A compromise between Tgfβ and Egfr signaling programs confers the ability to lead heterogeneous collective invasion

Apsra Nasir#, Sharon Camacho#, Alec T. McIntosh, Garrett T. Graham, Raneen Rahhal, Molly E. Huysman, Anna T. Riegel and Gray W. Pearson*

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

# Equal contribution

* To whom correspondence should be addressed.

Gray W. Pearson
Lombardi Comprehensive Cancer Center
Georgetown University
3970 Reservoir Road NW,
Washington, DC 20057
Phone- 202-687-0607
E-mail: **gp507@georgetown.edu**

Key Words: Heterogeneity; Epithelial-to-Mesenchymal Transition; Collective Invasion; Tgfβ; Egfr; Fra1

The authors declare no potential conflicts of interest.

ABSTRACT

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

INTRODUCTION

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

To better understand how subpopulation relationships are induced during tumor progression, we sought to understand how tumor cells are conferred with the ability to promote the collective invasion of intrinsically less invasive siblings. Collective invasion is the predominant mode of invasion in carcinomas and promotes the dispersion of tumor clusters that seed metastases (15-19). Subpopulations of trailblazer cells detected in breast, lung and thyroid tumor populations have an enhanced ability to reorganize the extracellular matrix (ECM) into micro-tracks that facilitate the initiation of collective invasion (20-23). Importantly, sibling opportunist populations invade

70 through the micro-tracks created by minority trailblazer populations in three-dimensional culture
71 models and xenograft tumors (21, 24, 25). Moreover, opportunist cells can have greater prolifera-
72 tive capacity than trailblazers in vitro and in xenografts, indicating that trailblazer-opportunist re-
73 lationships have the potential to increase the proliferative fitness of invading cell populations (24,
74 25). Thus, trailblazer cells promote phenotypic diversity in invading cell cohorts, which may in-
75 crease the risk of metastasis and elevate the risk of resistance to treatment (26, 27).
76
77 Trailblazer cells are distinguished from the bulk tumor population by the increased expres-
78 sion of trailblazer genes, including regulators of the actin cytoskeleton, integrins, receptor tyrosine
79 kinases, keratins, and cancer testes antigens (20, 21, 23, 24, 28). Trailblazer genes contribute to
80 the formation of actin-rich cellular protrusions that provide traction and exert tensile forces neces-
81 sary for ECM remodeling and other undefined functions (21, 23, 25). Interestingly, trailblazer
82 genes are not required for other forms of cell movement, including migration along two-dimen-
83 sional surfaces, movement within spheroids and opportunistic collective invasion (21). While
84 genes that specifically contribute to unique invasive characteristics of trailblazer cells have begun
85 to be defined, the regulatory programs underpinning the conversion to a trailblazer state remain
86 largely unknown. Trailblazer cells in breast and lung cancer cell line models are epigenetically
87 distinct from opportunist siblings, although the processes that initiate the conversion to the trail-
88 blazer state are not known (21, 23). Breast cancer trailblazer cells have basal characteristics, with
89 Keratin 14 (Krt14) expression proposed as a universal feature for identifying trailblazer cells in
90 mammary tumors (18, 20, 29, 30). Yet, the program controlling the induction of a basal trailblazer
91 state is undefined. Breast cancer trailblazer cells also have features of epithelial-mesenchymal
92 transition (EMT) re-programming (21). However not all EMT programs yield trailblazer activity

93 and whether EMT programs regulate the expression of trailblazer genes is not known (25, 31).
94 Thus, if and how the EMT process imbues a trailblazer state is unclear. Our poor understanding of
95 how trailblazer populations are induced limits options for potential interventions targeting trail-
96 blazer cells and advances in employing subpopulation identification to refine patient outcome pre-
97 dictions.
98
99 To define factors regulating trailblazer program activity, we determined how cells were
100 induced to collectively invade in the C3-TAg genetically engineered mouse model (GEMM) of
101 breast cancer. Based on histological analysis, breast cancer is defined as Estrogen Receptor-posi-
102 tive (ER-positive), Human Epidermal Growth Factor Receptor-positive (HER2-pos) or triple-neg-
103 ative (TNBC), which have low to undetectable levels of ER, HER2 and the Progesterone Receptor
104 (32). Breast cancer is further stratified into luminal A, luminal B, HER2-enriched and basal-like
105 based on mRNA expression. Luminal A are Estrogen Receptor-positive (ER-pos) while luminal B
106 are more proliferative and are ER-pos or HER2-pos (33). The HER2-enriched subtype is generally
107 HER2-pos and 70-80% of the Basal-like are TNBC (33). The C3-TAg model is most similar to
108 human basal-type triple negative breast cancer (TNBC) (34, 35). Through the analysis of tumor
109 organoid invasion dynamics, a data driven functional dissection of cellular re-programming and
110 quantitative imaging of fixed tumor sections, we discovered a Tgfβ regulated program that was
111 necessary to induce and sustain a trailblazer state in C3-TAg tumors and validated this discovery
112 in breast cancer patient tumor samples. Surprisingly, the Tgfβ regulated trailblazer cells were not
113 specified by Krt14 expression and were more aggressive than a previously defined Krt14 express-
114 ing trailblazer population detected in a luminal B GEMM, indicating that there are multiple differ-
115 ent programs for conferring trailblazer ability in breast tumors. In addition, trailblazer cells induced

by Tgfβ signaling were themselves heterogeneous, having different degrees of re-programming and capacities to initiate invasion. Tgfβ signaling collaborated with a parallel Egfr-Mek1/2-Erk1/2-Fra1 pathway to coordinate the expression of a subset of pro-motility genes, including Axl. Interestingly, the integration was also equivalently antagonistic in scope, with Egfr activity restricting the Tgfβ induced expression of a cohort of pro-invasive genes. Notably, clinically approved inhibitors of Egfr, Mek1/2 and Axl suppressed collective invasion. This indicated that the collateral inhibition of subset of pro-invasive genes by the Egfr pathway was a compromise that was necessary for manifestation of the trailblazer state. Thus, our results reveal how a new type of trailblazer cell is induced by a therapeutically addressable compromise between Tgfβ and Egfr signaling programs.

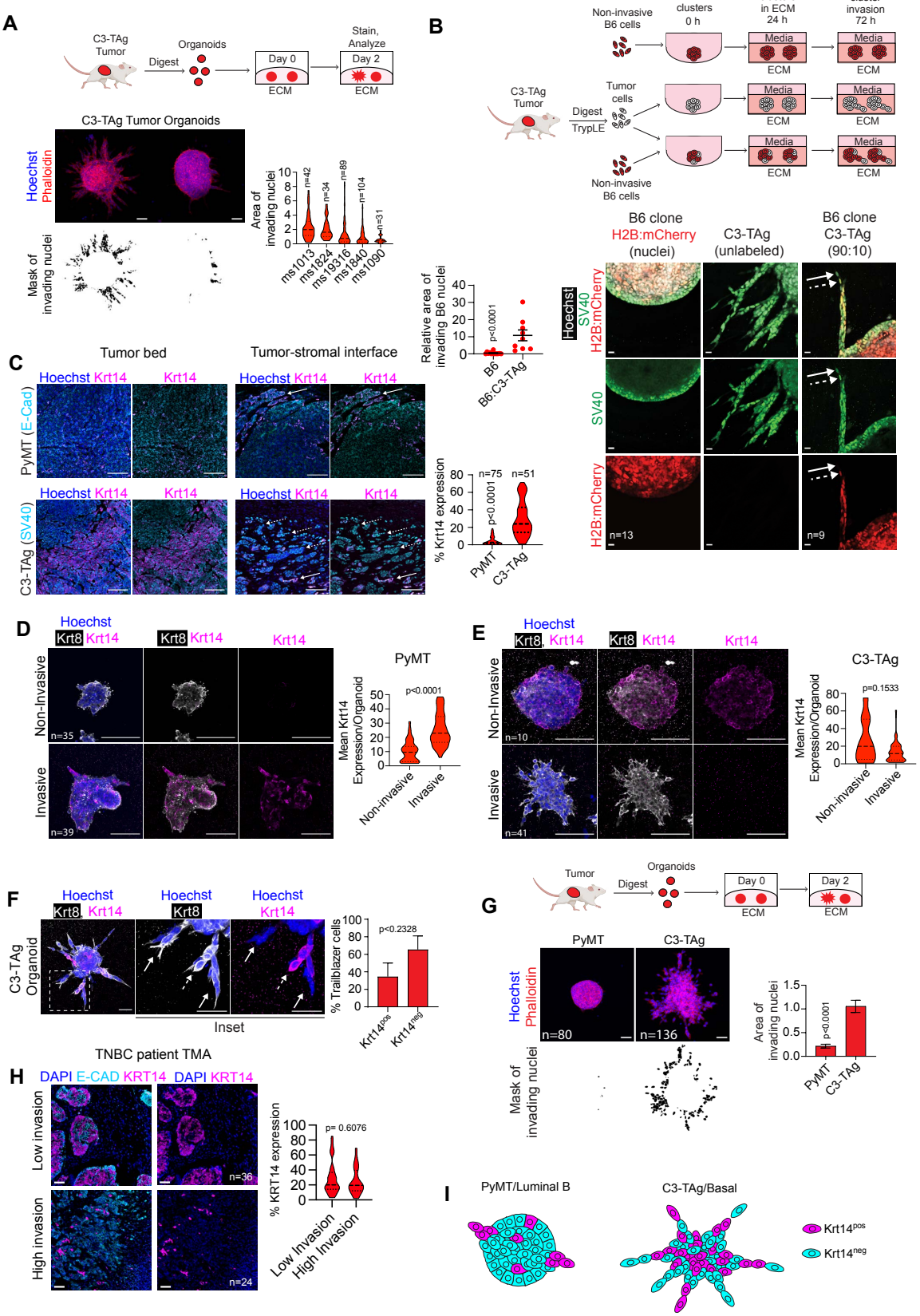
RESULTS

The trailblazer cells in basal-like tumors are distinct from trailblazer cells in luminal B tumors.

To define factors that initiate the conversion to a trailblazer state, we investigated the invasive properties of tumor cells isolated from the C3(1)/SV40 T-antigen (C3-TAg) GEMM of basal-like TNBC (34, 35) and **(Table S1)**. We first determined if there was heterogeneity with respect to the ability to invade in the tumor populations, which would indicate the potential for trailblazer-opportunist relationships. Indeed, there was extensive heterogeneity in the invasive ability of tumor organoids and single dissociated tumor cells **(Fig. 1A, S1A,B, Video S1,S2)**. Clonal invasive (D6 and C7) and clonal non-invasive cells (B6 and G2) were also derived from a single C3-TAg tumor **(Fig. S1C-E, Table S1)**. To test whether the invasive subpopulations in C3-TAg tumors could lead the invasion of siblings, we generated heterogeneous clusters containing C3-TAg tumor cells and intrinsically non-invasive B6 cells **(Fig. 1B, Videos S3-5)**. The C3-TAg tumor cells and the D6 cells led the collective invasion of B6 cells **(Fig. 1B, S1E, Videos S3,6)**. These combined results indicated that there were trailblazer subpopulations in C3-TAg tumors with an enhanced ability to lead the collective invasion of intrinsically less invasive siblings.

To understand how the trailblazer state was induced in C3-TAg tumor cells, we next evaluated the relationship between Krt14 expression and the ability of cells to lead collective invasion. Krt14 identifies cells that have converted from a luminal noninvasive state to a basal trailblazer state in the PyMT GEMM model of luminal B breast cancer, which is commonly used to study breast cancer invasion and metastasis (20, 36). Based on these results, it has been suggested that a basal gene expression program exemplified by Krt14 expression is a universal feature of breast

Figure 1


A
C3-TAg Tumor
Digest
Organoids
Day 0
ECM
Stain, Analyze
Day 2
ECM
C3-TAg Tumor Organoids
Hoechst
Phalloidin
Mask of invading nuclei
Area of invading nuclei
n=42
n=34
n=89
n=104
n=31
ms1013
ms1824
ms19316
ms1840
ms1090
B
Non-invasive B6 cells
Plate cells to form clusters 0 h
Plate clusters in ECM 24 h
Evaluate cluster invasion 72 h
Media
ECM
C3-TAg Tumor
Digest TrypLE
Tumor cells
Non-invasive B6 cells
Relative area of invading B6 nuclei
p<0.0001
B6
B6:C3-TAg
B6 clone H2B:mCherry (nuclei)
C3-TAg (unlabeled)
B6 clone C3-TAg (90:10)
Hoechst
SV40
H2B:mCherry
n=13
n=9
C
Tumor bed
Tumor-stromal interface
Hoechst Krt14
Krt14
PyMT (E-Cad)
C3-TAg (SV40)
% Krt14 expression
n=75
n=51
p<0.0001
PyMT
C3-TAg
D
Hoechst
Krt8 Krt14
Krt14
Non-Invasive
Invasive
n=35
n=39
PyMT
Mean Krt14 Expression/Organoid
p<0.0001
Non-invasive
Invasive
E
Hoechst
Krt8, Krt14
Krt8 Krt14
Krt14
Non-Invasive
Invasive
n=10
n=41
C3-TAg
p=0.1533
F
Hoechst
Krt8, Krt14
Krt8
Krt14
C3-TAg Organoid
Inset
% Trailblazer cells
p<0.2328
Krt14pos
Krt14neg
G
Tumor
Digest
Organoids
Day 0
ECM
Day 2
ECM
PyMT
C3-TAg
Hoechst
Phalloidin
n=80
n=136
Mask of invading nuclei
Area of invading nuclei
p<0.0001
PyMT
C3-TAg
H
TNBC patient TMA
DAPI E-CAD KRT14
DAPI KRT14
Low invasion
High invasion
n=36
n=24
% KRT14 expression
p= 0.6076
Low Invasion
High Invasion
I
PyMT/Luminal B
C3-TAg/Basal
Krt14pos
Krt14neg

