). Purification and characterization of an FSH releasing protein from porcine ovarian follicular fluid. Nature *321*, 776–779. https://doi.org/10.1038/321776A0.

50. Leitinger, B., and Hohenester, E. (2007). Mammalian collagen receptors. Matrix Biology *26*, 146–155. https://doi.org/10.1016/J.MATBIO.2006.10.007.

51. Zeltz, C., and Gullberg, D. (2016). The integrin-collagen connection - a glue for tissue repair? J Cell Sci *129*, 653–664. https://doi.org/10.1242/JCS.180992/260177/AM/THE-INTEGRIN-COLLAGEN-CONNECTION-A-GLUE-FOR-TISSUE.

52. Boraschi-Diaz, I., Wang, J., Mort, J.S., and Komarova, S. V. (2017). Collagen type i as a ligand for receptor-mediated signaling. Front Phys *5*, 249703. https://doi.org/10.3389/FPHY.2017.00012/BIBTEX.

53. Heino, J. (2000). The collagen receptor integrins have distinct ligand recognition and signaling functions. Matrix Biology *19*, 319–323. https://doi.org/10.1016/S0945-053X(00)00076-7.

54. Emsley, J., Knight, C.G., Farndale, R.W., Barnes, M.J., and Liddington, R.C. (2000). Structural basis of collagen recognition by integrin α2β1. Cell *101*, 47–56. https://doi.org/10.1016/S0092-8674(00)80622-4/ASSET/84C2B776-3C41-48BB-9242-1E0FD6D25A57/MAIN.ASSETS/GR4.JPG.

55. Győrffy, B. (2021). Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. Comput Struct Biotechnol J *19*, 4101–4109. https://doi.org/10.1016/J.CSBJ.2021.07.014.

56. Pandya, P., Orgaz, J.L., and Sanz-Moreno, V. (2017). Modes of invasion during tumour dissemination. Mol Oncol *11*, 5–27. https://doi.org/10.1002/1878-0261.12019.

57. Pal, A., Barrett, T.F., Paolini, R., Parikh, A., and Puram, S. V. (2021). Partial EMT in head and neck cancer biology: a spectrum instead of a switch. Oncogene *40*, 5049–5065. https://doi.org/10.1038/S41388-021-01868-5.

58. Zhao, X., Jiang, M., and Wang, Z. (2019). TPM4 promotes cell migration by modulating F-actin formation in lung cancer. Onco Targets Ther *12*, 4055–4063. https://doi.org/10.2147/OTT.S198542.

59. Giovannelli, P., Di Donato, M., Licitra, F., Sabbatino, E., Tutino, V., Castoria, G., and Migliaccio, A. (2024). Filamin A in triple negative breast cancer. Steroids *205*, 109380. https://doi.org/10.1016/J.STEROIDS.2024.109380.

60. Szu-Hua Pan; Pui-San Ieong; Shih-Han Kao; Tse-Ming Hong; Pan-Chyr Yang (2007). CapG can promote cancer cell invasion and predict lung cancer outcome. Tumor Biology. https://aacrjournals.org/cancerres/article/67/9_Supplement/1363/535655/CapG-can-promote-cancer-cell-invasion-and-predict.

61. Koorman, T., Jansen, K.A., Khalil, A., Haughton, P.D., Visser, D., Rätze, M.A.K., Haakma, W.E., Sakalauskaitè, G., van Diest, P.J., de Rooij, J., et al. (2022). Spatial collagen stiffening promotes collective breast cancer cell invasion by reinforcing extracellular matrix alignment. Oncogene *41*, 2458–2469. https://doi.org/10.1038/S41388-022-02258-1.

62. Gopal, S., Veracini, L., Grall, D., Butori, C., Schaub, S., Audebert, S., Camoin, L., Baudelet, E., Adwanska, A., Beghelli-De La Forest Divonne, S., et al. (2017). Fibronectin-guided migration of carcinoma collectives. Nature Communications 2017 8:1 *8*, 1–15. https://doi.org/10.1038/ncomms14105.

63. Nakashima, Y., Yoshinaga, K., Kitao, H., Ando, K., Kimura, Y., Saeki, H., Oki, E., Morita, M., Kakeji, Y., Hirahashi, M., et al. (2013). Podoplanin is expressed at the invasive front of esophageal squamous cell carcinomas and is involved in collective cell invasion. Cancer Sci *104*, 1718. https://doi.org/10.1111/CAS.12286.

64. Yoon, S.B., Chen, L., Robinson, I.E., Khatib, T.O., Arthur, R.A., Claussen, H., Zohbi, N.M., Wu, H., Mouw, J.K., and Marcus, A.I. (2024). Subpopulation commensalism promotes Rac1-dependent invasion of single cells via laminin-332. J Cell Biol *223*, e202308080. https://doi.org/10.1083/JCB.202308080.

65. Friedl, P., and Mayor, R. (2017). Tuning Collective Cell Migration by Cell-Cell Junction Regulation. Cold Spring Harb Perspect Biol *9*. https://doi.org/10.1101/CSHPERSPECT.A029199.

66. Khalil, A.A., and de Rooij, J. (2019). Cadherin mechanotransduction in leader-follower cell specification during collective migration. Exp Cell Res *376*, 86–91. https://doi.org/10.1016/J.YEXCR.2019.01.006.

67. Debaugnies, M., Rodríguez-Acebes, S., Blondeau, J., Parent, M.A., Zocco, M., Song, Y., de Maertelaer, V., Moers, V., Latil, M., Dubois, C., et al. (2023). RHOJ controls EMT-associated resistance to chemotherapy. Nature 2023 616:7955 *616*, 168–175. https://doi.org/10.1038/s41586-023-05838-7.

68. Lambert, A.W., and Weinberg, R.A. (2021). Linking EMT programmes to normal and neoplastic epithelial stem cells. Nature Reviews Cancer 2021 21:5 *21*, 325–338. https://doi.org/10.1038/s41568-021-00332-6.

69. Haeger, A., Alexander, S., Vullings, M., Kaiser, F.M.P., Veelken, C., Flucke, U., Koehl, G.E., Hirschberg, M., Flentje, M., Hoffman, R.M., et al. (2020). Collective cancer invasion forms an integrin-dependent radioresistant niche. J Exp Med *217*. https://doi.org/10.1084/JEM.20181184.

70. Flum, M., Dicks, S., Teng, Y.H., Schrempp, M., Nyström, A., Boerries, M., and Hecht, A. (2022). Canonical TGFβ signaling induces collective invasion in colorectal carcinogenesis through a Snail1- and Zeb1-independent partial EMT. Oncogene 2022 41:10 *41*, 1492–1506. https://doi.org/10.1038/s41388-022-02190-4.

71. Hao, Y., Baker, D., and Dijke, P. Ten (2019). TGF-β-Mediated Epithelial-Mesenchymal Transition and Cancer Metastasis. Int J Mol Sci *20*. https://doi.org/10.3390/IJMS20112767.

72. Zhao, B., Ye, X., Yu, J., Li, L., Li, W., Li, S., Yu, J., Lin, J.D., Wang, C.Y., Chinnaiyan, A.M., et al. (2008). TEAD mediates YAP-dependent gene induction and growth control. Genes Dev *22*, 1962–1971. https://doi.org/10.1101/GAD.1664408.

73. Margadant, C., and Sonnenberg, A. (2010). Integrin-TGF-beta crosstalk in fibrosis, cancer and wound healing. EMBO Rep *11*, 97–105. https://doi.org/10.1038/EMBOR.2009.276.

74. Cichon, M.A., and Radisky, D.C. (2014). Extracellular matrix as a contextual determinant of transforming growth factor-β signaling in epithelial-mesenchymal transition and in cancer. Cell Adh Migr *8*, 588–594. https://doi.org/10.4161/19336918.2014.972788.

75. Riikonen, T., Westermarck, J., Koivisto, L., Broberg, A., Kähäri, V.M., and Heino, J. (1995). Integrin α2β1 Is a Positive Regulator of Collagenase (MMP-1) and Collagen α1(I) Gene Expression. Journal of Biological Chemistry *270*, 13548–13552. https://doi.org/10.1074/JBC.270.22.13548.

76. Beerling, E., Seinstra, D., de Wit, E., Kester, L., van der Velden, D., Maynard, C., Schäfer, R., van Diest, P., Voest, E., van Oudenaarden, A., et al. (2016). Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity. Cell Rep *14*, 2281–2288. https://doi.org/10.1016/J.CELREP.2016.02.034.

77. Smits, D., and Khalil, A.A. (2023). Multimodal Techniques to Study Tumor Growth, Basement Membrane Breaching, and Invasion in 3D Matrices. Methods Mol Biol *2608*, 281–303. https://doi.org/10.1007/978-1-0716-2887-4_17.

78. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587.e29. https://doi.org/10.1016/J.CELL.2021.04.048.

79. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*. https://doi.org/10.1186/1471-2105-14-128.

80. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545–15550. https://doi.org/10.1073/PNAS.0506580102/SUPPL_FILE/06580FIG7.JPG.