162

Figure 1. Trailblazer cells in basal-like tumors are distinct from trailblazer cells in luminal B tumors. **A.** There is intra- and inter-tumor invasive heterogeneity in C3-TAg organoids. Graphic shows the process for the derivation and analysis of tumor organoids. Representative images show the invasive heterogeneity of C3-TAg organoids derived from a single tumor. Image mask defines invasive cell nuclei. Violin plot shows the quantification of organoid invasion from 5 C3-TAg tumors. (Mann Whitney test, n= organoids). Scale bars, 50 µm. **B.** C3-TAg tumors contain trailblazer cells that can lead the collective invasion of intrinsically less invasive cells. Graphic shows the process for generating homogeneous and heterogeneous clusters. Images show clusters immunostained with anti-SV40 antibody (detects TAg in all tumor cells), counterstained with Hoechst (nuclei) and H2B:mCherry fluorescence (B6 cells only). C3-TAg tumor cells (unlabeled) promoted the invasion of B6 cells (H2B:mCherry). Solid arrow indicates a C3-TAg tumor cell leading invasion. Dashed arrow indicates a B6 cell invading behind the C3-TAg cell. Graph shows the relative area of invading B6 nuclei from 3 independent experiments (mean±SEM, Mann Whitney test, n= clusters,). Scale bars, 20 µm. **C.** Krt14 expression does not correlate with C3-TAg tumor invasion. Krt14, SV40 (C3-TAg) and E-Cad (PyMT) expression in the tumor core and tumor-stromal interface of C3-TAg and PyMT tumors. Solid and dashed arrows indicate Krt14^{pos} and Krt14^{neg} cells at the leading collectively invading cells respectively. Violin plot shows percentage of tumor cell area expressing Krt14. (Mann Whitney Test, n=ROIs). Scale bars, 100 µm. **D.** Krt14 expression correlates with PyMT organoid invasion. Violin plot shows mean Krt14 fluorescence in invasive and non-invasive PyMT organoids. (Mann Whitney Test, n=organoids). Scale bars, 100 µm. **E.** Krt14 expression does not correlate with C3-TAg organoid invasion. Violin plot shows mean Krt14 fluorescence in invasive and non-invasive C3-TAg organoids. (Mann Whitney Test, n=organoids). Scale bars, 100 µm. **F.** Krt14 expression does not correlate with the ability to lead collective invasion in C3-TAg organoids. Solid and dashed arrows indicate Krt14^{neg} and Krt14^{pos} leader cells respectively. Bar graph shows Krt14 expression in leading cells of organoids from 3 tumors (mean±SEM, unpaired Student's t test, n=organoids) Scale bars, 50 µm. **G.** C3-TAg organoids were more invasive than PyMT organoids. Bar graph shows the area of invading nuclei (mean±SEM, Mann Whitney test, n=clusters, r=2) Scale bars, 50 µm. **H.** Krt14 expression in TNBC patient tumors did not correlate with invasion. Violin plot shows percentage of Krt14 expressing cells in regions of low and high invasion (Mann Whitney test, n=tumor sections). Scale bars, 50 µm. **I.** Model depicting the tumor subtype relationship between Krt14 expression and trailblazer cell activity.

cancer trailblazer cells (20, 29). Consistent with prior results, Krt14 was expressed in PyMT trailblazer cells and Krt14 expression correlated with PyMT tumor organoid invasion **(Fig. 1C,D)**. Krt14 was expressed in a greater proportion of cells in C3-TAg tumors compared to PyMT tumors, in line with Krt14 being more highly expressed in basal type compared to luminal B type breast cancer patient tumors **(Fig. 1C, S1F)**. The high Krt14 expression was consistent with the early acquisition of basal traits detected in C3-TAg and TNBC tumors during the initial stages of growth within mammary ducts (37, 38). However, Krt14 expression did not correlate with C3-TAg tumor collective invasion or a trailblazer state in our analysis of primary tumors, tumor organoids and trailblazer and opportunist cell lines **(Fig 1C,E,F, S1D,E,G)**. In fact, extensive collective invasion

173 was detected in C3-TAg tumor organoids with few, if any, Krt14 expressing cells **(Fig. 1E)**.
174 Krt14^{neg} C3-TAg tumor cells also led collective invasion as frequently as Krt14^{pos} cells **(Fig 1F)**.
175 C3-TAg organoids were relatively more invasive than PyMT organoids, demonstrating that the
176 C3-TAg tumor invasion was more aggressive than the Krt14 *s*pecified invasion of PyMT tumors
177 **(Fig. 1G, S1H, Videos S7-8)**. Moreover, Krt14 expression did not correlate with collective inva-
178 sion in our analysis of a TNBC patient tissue microarray **(Fig. 1H, S1I)**. This indicated that the
179 Krt14 independent collective invasion in C3-TAg tumors represented a feature of collective inva-
180 sion that was conserved in TNBC tumors. Together, these results suggested that essential features
181 of the C3-TAg trailblazer program were different than those of the PyMT trailblazer program, and
182 more than a single Krt14 specified trailblazer program that functioned in breast cancer patient
183 tumors **(Fig. 1I)**.
184

185 **Trailblazer cells are distinguished by a subset of EMT traits in basal-like mammary tumors.**
186
187 Given the lack of concordance with the established Krt14 specified trailblazer program, we
188 next took an unbiased approach to identify the key trailblazer regulatory program features in C3-
189 TAg tumors. To do this, we used next generation RNA sequencing to define differentially ex-
190 pressed genes (DEGs) in tumor organoids from 4 C3-TAg tumors compared with non-invasive
191 spheroids formed by 3 clonal C3-TAg tumor derived cell lines **(Fig. 2A, S2A)**. Among epithelial
192 keratins, Krt14 expression was reduced in C3-TAg tumor organoids compared to the non- of in-
193 vasive clonal C3-TAg cells, consistent with a lack of Krt14 involvement in C3-TAg collective
194 invasion (**Fig. 2B, Table S2**). Expression of the transcription factor p63 was also reduced in C3-
195 TAg tumor organoids and trailblazer clones, further indicating that canonical traits of normal basal
196 mammary epithelial cells were not essential to induce C3-TAg trailblazer cells (**Fig. 2B, S2B,**

Figure 2

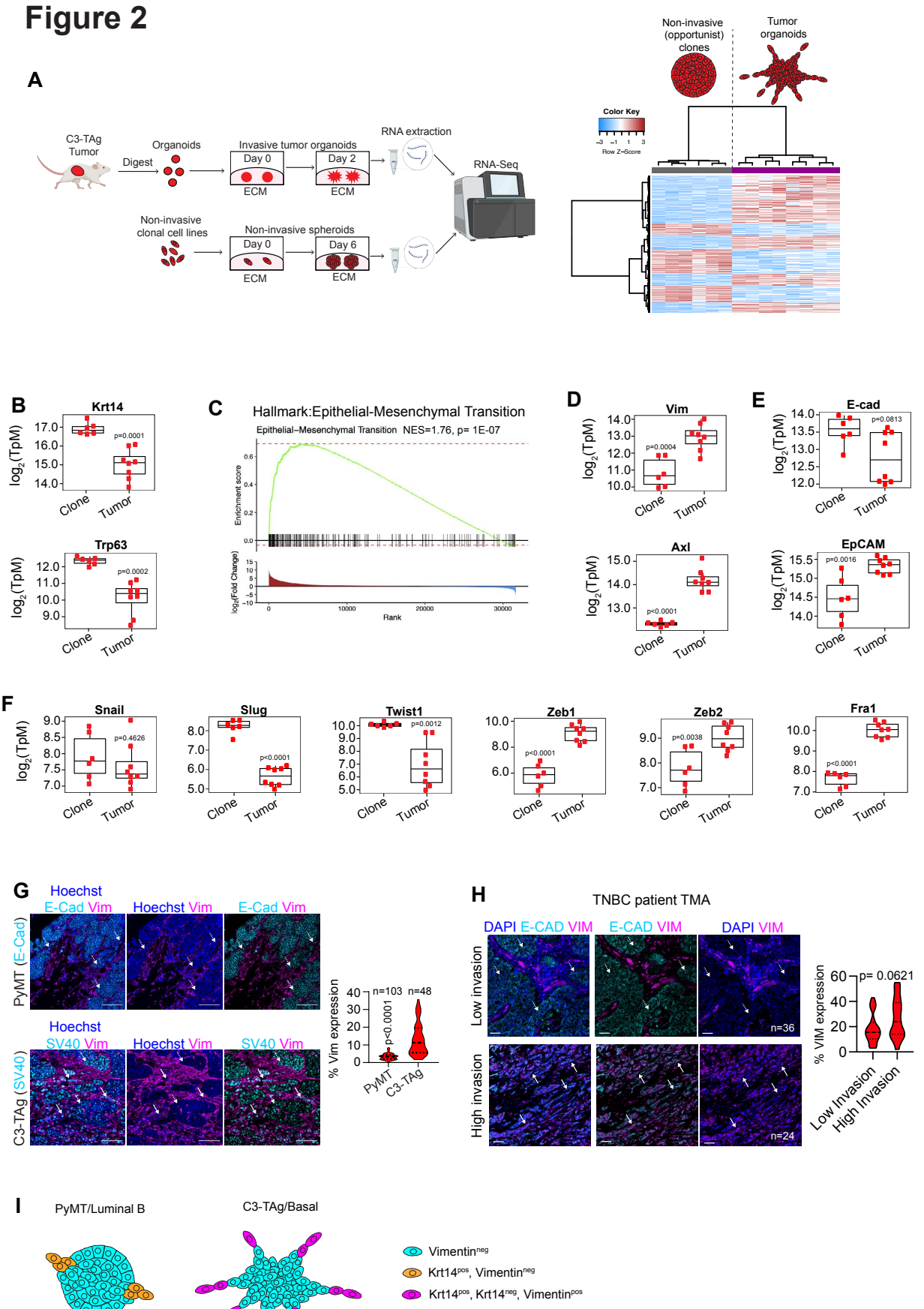

A
C3-TAg Tumor
Digest
Organoids
Invasive tumor organoids
Day 0
Day 2
ECM
ECM
RNA extraction
RNA-Seq
Non-invasive clonal cell lines
Non-invasive spheroids
Day 0
Day 6
ECM
ECM
Non-invasive (opportunist) clones
Tumor organoids
Color Key
Row Z-Score
B
Krt14
p=0.0001
Trp63
p=0.0002
Clone
Tumor
C
Hallmark:Epithelial-Mesenchymal Transition
Epithelial–Mesenchymal Transition NES=1.76, p= 1E-07
Enrichment score
Rank
D
Vim
p=0.0004
Axl
p<0.0001
E
E-cad
p=0.0813
EpCAM
p=0.0016
F
Snail
p=0.4626
Slug
p<0.0001
Twist1
p=0.0012
Zeb1
p<0.0001
Zeb2
p=0.0038
Fra1
p<0.0001
G
Hoechst
E-Cad Vim
Hoechst Vim
E-Cad Vim
PyMT (E-Cad)
Hoechst
SV40 Vim
Hoechst Vim
SV40 Vim
C3-TAg (SV40)
% Vim expression
n=103
n=48
p<0.0001
PyMT
C3-TAg
H
TNBC patient TMA
DAPI E-CAD VIM
E-CAD VIM
DAPI VIM
Low invasion
High invasion
n=36
n=24
% VIM expression
p= 0.0621
Low Invasion
High Invasion
I
PyMT/Luminal B
C3-TAg/Basal
Vimentin^neg
Krt14^pos, Vimentin^neg
Krt14^pos, Krt14^neg, Vimentin^pos

Figure 2. Trailblazer cells are distinguished by a subset of EMT traits in basal-like mammary tumors. **A.** Graphic shows preparation of C3-TAg organoids and clonal cell line spheroids for RNA-seq. Heatmap shows unsupervised clustering and differentially expressed genes in non-invasive C3-TAg clones (B6, C6, E8) and organoids from 4 C3-TAg tumors. **B.** Boxplots showing the expression of the basal genes Krt14 and Trp63. **C.** There was increased expression of Hallmark EMT genes in C3-TAg tumor organoids compared to non-invasive C3-TAg clones. Plot shows the details of the Hallmark EMT generated by GSEA. **D-F.** Boxplots showing the expression of canonical mesenchymal (Vimentin, Axl), epithelial (E-cadherin, EpCAM) and EMT-TF (Snai1, Slug, Twist1, Zeb1, Zeb2, Fra1) genes in C3-TAg tumor organoids compared to non-invasive clonal spheroids from RNA-seq data (Mann Whitney test for E-cad, Student's t-test for all others, n=6 for clones, (3 biological replicates); n=8 for tumor (4 biological replicates)). **G.** Vimentin is more highly expressed in C3-TAg primary tumors compared to PyMT tumors. Solid arrows indicate collectively invading Vimentin^{pos} tumor cells. Dashed arrows indicate Vimentin^{neg} tumor cells. Violin plot shows percentage of tumor cell area expressing Vimentin (Mann Whitney Test, n=ROIs from 5 tumors). Scale bars, 100 µm. **H.** Invasive regions in TNBC patient tumors trend towards higher Vimentin expression compared to regions of low invasion. Solid arrows indicate collectively invading Vimentin^{pos} tumor cells. Dashed arrows indicate Vimentin^{neg} tumor cells. Violin plot shows percentage of Vimentin expressing tumor cells (Mann Whitney test, n=tumor sections). Scale bars, 50 µm. **I.** Model showing the tumor subtype relationship between an EMT exemplified by Vimentin expression and trailblazer activity.

198

199 **Table S2)**. Genes highly expressed in C3-TAg tumor organoids were associated with biological
200 processes that contribute to invasion (ECM Receptor Interaction, Focal Adhesions and Regulation
201 of the Actin Cytoskeleton) and EMT **(Fig. 2C,D, S2C,D, Table S2)**. The C3-TAg tumor organoids
202 retained canonical epithelial gene expression as well, indicating that the majority of cells were in
203 a hybrid EMT state (**Fig. 2E, Table S2)**. The EMT inducing transcription factors (EMT-TFs)
204 Zeb1, Zeb2 and Fra1 were more highly expressed in C3-TAg organoids **(Fig. 2F)**. However, the
205 EMT-TFs Snail, Slug and Twist1 were expressed at similar or higher levels in the noninvasive
206 clones **(Fig. 2F)**. These results indicated that the specific features of an EMT program determined
207 the invasive phenotype of C3-TAg tumor cells, rather than simply the presence or absence of EMT
208 activity, consistent with results obtained analyzing human breast cancer cell lines (25, 31).
209

210 To understand the relationship between the EMT program features of the C3-TAg organ-
211 oids and collective invasion, we evaluated the expression of the microfilament protein Vimentin,
212 a common EMT marker that was more highly expressed in C3-TAg organoids **(Fig. 2D)**. Vimentin

213 expression was more frequently and extensively detected in C3-TAg tumors compared to PyMT
214 tumors **(Fig. 2H)**. The heterogeneity of Vimentin expression in C3-TAg tumors was consistent
215 with the heterogeneity in the extent of C3-TAg organoid invasion **(Fig. 1A)**. Vimentin^{pos} and Vi-
216 mentin^{neg} cells also collectively invaded together in C3-TAg primary tumors and organoids, sug-
217 gesting that there was heterogeneity in collectively invading cell populations with respect to EMT
218 state **(FIg. 2H, S2D)**. In addition, Vimentin expression was heterogeneous in TNBC sections in a
219 TMA, with a trend towards more extensive Vimentin expression in regions of collective invasion
220 **(Fig. 2I, S2E)**. Vimentin expression was higher in the basal type compared to other subtypes of
221 breast cancer as well, indicating that the difference in trailblazer program activation in C3-TAg
222 (basal-like) and PyMT (luminal B) tumors reflected a feature of inter-tumor heterogeneity found
223 in human breast cancer **(Fig. S2F)**. Our combined results demonstrated that an EMT program ex-
224 emplified by Vimentin expression conferred a trailblazer state in TNBC tumor subpopulations
225 **(Fig. 2J)**.
226
227 **TGFβ induces a trailblazer state in C3-TAg tumors.**
228 Gene set enrichment analysis (GSEA) indicated that TGFβ signaling was more active in
229 C3-TAg tumor organoids than non-invasive spheroids **(Fig. 3A, Table S2)**. Thus, we evaluated
230 the contribution of Tgfβ towards regulation of a trailblazer state. The pharmacological inhibitor of
231 Tgfβ receptor 1 (Tgfbr1) A83-01 suppressed C3-TAg tumor organoid invasion **(Fig. 3B)**. 1863
232 cells derived in the presence of A83-01 were also less invasive than 1863T cells established from
233 the same tumor in the absence of A83-01 **(Fig. S3A, Table S1)**. Depletion of the Tgfβ regulated
234 transcription factor SMAD3 reduced the invasion of 1863T and Tgfβ stimulated 1863 cells, indi-
235 cating that TGFβ regulated gene expression contributed to the observed phenotypes **(Fig. 3C,**

Figure 3

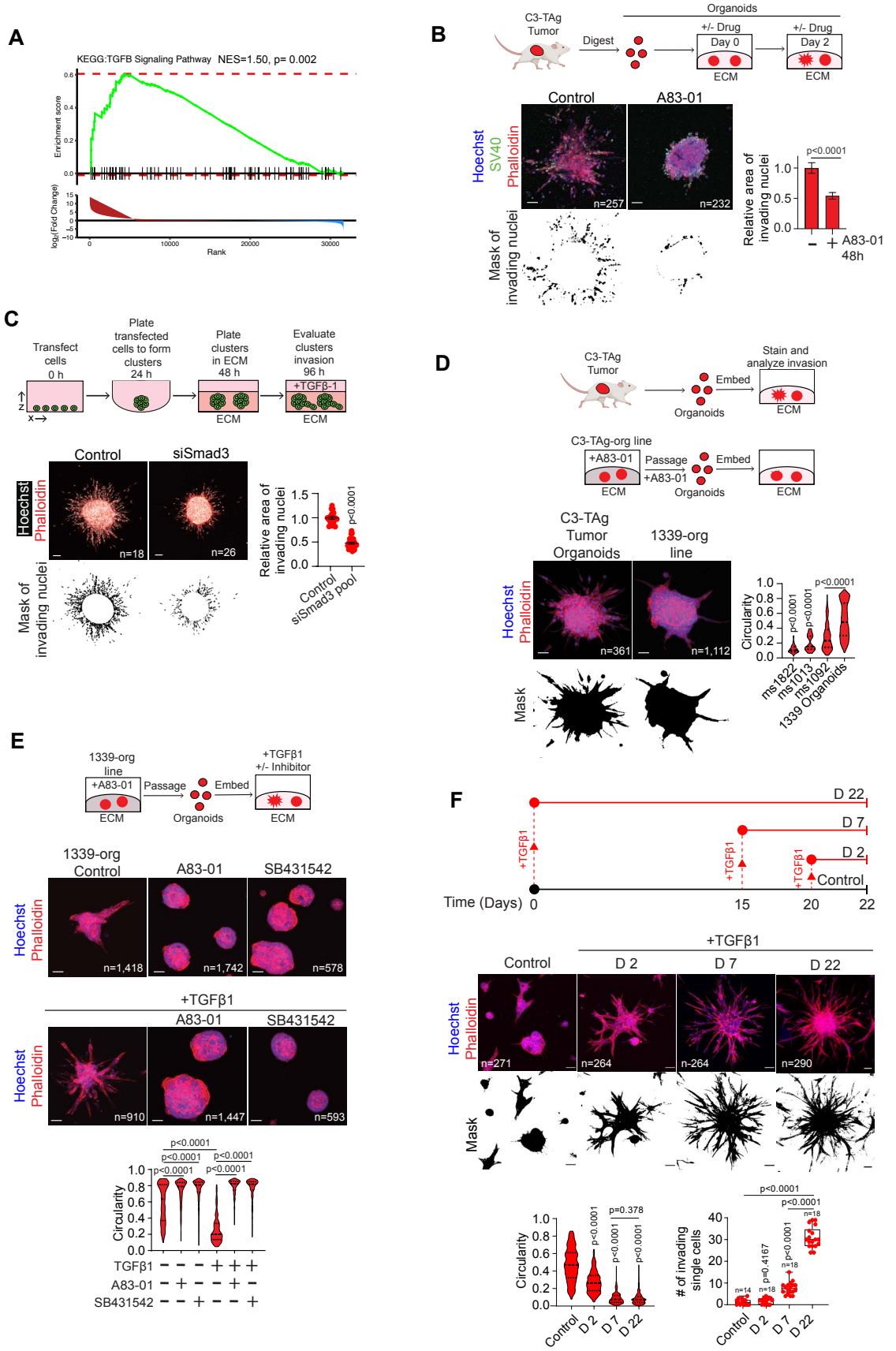

A
KEGG:TGFB Signaling Pathway NES=1.50, p= 0.002
Enrichment score
Rank
B
C3-TAg Tumor
Digest
Organoids
+/- Drug
Day 0
Day 2
ECM
Control
A83-01
Hoechst
SV40
Phalloidin
n=257
n=232
Mask of invading nuclei
Relative area of invading nuclei
p<0.0001
A83-01 48h
C
Transfect cells 0 h
Plate transfected cells to form clusters 24 h
Plate clusters in ECM 48 h
Evaluate clusters invasion 96 h
+TGFβ-1
ECM
Control
siSmad3
n=18
n=26
Relative area of invading nuclei
p<0.0001
Control
siSmad3 pool
D
C3-TAg Tumor
Embed
Organoids
Stain and analyze invasion
C3-TAg-org line
+A83-01
Passage
C3-TAg Tumor Organoids
1339-org line
n=361
n=1,112
Circularity
p<0.0001
ms1822
ms1013
ms1092
1339 Organoids
Mask
E
1339-org line
+A83-01
Passage
Embed
+TGFβ1 +/- Inhibitor
1339-org Control
A83-01
SB431542
n=1,418
n=1,742
n=578
+TGFβ1
n=910
n=1,447
n=593
Circularity
p<0.0001
TGFβ1
A83-01
SB431542
F
D 22
D 7
D 2
Control
+TGFβ1
Time (Days)
0
15
20
22
n=271
n=264
n=264
n=290
Mask
Circularity
p=0.378
# of invading single cells
p=0.4167
n=14
n=18
Control
D 2
D 7
D 22

Figure 3. Tgfβ induces trailblazer state in C3-TAg tumors.
A. GSEA was performed on RNA-seq data from C3-TAg organoids and non-invasive C3-TAg clones to identify pathway enrichment in C3-TAg organoids. Plot shows the GSEA for the KEGG: TGFβ signaling pathway. **B.** The Tgfbr1 inhibitor A83-01 suppresses C3-TAg organoid invasion. Graph shows the quantification of invasion (mean±SEM, Mann Whitney test, n=organoids). Scale bars, 50μm. **C.** Smad3 siRNA suppressed the invasion of TGFβ1 treated 1863 clusters (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 μm. **D.** 1339-org cells cultured in A83-01 were less invasive than organoids tested immediately after isolation from primary tumors. Graph shows organoid circularity. Reduced circularity indicates increased invasion (mean±SEM, Mann Whitney test, n=organoids). Scale bars, 50 μm. **E.** Tgfbr1 inhibition suppresses intrinsic and extrinsic TGFβ1 induced invasion of 1339-org cells. Organoids were treated with the Tgfbr1 inhibitors A83-01 (500 nM) and SB431542 (1 μM) and extrinsic TGFβ1 (2 ng/ml) as indicated. Violin plots show the quantification of organoid circularity (Mann Whitney test, n=organoids). Scale bars, 50 μm. **F.** Persistent exposure to exogenous TGFβ1 progressively enhances the invasive phenotype in C3-TAg organoid lines. Graphs show the quantification of organoid circularity and the number of invading single cells (Mann Whitney test, n=organoids, r=3) Scale bars, 50 μm.

S3B,C). These results indicated that Tgfβ signaling was necessary to sustain a trailblazer state in C3-TAg cells.

To determine if Tgfβ could initiate the conversion from an opportunist to a trailblazer state, we established organoid lines from 3 different tumors in the presence of A83-01 for at least 4 weeks (**Fig. S3D, Table S1**). These conditions yielded 1339-org, 1863-org and 1788-org tumor organoid lines that had a diminished Tgfβ activity and a reduced ability to invade relative to tumor organoids tested immediately after isolation from C3-TAg tumors (**Fig. 3D**). Removal of A83-01 or SB431542, a second inhibitor of Tbfgr1, promoted tumor organoid line invasion (**Fig. 3E, S3E,F**). The invasion of the tumor organoid lines was further enhanced by exogenous TGFβ1 within 48 h (**Fig. 3E, S3E,F, Videos S9-11**). Persistent exposure of 1339-org cells to Tgfβ signaling for up to 22 days progressively increased the rate of collective invasion (**Fig. 3F, Videos S9-12**). TGFβ signaling for at least 7 days also increased the frequency that trailblazer cells detached from collectively invading cohorts and invaded as single cells, suggesting that at least some single cell invasion reflected a loss of cohesion rather than the activation of a separate invasion program

253 **(Fig. 3F, Video S12)**. Thus, Tgfβ induced signaling promoted the conversion of C3-TAg cells to
254 a trailblazer state.
255
256 **TGFβ induces multiple trailblazer states through distinct phases of re-programming.**
257 To understand how Tgfβ induced a trailblazer state, we used RNA-sequencing (RNA-seq)
258 to analyze gene expression changes 1339-org cells after 2, 7 and 22 days after Tgfβ pathway acti-
259 vation. Hierarchical clustering revealed 4 different groups of genes regulated by TGFβ **(Fig. 4A,**
260 **Table S3).** If cells were undergoing binary conversion between states, the population averaged
261 mRNA expression obtained by RNA-seq would have shown 2 groups of genes (suppressed and
262 induced) undergoing a gradual change in expression as percentage of cells in the 2 different states
263 shifted (39). The 4 different gene groups, each with a unique timing in their onset of change and
264 maximal expression difference, therefore indicated that the Tgfβ stimulated C3-TAg tumor cells
265 progressed through different stages of re-programming. Consistent with inducing multiple differ-
266 ent states, TGFβ promoted a shift in the distribution of EpCAM expression rather than a switch
267 between two unique EpCAM-high and EpCAM-low states, **(Fig. S4A)**. TGFβ also induced a shift
268 in the distribution of Vimentin expression intensity in organoids at day 7 and day 22, as opposed
269 to a switch between two distinct Vimentin^{neg} and Vimentin^{pos} states **(Fig. S4B)**.
270
271 The gene groups identified by RNA-seq were associated with distinct combinations of bi-
272 ological processes **(Fig. 4B, Table S3)**. Of note, genes in Groups 3 and 4 were involved with ECM
273 regulation **(Fig. 4B, Table S3)**. This indicated that there were different mechanisms for regulating
274 genes associated with the same biological functions in response to TGFβ. Group 3 and 4 genes
275 were also connected with unique functions, which indicated that the nature of reprogramming

Figure 4

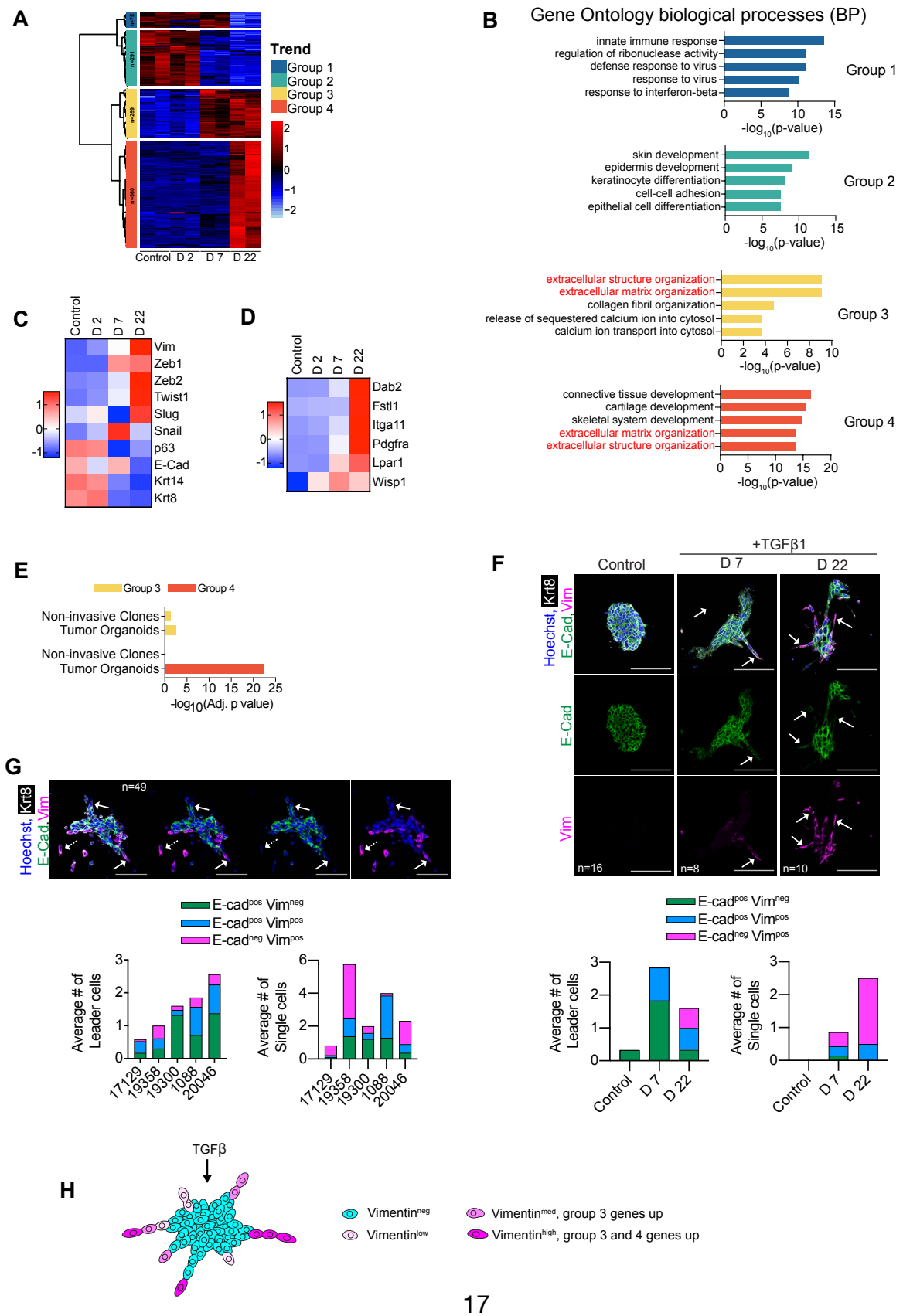

A
Trend
Group 1
Group 2
Group 3
Group 4
Control D 2 D 7 D 22
B
Gene Ontology biological processes (BP)
innate immune response
regulation of ribonuclease activity
defense response to virus
response to virus
response to interferon-beta
Group 1
-log10(p-value)
skin development
epidermis development
keratinocyte differentiation
cell-cell adhesion
epithelial cell differentiation
Group 2
extracellular structure organization
extracellular matrix organization
collagen fibril organization
release of sequestered calcium ion into cytosol
calcium ion transport into cytosol
Group 3
connective tissue development
cartilage development
skeletal system development
extracellular matrix organization
extracellular structure organization
Group 4
C
Control D 2 D 7 D 22
Vim
Zeb1
Zeb2
Twist1
Slug
Snail
p63
E-Cad
Krt14
Krt8
D
Dab2
Fstl1
Itga11
Pdgfra
Lpar1
Wisp1
E
Group 3 Group 4
Non-invasive Clones
Tumor Organoids
-log10(Adj. p value)
F
+TGFβ1
Control D 7 D 22
Hoechst, E-Cad, Vim
Krt8
E-Cad
Vim
n=16
n=8
n=10
G
Hoechst, E-Cad, Vim
Krt8
n=49
E-cadpos Vimneg
E-cadpos Vimpos
E-cadneg Vimpos
Average # of Leader cells
Average # of Single cells
17129 19358 19300 1088 20046
Control D 7 D 22
H
TGFβ
Vimentinneg
Vimentinlow
Vimentinmed, group 3 genes up
Vimentinhigh, group 3 and 4 genes up

Figure 4. Tgfβ induces multiple trailblazer states through distinct phases of re-programming. **A.** Tgfβ induces 4 distinct patterns of gene expression changes in 1339-org organoids. Heatmap shows unsupervised clustering and 4 groups of DEGs. **B.** Graphs show the top 5 Gene Ontology terms associated with each group. **C.** Heatmap shows changes in expression of mesenchymal (Vim), EMT-TFs (Zeb1 and 2, Twist1, Snail, Slug), epithelial (E-Cad), basal (p63, Krt14) and luminal (Krt8) markers after Tgf. **D.** Heatmap shows persistent TGFβ pathway activation increases the expression of trailblazer genes previously reported in human cell lines. **E.** Graph shows the enrichment of Group 3 and Group 4 genes in non-invasive clones and invasive C3-TAg tumor organoids. **F, G.** Multiple different trailblazer states lead collective invasion in C3-TAg tumor and TGFβ1 treated 1339-org organoids. Paraffin sections were immunostained with Vimentin and E-cadherin antibodies to define different EMT states. The phenotype of trailblazer cells leading invasion or invading as single cells was quantified. Bar graph indicates the mean number of trailblazer cells or single invading cells per organoids and their EMT phenotype. **H.** Model showing the heterogeneity in trailblazer states induced by Tgfβ.

changed over time and was controlled by distinct mechanisms. Interestingly, genes commonly associated with EMT were included in Group 3 (Vcam1, Itgb3, Fn1) and Group 4 (Vimentin, N-cadherin, Periostin) **(Table S3)**. EMT-TFs were also split between Group 3 (Zeb1), Group 4 (Zeb2, Twist), or not included in any group (Snail) **(Fig. 3C)**. Previously defined trailblazer genes were in Group 3 (Wisp1) and Group 4 (Dab2, Fstl1, Itga11, Pdgfra, Lpar1) **(Fig. 4D)**. Group 3 and Group 4 genes were enriched in the genes that were highly expressed in C3-TAg organoids, which indicated that Tgfβ re-activated a program that was present in C3-TAg tumors **(Fig. 2A, 4E, Table S3)**. The distinct expression profile of EMT-TFs, EMT effectors, ECM regulatory genes and different indicated that TGFβ induced re-programming yielded multiple different trailblazer states in response to evolving regulatory mechanisms. To test this possibility, we immunostained 1339-org and C3-TAg tumor organoids with Vimentin and E-Cadherin antibodies, to specify different states within the spectrum of EMT re-programming. Vimentin^{neg}/E-cadherin^{pos}, Vimentin^{pos}/E-cadherin^{pos} and Vimentin^{pos}/E-cadherin^{neg} cells were all capable of leading collective invasion (**Fig. 4F,G**). Notably, the presence of more Vimentin^{pos}/E-cadherin^{pos} and Vimentin^{pos}/E-cadherin^{neg} trailblazer cells with increasing time after Tgfβ pathway activation correlated with an increased extent of 1339-org collective invasion (**Fig. 3F**). These combined results suggested that there were multiple

294 trailblazer states induced by Tgfβ and that they had different capacities to initiate collective inva-
295 sion **(Fig. 4H)**.
296
297 **The transcription factor Fra1 confers a trailblazer state.**
298 We next investigated how EMT-TF expression influenced the trailblazer state. To prioritize
299 our investigation, we first determined how siRNAs targeting EMT-TFs influenced the invasion of
300 the 1863T cell line, which contains a high proportion of trailblazer cells **(Fig. S3A)**. The siRNAs
301 targeting Fra1, Zeb1 and Zeb2 suppressed invasion to the greatest extent whereas siRNAs targeting
302 Twist, Slug and Snail failed to suppress invasion, consistent with Fra1, Zeb1 and Zeb2 being the
303 EMT-TFs that were more highly expressed in C3-TAg tumor organoids **(Fig. 5A, S4A)**. Fra1 ex-
304 pression was higher in basal like breast cancer whereas Zeb1 expression was lower in basal-like
305 tumors and Zeb2 expression was similar across tumor subtype **(Fig. 5B)**. Fra1 expression also
306 correlated with Vimentin expression in breast cancer patients, suggesting that Fra1 potentially
307 contributes to induction of a Vimentin specified trailblazer state during tumor progression **(Fig.**
308 **5C)**. We therefore selected Fra1 for further analysis.
309
310 Fra1 depletion by 4 different siRNAs and 2 different siRNA pools reduced 1863T and
311 TGFβ stimulated 1863 invasion to a similar degree **(Fig. 5D,E, S5A-D)**. Fra1 depletion also re-
312 duced the invasion of D6 trailblazer clones and Tgfβ stimulated 1339 cells, which were derived
313 from a different C3-TAg tumor than the 1863 and 1863T cells **(Fig. S5F)**. Fra1 interacts with Jun
314 to regulate gene expression as an AP-1 complex (40). Indeed, Jun was required for 1863T and
315 Tgfβ stimulated 1863 cell invasion, suggesting Fra1 containing AP-1 complexes regulate the ex-
316 pression of genes necessary for the induction of a trailblazer state **(Fig. 5F,H, S5G)**. Exogenous

Figure 5

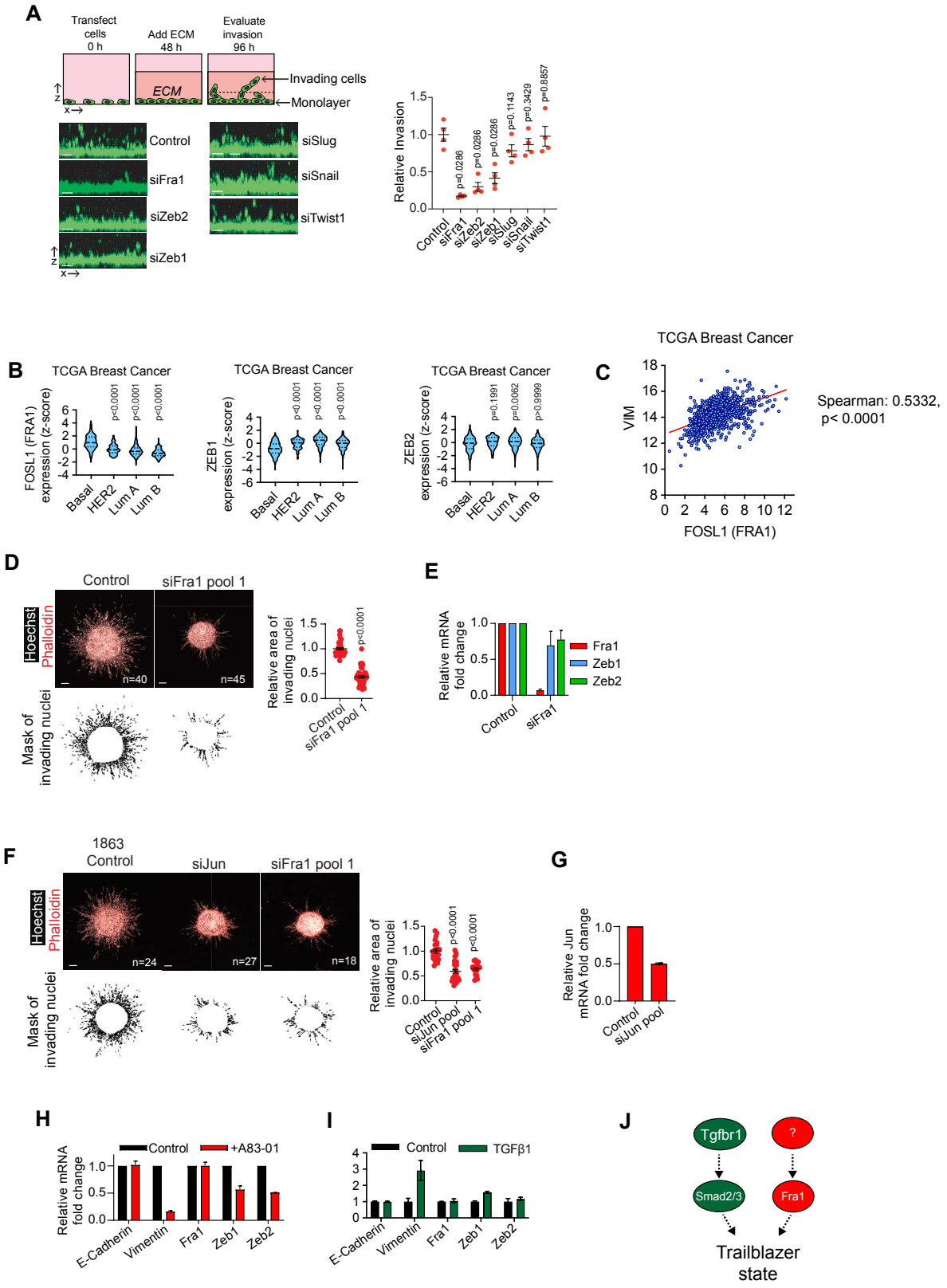

A
Transfect cells 0 h
Add ECM 48 h
Evaluate invasion 96 h
ECM
Invading cells
Monolayer
Control
siFra1
siZeb2
siZeb1
siSlug
siSnail
siTwist1
Relative Invasion
p=0.0286
p=0.0286
p=0.0286
p=0.1143
p=0.3429
p=0.8857
B
TCGA Breast Cancer
FOSL1 (FRA1) expression (z-score)
ZEB1 expression (z-score)
ZEB2 expression (z-score)
p<0.0001
p=0.1991
p=0.0062
p>0.9999
Basal
HER2
Lum A
Lum B
C
TCGA Breast Cancer
VIM
FOSL1 (FRA1)
Spearman: 0.5332, p< 0.0001
D
Control
siFra1 pool 1
Hoechst
Phalloidin
n=40
n=45
Mask of invading nuclei
Relative area of invading nuclei
p<0.0001
E
Relative mRNA fold change
Fra1
Zeb1
Zeb2
F
1863 Control
siJun
siFra1 pool 1
n=24
n=27
n=18
siJun pool
G
Relative Jun mRNA fold change
H
+A83-01
E-Cadherin
Vimentin
I
TGFβ1
J
Tgfbr1
?
Smad2/3
Fra1
Trailblazer state

Figure 5. The transcription factor Fra1 confers a trailblazer state.
A. Invasion of 1863T cells transfected with siRNAs targeting EMT-TFs. Graph shows the quantification of cells invading ≥40 μm into the ECM (mean±SEM, Mann Whitney test, n=4 wells from 2 independent experiments). Scale bars, 50 μm. **B.** FRA1 (FOSL1) mRNA is more highly expressed in the basal-like subtype compared to other subtypes of breast cancer. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to the basal-like tumors. **C.** FRA1 (FOSL1) correlates with VIM expression in breast cancer patient tumors. **D.** Fra1 depletion decreases the invasion of TGFβ1 treated 1863 clusters (mean±SEM, Mann Whitney test n=clusters, r=5) Scale bars, 50 μm. **E.** Fra1 depletion does not suppress Zeb1 and Zeb2 expression (mean±SEM, n=2) **F.** Jun depletion suppresses TGFβ1 treated 1863 cluster invasion (mean±SEM, Mann Whitney test, n=clusters, r=2) Scale bars, 50 μm. **G.** Depletion of Jun by siRNAs (mean±SEM, n=2). **H.** E-cadherin, Vimentin, Fra1, Zeb1 and Zeb2 expression after 48 h treatment with Tgfbr1 inhibitor A83-01 (mean±SEM, n=2). **I.** Expression of E-cadherin, Vimentin, Fra1, Zeb1 and Zeb2 48 h after TGFβ1 treatment. **J.** Model showing the parallel regulation of the trailblazer state by the Tgfβ signaling and a second pathway that induces Fra1.

Fra1 expression promotes Zeb1 and Zeb2 expression (41). However, Fra1 depletion in 1863T cells did not alter Zeb1 or 2 expression **(Fig. 5E)**. Zeb1 and Vimentin expression were regulated by autocrine TGFβ signaling and induced by acute or persistent exogenous TGFβ treatment (**Fig. 4C, 5I,H)**. Zeb2 was included in the Group 4 Tgfβ regulated genes **(Fig. 4C)**. However, Fra1 expression was not influenced by autocrine or extrinsic TGFβ pathway stimulation **(Fig. 5I,H)**. Fra1 was also not included in Group 3 or 4 genes detected as part of a persistent TGFβ induced gene regulatory program **(Fig. 4C and Table S3)**. Thus, the trailblazer state required the activity of at least one additional signaling pathway acting in parallel with TGFβ signaling to induce the expression of Fra1 **(Fig. 5J)**.

Egfr pathway activity promotes Fra1 expression and a trailblazer state.

We next determined how Fra1 expression was regulated in C3-TAg tumor cells. Fra1 mRNA was more highly expressed in C3-TAg tumor organoids compared to opportunist spheroids in our RNA-seq analysis, indicating that the signaling pathway responsible for the control of Fra1

333 was more active in the C3-TAg organoids **(Fig. 2A)**. GSEA identified multiple active KRAS reg-
334 ulated expression signatures in the C3-TAg tumor organoids **(Fig. S6A,B, Table S2)**. The Kras
335 gene can be amplified, but is not frequently mutated in C3-TAg tumors , indicating that pathway
336 activation occurred upstream of Kras (42). Analysis of the genes in the active KRAS signature
337 showed that the Egfr ligand Hbegf was more highly expressed in C3-TAg organoids than oppor-
338 tunist clones **(Fig. 6A, Table S2)**. Like FRA1, HBEGF was more highly expressed in Basal type
339 breast tumors **(Fig. S6C)**. HBEGF expression correlated with Fra1 expression in breast cancer
340 patient tumors as well, consistent with the possibility that Hbegf promotes Fra1 expression during
341 tumor development **(Fig. S6D)**. Further inspection found that the Egfr ligands Areg and Epgn were
342 also expressed at a higher level in C3-TAg organoids **(Fig. 6A, Table S2)**. These combined results
343 prompted us to investigate how the Egfr-Ras-Mek1/2-Erk1/2 signaling pathway contributed to
344 Fra1 expression and the induction of a trailblazer state. Erlotinib (Egfr inhibitor) and trametinib
345 (Mek1/2 inhibitor) suppressed Fra1 mRNA and protein expression in C3-TAg cell lines and or-
346 ganoids **(Fig. 6B,C, S6E)**. By comparison, Jun expression was unchanged by treatment with either
347 inhibitor **(Fig. S6F)**. In addition, Erlotinib and Trametinib did not inhibit the activating phosphor-
348 ylation of Smad2 or Zeb1 expression, indicating that Fra1 did not influence cell state through
349 autocrine Tgfβ signaling **(Fig. 6C, S6G)**. A83-01 suppressed Smad2 phosphorylation, but did not
350 influence Fra1 expression and phosphorylation **(Fig. 6C)**. Erlotinib and Trametinib reduced the
351 invasion of TGFβ stimulated 1339-org cells **(Fig. 6D,E)**. EGF also enhanced the invasion of 1339-
352 org, 1863-org and 1788-org organoids and this enhancement was blocked by A83-01, further in-
353 dicating that the integration of Tgfbr1 and Egfr pathways promoted a trailblazer state **(Fig. 6F,**
354 **S6H,I)**. Trametinib inhibited the invasion of C3-TAg primary tumor organoids and phosphory-
355 lated Erk1/2 was detected in collectively invading cells in C3-TAg primary tumors **(Fig. 6G,H)**.

Figure 6

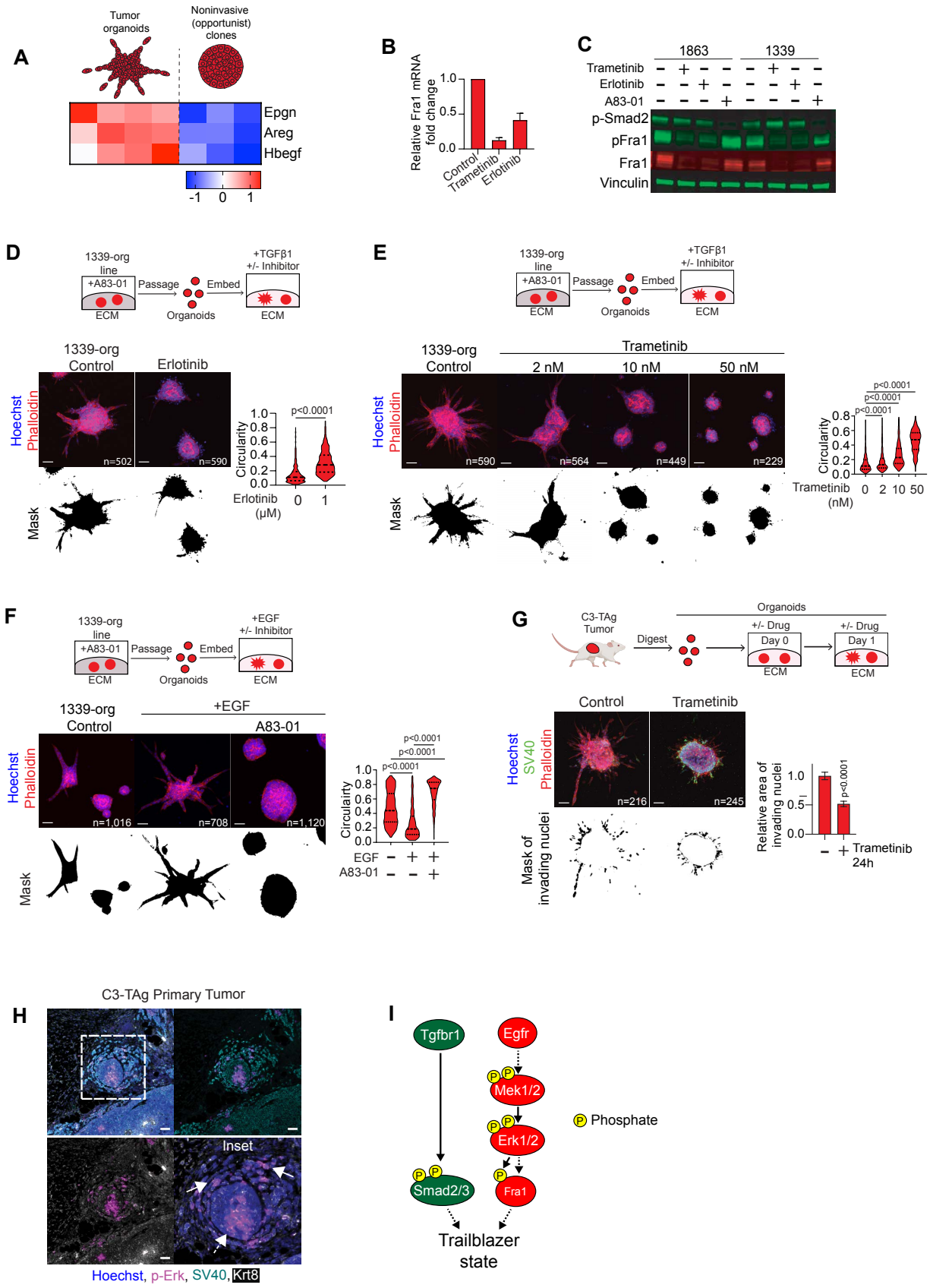

A
Tumor organoids
Noninvasive (opportunist) clones
Epgn
Areg
Hbegf
-1 0 1
B
Relative Fra1 mRNA fold change
1.0
0.5
0.0
Control
Trametinib
Erlotinib
C
1863
1339
Trametinib
Erlotinib
A83-01
p-Smad2
pFra1
Fra1
Vinculin
D
1339-org line
+A83-01
Passage
Embed
+TGFβ1 +/- Inhibitor
ECM
Organoids
ECM
1339-org
Control
Erlotinib
Hoechst
Phalloidin
n=502
n=590
Mask
Circularity
p<0.0001
Erlotinib (μM)
0
1
E
1339-org line
+A83-01
Passage
Embed
+TGFβ1 +/- Inhibitor
ECM
Organoids
ECM
1339-org
Control
Trametinib
2 nM
10 nM
50 nM
Hoechst
Phalloidin
n=590
n=564
n=449
n=229
Mask
Circularity
p<0.0001
p<0.0001
p<0.0001
Trametinib (nM)
0 2 10 50
F
1339-org line
+A83-01
Passage
Embed
+EGF +/- Inhibitor
ECM
Organoids
ECM
1339-org
Control
+EGF
A83-01
Hoechst
Phalloidin
n=1,016
n=708
n=1,120
Mask
Circularity
p<0.0001
p<0.0001
p<0.0001
EGF
A83-01
G
C3-TAg Tumor
Digest
Organoids
+/- Drug
Day 0
+/- Drug
Day 1
ECM
ECM
Control
Trametinib
Hoechst
SV40
Phalloidin
n=216
n=245
Mask of invading nuclei
Relative area of invading nuclei
p<0.0001
Trametinib 24h
H
C3-TAg Primary Tumor
Inset
Hoechst, p-Erk, SV40, Krt8
I
Tgfbr1
Egfr
Mek1/2
Erk1/2
Smad2/3
Fra1
Phosphate
Trailblazer state

Figure 6. The transcription factor Fra1 confers a trailblazer state.
A. Expression of the Egfr ligands Epigen (Epgn), Amphiregulin (Areg) and Heparin-binding epidermal growth factor (Hbegf) in C3-TAg organoids and non-invasive clones. **B.** The Mek1/2 inhibitor trametinib (10 nM) and Egfr inhibitor erlotinib (1 μM) reduces Fra1 expression in 1339-org organoids (mean±SEM, n=2). **C.** Immunoblot showing the response of 1863 and 1339 cells to trametinib, erlotinib and A83-01 treatment. **D.** Erlotinib (1 μM) decreases the TGFβ1 induced invasion 1339-org organoids. Violin plot shows the quantification of circularity (Mann Whitney test, n=organoids, r=3) Scale bars, 50 μm. **E.** Trametinib suppresses the TGFβ1 induced invasion of 1339-org organoids in a dose dependent manner. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=3). Scale bars, 50 μm. **F.** A83-01 suppresses EGF induced invasion in the 1339-org organoids. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=4) Scale bars, 50 μm. **G.** Trametinib suppresses C3-TAg tumor organoid invasion. Graph shows the quantification of invading nuclei from 4 different tumors (mean±SEM, Mann Whitney test, n=organoids) Scale bars, 50 μm. **H.** Images show p-Erk immunostaining in regions of invasion (solid arrows) and no invasion (dotted arrow) in primary C3-TAg tumors. Scale bars, 50 μm. **I.** Model showing the parallel regulation of trailblazer cells by signaling pathways coordinated by Tgfbr1 and Egfr.

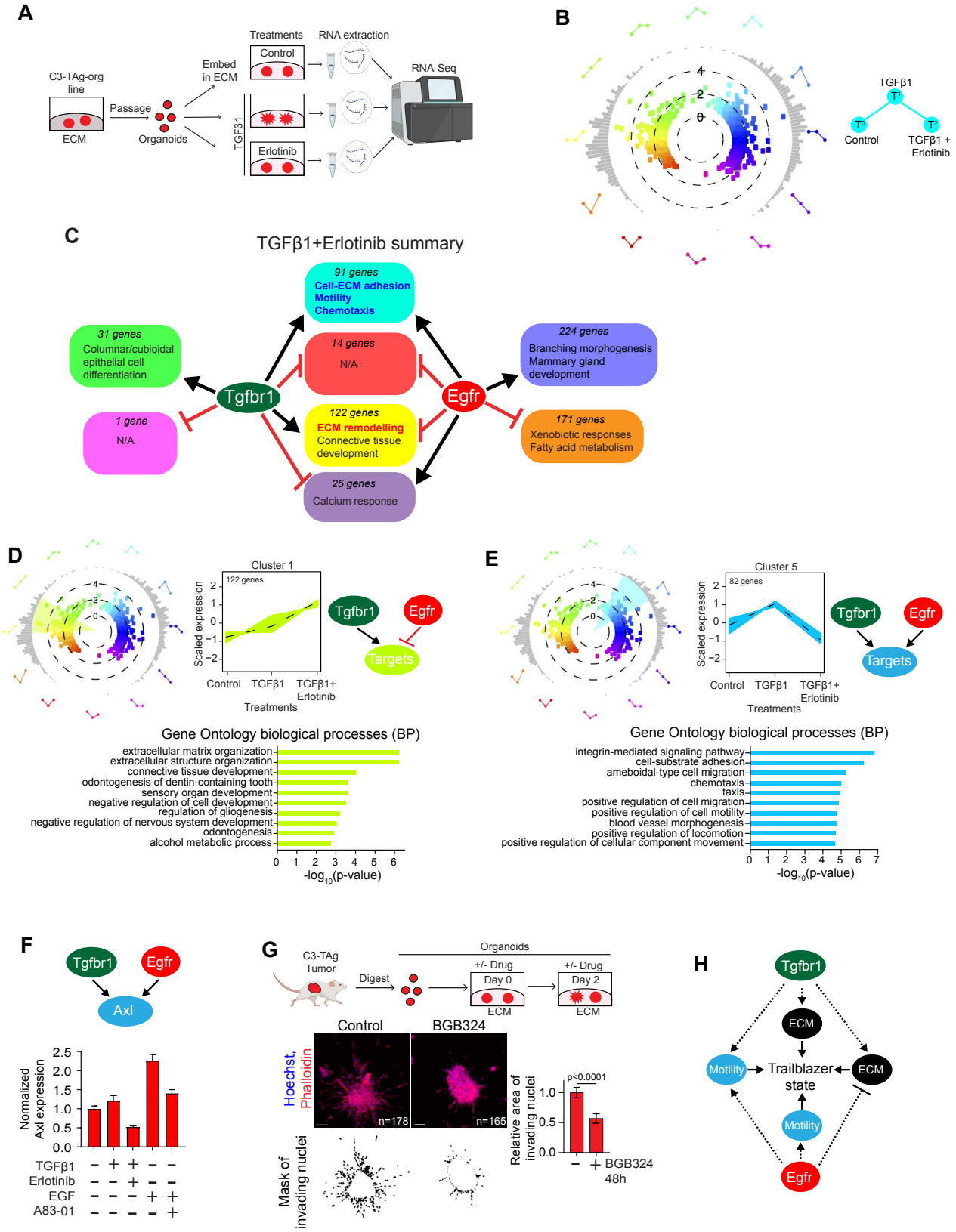
Erk1/2 activity is pulsatile in response to EGF stimulation, and these pulses promote Fra1 expression (43). Thus, persistent Erk1/2 phosphorylation was not necessarily expected in all collectively invading cells or to promote Fra1 expression. Together, these results indicate that the regulation of Fra1 by Egfr and Erk1/2 promotes the induction of a trailblazer EMT **(Fig. 6I)**.

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

To understand the contribution of the Tgfbr1 and Egfr regulated signaling pathways towards the induction of the trailblazer state, we performed RNA-seq on organoids treated with diluent, TGFβ, TGFβ+erlotinib, EGF and EGF+A83-01 **(Fig. 7A, S7A)**. Polar plots were used to cluster genes based on the nature and the degree of integration between the Tgfbr1 and Egfr signaling pathways (44). This analysis revealed multiple modes of interaction between the two pathways in response to either TGFβ or EGF and that the type of interaction (positive or negative) was associated with different biological processes **(Fig. 7B,C, S7B-E, Tables S4,5)**. Thus, the relationship was far more complex than the simple positive signal integration previously suggested by

373 phenotypic analysis and the expression of specific subsets of EMT-TFs and canonical EMT pro-
374 gram genes (45, 46). In addition, the gene expression changes induced by extrinsic TGFβ and EGF
375 were associated with unique combinations of biological processes, indicating that extrinsic TGFβ
376 and extrinsic EGF induced distinct types of trailblazer cells **(Fig. 7C, S7F, Tables S4,5)**. Never-
377 theless, there were conserved features of pathway integration that influenced the regulation of pro-
378 cesses associated with invasion regardless of the external stimulus. Genes involved in controlling
379 the organization and composition of the ECM, including previously defined trailblazer genes, were
380 predominantly induced by Tgfbr1 and not positively impacted by Egfr signaling **(Fig. 7C,D,**
381 **S7F,G, Tables S4,5).** Notably, autocrine and extrinsic Egfr activation suppressed a number of
382 Tgfbr1 regulated ECM reorganization genes and Hallmark EMT genes, including Postn, Bmp2,
383 Itgb3, Orai2, Mmp11 and Erg **(Fig. 7C,D, S7F,G Tables S4,5)**, which contribute to invasion and
384 metastasis (47-52). By comparison, autocrine or extrinsic EGF activity alone, or in combination
385 with Tgfbr1 signaling, induced the expression of genes that promote cell motility, such as Axl,
386 Sema7a, Ncam1 and Itga2 **(Fig. 7C,E,F, S7F,H, Tables S4,5)** and (53-56). Thus, a compromise
387 between the Tgfbr1 and Egfr signaling programs was necessary to induce a set of pro-motility
388 genes. To determine how the Egfr dependent motility genes influenced the induction of the trail-
389 blazer state, we tested the function of the receptor tyrosine kinase Axl since it can be directly
390 regulated by Fra1 (57). The Axl inhibitor BGB324 suppressed C3-TAg organoid invasion, demon-
391 strating that a gene positively regulated by combined Tgfbr1 and Egfr activity was necessary for
392 trailblazer invasion **(Fig. 7G)**. Together, these results suggested that the induction of the trailblazer
393 state involved a compromise between the Tgfbr1 and Egfr signaling programs that limited features
394 of ECM reorganization and metastatic potential to promote cell motility that was essential for
395 trailblazer cell invasion **(7H)**.

Figure 7

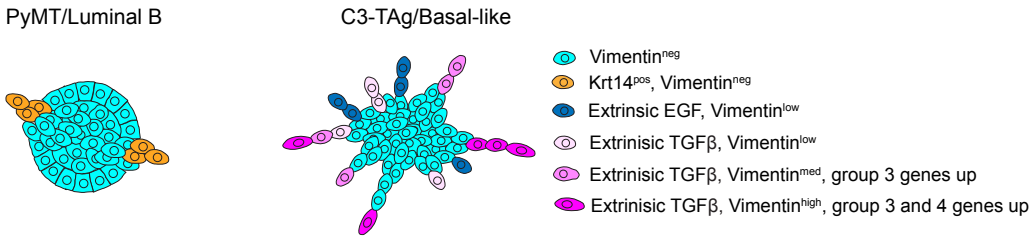

A
C3-TAg-org line
ECM
Passage
Organoids
Embed in ECM
TGFβ1
Treatments
Control
Erlotinib
RNA extraction
RNA-Seq
B
TGFβ1
Control
TGFβ1 + Erlotinib
C
TGFβ1+Erlotinib summary
91 genes
Cell-ECM adhesion
Motility
Chemotaxis
31 genes
Columnar/cubioidal epithelial cell differentiation
14 genes
N/A
224 genes
Branching morphogenesis
Mammary gland development
Tgfbr1
Egfr
1 gene
N/A
122 genes
ECM remodelling
Connective tissue development
171 genes
Xenobiotic responses
Fatty acid metabolism
25 genes
Calcium response
D
Cluster 1
122 genes
Scaled expression
Control
TGFβ1
TGFβ1+ Erlotinib
Treatments
Tgfbr1
Egfr
Targets
Gene Ontology biological processes (BP)
extracellular matrix organization
extracellular structure organization
connective tissue development
odontogenesis of dentin-containing tooth
sensory organ development
negative regulation of cell development
regulation of gliogenesis
negative regulation of nervous system development
odontogenesis
alcohol metabolic process
-log10(p-value)
E
Cluster 5
82 genes
Scaled expression
Control
TGFβ1
TGFβ1+ Erlotinib
Treatments
Tgfbr1
Egfr
Targets
Gene Ontology biological processes (BP)
integrin-mediated signaling pathway
cell-substrate adhesion
ameboidal-type cell migration
chemotaxis
taxis
positive regulation of cell migration
positive regulation of cell motility
blood vessel morphogenesis
positive regulation of locomotion
positive regulation of cellular component movement
-log10(p-value)
F
Tgfbr1
Egfr
Axl
Normalized Axl expression
TGFβ1
Erlotinib
EGF
A83-01
G
C3-TAg Tumor
Digest
Organoids
+/- Drug
Day 0
Day 2
ECM
Control
BGB324
Hoechst, Phalloidin
n=178
n=165
Mask of invading nuclei
Relative area of invading nuclei
p<0.0001
BGB324 48h
H
Tgfbr1
ECM
Motility
Trailblazer state
Egfr

396

Figure 7. A compromise between the Tgbβ and Egfr signaling programs confers the trailblazer state.
A. Overview of the approach to test the interactions between the Egfr and TGFβ1 signaling pathways. **B.** Polar plot organizing genes represented as individual dots based on their change in expression in response to treatment with TGFβ1 and co-treatment with TGFβ1 and Erlotinib. Genes that underwent ≥2 fold-change with a $p < 0.05$ after Benjamini-Hochberg correction in at least one condition relative to one of the other conditions are shown (3 total comparisons). The distance from the center of the plot represents the $\Delta\log_2$ value of each gene (increased difference in expression, further from the center). The bar plots on the surface of the circle show the number of genes at the degree point in the plot. Genes are grouped into 12 color-coded groups along 30 degree increments over the 360-degree plot. As an example, genes in group 5 are increased in expression by TGFβ1 and this increase in expression is inhibited by Erlotinib. In contrast, genes in group 1 are increased in expression by TGFβ1, and this induction is further enhanced by treatment with Erlotinib. **C.** Summary of the number of genes and represented biological processes associated with different modes of interaction between the Tgfbr1 and Egfr pathways in cells treated with TGFβ. Also see **Fig. S7C,D** and **Table S4** for the assignment of polar plot clusters and detailed biological processes used to define the summary integration groups. **D.** Egfr suppressed the expression of a subset of genes induced by extrinsic TGFβ1. Line plots indicate the expression of genes in Cluster 1. Dashed line indicates the average scaled expression of all Cluster 1 genes. Lower bar graph shows the top 10 biological processes (GO-BP) associated with Cluster 1 genes (Fisher's exact test). **E.** Egfr was necessary for the expression of a subset of genes induced by extrinsic TGFβ**.** Line plots show the expression of genes in Cluster 5. Dashed line indicates the average scaled expression of all Cluster 5 genes. Lower bar graph shows the top 10 biological processes (GO-BP) associated with Cluster 5 genes (Fisher's exact test). **F.** Graph of Axl expression (n=2). **G.** The Axl inhibitor BGB324 (1 μM) suppressed C3-TAg tumor organoid invasion. Graph shows the quantification of invading nuclei (mean±SEM, Mann Whitney test, n=organoids, r=2). **H.** Model showing how the induction of the trailblazer EMT state requires a compromise between the Tgfβ and Egfr regulatory programs. The Egfr dependent induction of genes, such as Axl, that are required for trailblazer cell motility has the cost of restricting the expression ECM remodeling genes known to promote invasion and metastasis.

397

Figure 8



398

Figure 8. Model of inter- and intra-tumor trailblazer cell heterogeneity. The characteristic features and functional requirements of the trailblazer cells induced by Tgfβ and Egfr signaling programs in C3-TAg/basal-like tumors are fundamentally distinct from the trailblazer cells in PyMT/luminal B tumors.

399

400 **DISCUSSION**

401 Our findings indicate that defining how cells metastasize requires an understanding of how
402 trailblazer cells are induced in different breast tumors. **(Fig. 8)**. The Tgfβ regulated trailblazer
403 program that we have defined in the basal like C3-TAg tumor cells induced a more aggressive
404 form of collective invasion than was detected in luminal B PyMT tumors and functioned inde-
405 pendently of increased Krt14 and p63 expression. This indicated that the Tgfβ program was mech-
406 anistically and phenotypically distinct from a previously described Krt14-high trailblazer program
407 (20). In addition, PyMT tumor cells rarely expressed detectable Vimentin, which suggested that
408 the Tgfβ regulated trailblazer program that we identified was not active in PyMT tumors. Thus,
409 our results indicate that there are at least 2 different trailblazer programs and that their activity and
410 requirement for invasion varies between different types of mammary tumors associated with pre-
411 viously defined breast cancer intrinsic subtypes **(Fig. 8)**. Genes induced as part of the C3-TAg
412 trailblazer program are more frequently detected at a high level in Basal-type breast cancer pa-
413 tients. Vimentin expression also correlated with collective invasion in TNBC patients, which are
414 most often classified as basal-like. In addition, genes induced as part of Tgfβ trailblazer program
415 have previously been shown to correlate with reduced odds of survival in triple negative breast
416 cancer patients, but not patients classified as ER-positive or HER2-positive (21). These combined
417 results indicate that the Tgfβ trailblazer program is active in basal-like TNBC and influences pa-
418 tient outcome. Tumor subtype specific mechanisms for inducing a trailblazer state are potentially
419 a general property of cancer progression, as molecular subtypes are a feature of many carcinomas
420 (58-61). Further investigation is required to determine the extent to which the Tgfβ and Krt14
421 trailblazer programs align with existing tumor subtypes or define new categories of inter-tumor

422 heterogeneity. Thus, our results provide new insight into approaches used to stratify patients into
423 outcome groups
424
425 The heterogeneity in C3-TAg tumor organoid invasion indicated that the trailblazer popu-
426 lation itself can be heterogeneous within tumors with the same primary driving genetic factors and
427 molecular subtype **(Fig. 8)**. Previous analysis of other tumor cell models had suggested that trail-
428 blazer cells were in a single uniform state (20, 21, 24). Notably, our analysis of 1339-org cells
429 revealed that invasive variability can be a consequence of differences in cellular responses to a
430 single stimulus, in this case TGFβ signaling. Therefore trailblazer cell phenotypic heterogeneity
431 within a population did not necessarily reflect differences between EMT programs that were initi-
432 ated by different activating signals (62). Gene expression changes in response to TGFβ are heter-
433 ogeneous and temporally regulated in other systems as well (63). However, the influence of this
434 gene expression heterogeneity on cell phenotype has not been clear and has often been modeled
435 as a progressive accumulation of new functions (64). The phased changes in gene expression we
436 defined in 1339-org cells suggested how trailblazer phenotypes evolved over time in response to
437 TGFβ activity, with early onset gene expression changes inducing trailblazer cells that were further
438 modulated by later stage reprogramming events **(Fig. 8)**. Interestingly, long-term TGFβ induced
439 changes in gene expression that were associated with other cell lineages. This suggests that the
440 enhanced invasion of trailblazer cells in response to persistent TGFβ stimulation may be associated
441 with a more general lineage plasticity and not simply a re-balancing of epithelial and mesenchymal
442 gene expression.
443

444 Our results revealed that a new compromise mode of integration between Tgfbr1 and Egfr
445 regulated signaling programs was essential for the manifestation of a trailblazer state. The integra-
446 tion of parallel signaling pathways is frequently required to confer specific cell states and is typi-
447 cally classified from the standpoint of being a positive or negative interaction (46, 65, 66). Indeed,
448 prior studies have suggested that TGFβ and Egfr-Ras regulated pathways positively integrate at
449 the point of EMT-TF expression to initiate an EMT program (45, 46, 67, 68). While the combined
450 activity of the Tgfbr1 and Egfr signaling pathways was necessary to induce a trailblazer state, there
451 was an equivalent degree of positive and antagonistic interaction between the programs in C3-TAg
452 tumor cells, indicating a new type of integration between the two signaling programs. Tumor cells
453 must obtain the ability to resist TGFβ induced growth suppression and apoptosis (69). This is
454 accomplished either through the clonal selection of mutants that inactivate Tgfβ signaling, which
455 does not occur in C3-TAg tumor cells, or an intrinsic reprogramming of cellular responses to TGFβ
456 (70). The type of TGFβ reprogramming that is selected for during tumor progression may deter-
457 mine whether Tgfbr1 and Egfr signaling engage in a compromise mode of integration or if there
458 is a predominantly positive or negative interaction between the pathways. The selective activation
459 and restriction of Tgfβ and Egfr dependent gene expression suggests that the compromise integra-
460 tion is based on differences in the expression or activity of co-factors that modulate chromatin
461 accessibility or DNA binding. Tgfβ can increase Egfr expression and activate Erk1/2 signaling in
462 a cell line dependent manner (71, 72). Similarly, Egfr-Mek1/2-Erk1/2 signaling can activate Tgfβ
463 signaling through the phosphorylation of SMAD2/3 and increased expression of Tgfβ expression
464 depending on cell context (41, 73, 74). However, this manner of integration is unlikely contributing
465 to the compromise integration in the C3-TAg tumor cells given that Tgfbr1 inhibition did not alter
466 ERK1/2 phosphorylation or Fra1 expression, Egfr and Mek1/2 inhibition failed to impact Smad2/3

467 phosphorylation. The direct control of the activity of one pathway by the other would also be
468 expected to yield a more uniform positive or negative integration.
469
470 Our findings indicate that defining the nature of Tgfβ and Egfr pathway integration is nec-
471 essary to understand tumor cell properties that influence metastasis and treatment response. In C3-
472 TAg cells in which Tgfβ and Egfr activity combined to promote a trailblazer state, Egfr also re-
473 stricted the expression of genes that could endow trailblazer cells with additional invasive and
474 metastatic properties (50, 75-78). Similarly, Tgfβ limited the Egf induced expression of chemo-
475 kines that increase invasion and the recruitment of tumor promoting immune cells (79, 80). Thus,
476 the compromise mode of integration limited the expression of genes that could promote tumor
477 progression whereas a strictly positive integration would have no such restrictions. The type of
478 interaction between the Tgfβ and Egfr programs also influences the rationale for the therapeutic
479 targeting of trailblazer cells. Inhibiting the Egfr pathway with the clinically approved drugs erlo-
480 tinib and trametinib suppressed C3-TAg tumor cell invasion (81-83). However, combining either
481 drug with newly developed anti-Tgfβ antibodies (84) may have additional benefit due to the fact
482 that the Egfr and Tgfβ pathways function in parallel and each independently regulate large cohorts
483 of genes. By comparison, cells in which Tgfβ activates Egfr-ERK1/2 signaling or vice versa may
484 derive little or no benefit from combination therapy (71, 72). Thus, understanding how the mode
485 of Tgfβ and Egfr pathway integration influences the risk of metastasis and options for treatment is
486 an important line of future investigation.
487

488 Together, our results indicate that an improved understanding of how trailblazer cells are
489 regulated and function in different breast cancer subtypes has the potential to improve the accuracy
490 of patient diagnosis and reveal new treatment options.
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506

507 **MATERIALS AND METHODS**

508 ***Mice***. The C3-TAg [FVB-Tg(C3-1-TAg)cJeg/Jeg] were a gift from Anna T. Riegel (Georgetown
509 University) and purchased from The Jackson Laboratory (013591). REAR [FVB/N-Tg(C3-1-
510 TAg)cJeg/2JegJ] mice and PyMT [FVB/N-Tg(MMTV-PyVT)634Mul/J] mice were purchased
511 from The Jackson Laboratory (REAR, 030386 and PyMT, 002374) and (**Table S1**). For orthotopic
512 tumor experiments, 1339-org organoids were dissociated using TryPLE (Gibco, 12605-010) and
513 500,000 cells were bilaterally injected in the 4th fat pads of 8-12 weeks old female REAR mice.
514 Mice were euthanized and tissue collection was performed on day 27 post injection. Mice were
515 housed and bred in accordance with a protocol approved by the Institutional Animal Use and Care
516 Committee at Georgetown University (IRB# 2017-0076) and in compliance with the NIH Guide
517 for the Care and Use of Laboratory animals. Female mice were used for all experiments.

518

519 ***Tumor organoid derivation.*** The largest tumors from female C3-TAg and PyMT mice were
520 minced and dissociated for up to 120 min at 37°C in a mixture of 1 mg/ml Collagenase (Worthing-
521 ton, LS004188), 20 U/ml DNase (Invitrogen, 18047-019), 5% Fetal Bovine Serum (FBS) (Peak
522 Serum, PS-FB2), ITS (Lonza, 17-838Z), non-essential amino acids (Sigma, M7145) and 10 ng/ml
523 FGF-2 (Peprotech, 100-18C) in DMEM/F-12 (Corning, 10-092-CV). Dissociated tumors were pel-
524 leted at 80 x *g* for 2 min and the supernatant was discarded. Tumor organoids were then rinsed up
525 to 5 times in 5% FBS in DMEM/F-12 followed by filtering through a 250 μm tissue strainer.
526 Organoids were then tested in invasion assays or to establish organoid lines and cell lines (**Table**
527 **S1**).

528

529 ***Cell and organoid culture.*** Cell and organoid lines were generated from C3-TAg tumors **(Table**
530 **S1)** and tested for mycoplasma (Lonza, LT07-703) prior to use and the creation of frozen stocks.
531 Cells and organoids were routinely used within 25 passages. Culture conditions for all cell and
532 organoid lines are summarized in **Table S1.** Organoids were passaged at least once per week by
533 dissociating tumor organoids into single cell suspensions with Dispase (Sigma, SCM133) and Try-
534 PLE (Gibco, 12605-010). The cell suspensions were plated at a density of 200,000-500,000 cells
535 per well in a 24-well ultra-low adhesion plate (Sigma, CLS3474) to form multicellular aggregates
536 for at least 16 h. The aggregates were then embedded in Matrigel (Corning, 354230) or Cultrex
537 (Biotechne, 3533-005-02) and overlaid with Organoid Media (**Table S1**) in 6-well ultra-low ad-
538 hesion plates (Sigma, CLS3471). H2B:GFP retrovirus (85) and H2B:mCherry and GFP lentivirus
539 were produced and cells were infected as described previously (86).
540
541 ***Organoid invasion.*** Organoids in 30 μl of ECM were plated onto a 20 μl base layer of ECM in 8
542 well chamber slides (Falcon, 354108) and allowed to invade for 48 h unless otherwise indicated.
543 The ECM was a mixture of 2.4 mg/ml rat tail collagen I (Corning, CB-40236) and 2 mg/ml of
544 reconstituted basement membrane (Matrigel or Cultrex) unless otherwise indicated. Tumor organ-
545 oids were analyzed in a base media of DMEM/F-12 and ITS in Fig. 1A,F, 5G and 6G, supple-
546 mented with 10 ng/ml FGF2 in 1E and 1% FBS + 10 ng/ml FGF2 in 1B, 1D. A base media of
547 DMEM/F-12 and ITS was used for the analysis of 1339-org, 1863-org and 1788-org lines. Growth
548 factors (2 ng/ml TGFβ1; 50 ng/ml EGF) and inhibitors were added to the base media and tested as
549 indicated in the text and **Tables S1**. Organoids from established organoid lines **(Table S1)** were
550 tested after dissociation and aggregation for at least 16 h, as described for organoid passaging.
551 Organoids were fixed and stained with Hoechst and Phalloidin. Images were acquired with a Zeiss

552 LSM800 laser scanning confocal microscope using 10x/0.45 (Zeiss, 1 420640-9900-000) or
553 20x/0.8 (Zeiss, 1 420650-9902-000) objectives. To determine the area of invasion and circularity,
554 an image mask for each organoid was generated using ImageJ (NIH) based on the Hoechst (inva-
555 sion area) or Phalloidin (circularity) signal. The total area of the invading nuclei was determined
556 using the "Measure" function in ImageJ. Circularity was quantified using the "Analyze Particle"
557 function.
558
559 ***siRNA experiments.*** Cells were transfected with 50 nM of siRNA using RNAiMax transfection
560 reagent (Invitrogen, 13778-150) for 48 h before testing. Cells in all conditions designated as "Con-
561 trol" were transfected with a pool of siRNAs that does not target mouse genes. The catalog num-
562 bers for each siRNA are located in **Table S1**.
563
564 ***Vertical invasion assay.*** Vertical invasion assays were performed and quantified as described pre-
565 viously using a Collagen/Basement Membrane (BM) mix (2.4 mg/ml Collagen I and 2 mg/ml
566 Matrigel or Cultrex) (25). Cells expressing H2B:GFP (10,000-18,000/well) were transfected in 96-
567 well plates (Thermo, 165305) for 48 h. Media was then replaced with 50 μl Collagen/BM) ECM
568 and overlaid with fresh media (DMEM-F12, ITS, 1% FBS, 10 ng/ml FGF2) for 48 h. Cells were
569 imaged with a 10X/0.30 objective (Zeiss, 1 420340-9901-000) on a Zeiss LSM800 laser scanning
570 confocal microscope. Twenty-one z-slices at 10 μm intervals over a total span of 200 μm were
571 acquired for 9 tiled x, y positions per well. At least 3 wells (27 x,y positions total) were imaged
572 per condition in each experiment. ImageJ software (NIH) was used to process images and quantify
573 invasion using a custom macro as described (21). Invasion was normalized by dividing the total
574 number of cells that had migrated ≥ 40 μm above the monolayer by the number of cells in the

575 monolayer. The normalized invasion value for each condition was divided by the normalized in-
576 vasion for the control wells transfected with an siRNA pool that does not target any mouse genes
577 to provide a relative invasion value.
578
579 ***Cluster invasion.*** Cells (10,000-18,000/well) were transfected in 96-well plates (Corning, 3904).
580 After 24 h, the transfected cells were trypsinized and clustered in 96-well Nunclon Sphera U-
581 bottom low adhesion plates (Thermo Scientific, 174925) overnight (1000 cells/well). Homogene-
582 ous and mosaic clusters were also generated in U-bottom Nunclon Sphera low adhesion plates.
583 Cells then aggregated into clusters for at least 16 h before embedding in ECM. Clusters were re-
584 suspended in 30 µl Collagen I/BM mix (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex),
585 plated on 20 µl of a base layer of Matrigel/Collagen I, overlaid with media containing 1% FBS
586 and 2 ng/ml TGFβ1 and allowed to invade for 24 h-48 h unless stated otherwise. The spheroids
587 were fixed and stained with Hoechst and Phalloidin and imaged with a 10x/0.45 objective (Zeiss,
588 1 420640-9900-000) using a Zeiss LSM800 laser scanning confocal microscope. To determine the
589 area of invasion, Hoechst staining was used to identify and mask nuclei invading into the ECM
590 using ImageJ. The total area of the invading nuclei was then determined using the "Measure" func-
591 tion in ImageJ.
592
593 ***Live cell imaging***. Imaging was performed using a Zeiss LSM800 laser scanning confocal micro-
594 scope enclosed in a 37° C chamber supplemented with humidified CO_2. Images were acquired
595 every 30 min for 14-18 h with a 10X/0.30 objective (Zeiss, 1 420340-9901-000). At least 15 dif-
596 ferent x,y coordinates with 5 to 7 z-slices over 50-100 µm span for each condition were acquired
597 in parallel.

598

599 ***Immunofluorescence and immunoblotting.*** Experiments and analysis were performed as de-
600 scribed (25) using antibodies detailed in **Table S1.** Immunoblots were imaged using an Odyssey
601 scanner (Licor, 9120). Immunofluorescence images were acquired using 10x/0.45 (Zeiss, 1
602 420640-9900-000) and 20x/0.8 (Zeiss, 1 420650-9902-000) objectives. For analysis of the rela-
603 tionship between Krt14 expression and invasion, organoids were classified as invasive if they had
604 at least two protrusions with each protrusion having at least two cells invading outward from the
605 center mass. Images were exported as TIFFs and analyzed with ImageJ. The freehand selection
606 tool was used to define ROIs containing individual organoids for analysis. The Krt14 fluorescence
607 for each organoid was then defined as the "Mean Gray Value", as determined using the "Measure"
608 function. To define Krt14 expression in cells leading invasion, >10 organoids stained with Phal-
609 loidin, Hoechst, Krt14 and Krt8 were analyzed. Seven Z slices at 10 μm intervals over a total span
610 of 100 μm were acquired, and the leading cells of each protrusion were designated as Krt14^{low} or
611 Krt14^{high}. Vimentin and E-cadherin expression were analyzed in 5 μm histology sections. Organ-
612 oids in 8-well chamber slides (Falcon, 354108) were washed 3x with distilled water. After detach-
613 ing the chamber walls from the slides, the organoid/ECM gels were gently picked up using a single
614 edge blade (Personna, 94-120-71) and laid out on a strip of parafilm (Bemis, PM-999). The 8-well
615 chamber walls were placed around the gels and 150-200 μl of hydroxyethyl agarose processing
616 gel (Thermo Fisher, HG 4000-012) was poured into the well. After the hydroxyethyl agarose pro-
617 cessing gel solidified, it was transferred into histosettes (Simport, M498-3) and stored in 70%
618 EtOH until embedding in paraffin by the Georgetown Histology and Tissue Shared Resource. The
619 5 μm sections were deparaffinized and immunostained as described (21) to detect Vimentin and
620 E-cadherin expression (**Table S1**). For individual organoids, Vimentin expression was defined as

621 the “Mean Gray Value” using the “Measure” function in ImageJ. Using ImageJ, Vimentin expres-
622 sion was defined as the “Mean Gray Value” for individual organoids using the “Measure” function.
623 For characterization of the trailblazer cell state, the cell leading a collectively invading group of
624 cells was classified as a “trailblazer cell” while cells detached from the tumor organoid were clas-
625 sified as “single cells”. Each cell was assigned E-caderhin^{pos}/ Vimentin^{neg}, E-Cadherin^{pos}/Vimentin-
626 ^{pos} or E-Caderin^{neg}/ Vimentin^{pos} status based on E-Cadherin and Vimentin expression.
627
628 ***Immunohistochemistry.*** Sample processing and staining was performed as described (25). Images
629 were acquired using 10x/0.45 (Zeiss, 1 420640-9900-000) and 20x/0.8 (Zeiss, 1 420650-9902-
630 000) objectives. At least 10 regions of interest (ROIs) were randomly selected for each tumor. Cell
631 Profiler (Broad Institute) was used to define image masks based on Krt8, Krt14 or Vimentin signal.
632 To define their expression in tumor cells, masks based on the area of overlap between the Krt14
633 and Krt8 and the Vimentin and Krt8 signals were determined for each ROI. The area of individual
634 channel and overlap masks created in CellProfiler were then loaded onto ImageJ and quantified
635 using the “Measure” function.
636
637 ***Quantitative real-time PCR.*** Total RNA was isolated using the RNeasy Kit (Qiagen, 74104) or
638 the PureLink™ RNA Mini Kit (Invitrogen, 12183025). Reverse transcription was performed using
639 iScript cDNA Synthesis Kit (Bio-Rad, 1708891) following the manufacturer's protocol. Real time
640 PCR was performed with iQ SYBR Green Supermix (Bio-Rad, 1708882) using a RealPlex 2 ther-
641 mocycler (Eppendorf, 2894). All primers were designed and purchased from Integrated DNA
642 Technologies and their sequences are listed in **Table S1**.
643

644 ***Flow Cytometry.*** Cells were dissociated with TrypLe (Gibco, 12605-010) into single cell suspen-
645 sions, and then washed 2x with FACs buffer (PBS + 2% FBS). Cells were then incubated with
646 anti-CD326 (EpCAM)-PE/Dazzle 594 (Biolegend, 118236) at 4°C for 30 minutes. After staining,
647 cells were washed with FACs buffer 2x, and stained with Helix NP Blue (Biolegend, 425305) in
648 order to exclude dead cells. Cells were then subjected to flow cytometry analysis using BD LSR
649 Fortessa. Results were analyzed using FCSExpress 7 (De Novo Software).
650
651 ***RNA sequencing.*** For the analysis of C3-TAg tumor organoids and non-invasive spheroids in **Fig.**
652 **2 and S2**, at least 1000 C3-TAg tumor organoids or 60,000-80,000 non-invasive cells were em-
653 bedded in a 500 μl ECM mixture (2.4 mg/ml Collagen I and 2 mg/ml Matrigel or Cultrex) and
654 plated on top of 500 uL base layer of ECM in 6 cm dishes (Falcon, 353004). C3-TAg tumor or-
655 ganoids were allowed to invade for 48 h. Non-invasive cells were allowed to grow and form sphe-
656 roids for 6 days. Experiments were performed twice using frozen stocks of tumors and cell lines.
657 For the analysis of 1339-org responses to growth factors and inhibitors in **Fig. 3, 6 and S6**, 250,000
658 1339-org cells were embedded in a 100 μl ECM mixture (2.4 mg/ml Collagen I and 2 mg/ml
659 Matrigel or Cultrex) and plated on top of a 200 μl ECM base layer in 6 cm dishes (Falcon, 353004).
660 The gels were overlaid with control or treatment media. Organoid/ECM gels were collected in 15
661 mL conical tubes (Fisher, 339651), snap frozen and stored in -80^0 C freezer. Experiments were
662 performed twice. For RNA isolation in all experiments, the gels were thawed on ice and disinte-
663 grated in 1 mL Trizol (Invitrogen, 15596026) by repeatedly passing the mixture through a 23G
664 needle (BD, 305143). After the gels were homogenized, the Trizol-gel mixture was allowed to sit
665 at RT for 5 min. Then 200 μL of chloroform (Sigma-Aldrich, C2432-25ML) was added. Following
666 vigorous shaking for 15 sec, the mix was left for 2 min at RT to allow for phase separation. The

667 mix was then centrifuged at 12,000 x g for 15 m at 4^0 C. The upper aqueous phase was then trans-
668 ferred to a new tube and RNA was then further purified using the RNeasy Kit (Qiagen, 74104) and
669 or PureLink™ RNA Mini Kit (Invitrogen, 12183025). For **Fig. 2 and S2**, RNA sequencing, was
670 performed by the Genomics and Epigenomics Shared Resource at Georgetown University. The
671 cDNA libraries were generated using the Illumina TruSeq Stranded Total RNA Library Prepara-
672 tion Kit (Illumina, 20020597). Sequencing was performed using an Illumina NextSeq 500 at a
673 depth of 12 million paired-end 75 base pair reads per sample. Data is available from GEO at
674 GSE188312. Reads were quantified using Salmon with mouse reference genome
675 (GRCm38/mm10). Counts were normalized and differential expression was calculated using
676 DESeq2. For **Fig. 3, 6 and S6**, RNA sequencing was performed by Novogene. RNA quality anal-
677 ysis was assessed by Bioanalyzer using the RNA Nano 6000 Assay Kit (Agilent, 5067-1511). The
678 cDNA libraries were prepared and indexed using the NEBNext Ultra RNA Library Prep Kit for
679 Illumina (NEB, E7530L). Libraries were sequenced with the NovaSeq 6000 at a depth of 20 mil-
680 lion paired-end 150 base pair reads per sample. Sequencing data available from GEO at
681 GSE188300 and GSE188323. Reads were mapped to the mouse reference genome
682 (GRCm38/mm10) using HISAT2 and reads per gene determined using HTSeq. Counts were nor-
683 malized and differential expression was calculated using DESeq2.

684

685 ***Polar plot analysis.*** Expression data was normalized and differentially expressed genes with a ≥2-
686 fold change in expression with $p < 0.05$ after Benjamini-Hochberg (BH) correction were identified
687 using the *DESeq2* package in R (87, 88). Normalized expression data was collapsed for all genes
688 as the average between condition replicates and were subset to only include genes that were dif-
689 ferentially expressed between at least one of the three conditions (control, growth factor, growth

690 factor plus inhibitor) assessed. To define the integration between the pathways, we modified a
691 previously described approach for generating polar plots that has been used to identify dynamic
692 trends in gene expression (44, 89). Relative mean signal changes over successive conditions, of
693 control, growth factor, and growth factor plus inhibitor treatment (T^0, T^1, and T^2, respectively),
694 were transformed to an angle between -180 and +180 degrees (θ). This angle was further trans-
695 formed to a scale from 0 to 360 degrees and then interpolated into a color based on the angle.
696 Differentially expressed genes were plotted using polar coordinates (r,θ) with the radius (r), or
697 distance from the focus, being equal to the value of largest absolute $\log_2$ fold change between
698 conditions and theta (θ) being equal to the angle corresponding to the trend of change and calcu-
699 lated using the following equations:

700

701 $$I = max\ (T^0, T^1, T^2)$$

702 $$R = 1 - \frac{min\ (T^0, T^1, T^2)}{I}$$

703 $$\theta = \left(\frac{T^0 + T^1 - T^2 - I}{I \times R} + 2\right) \times sign(T^1 - T^0) \times 60$$

704 The intensity (I) is equivalent to the maximum expression for a gene across conditions and the
705 ratio (R) between the minimum expression of a gene across conditions over the intensity, are used
706 to calculate the angle of expression trend (θ). As such, genes are plotted as a point with an angle
707 (θ) corresponding to relative changes in a gene's expression with subsequent treatments at a dis-
708 tance from the focus corresponding to the largest magnitude of change between any pair of condi-
709 tions in the treatment series. Three-point profile annotations of expression change (corresponding
710 from left to right as T^0, T^1, and T^2, respectively) are indicated on the outer edge of the polar scatter

711 plots every 30 degrees in the color assigned to that particular θ. These profiles are accompanied
712 by a radial histogram denoting the density of differentially changed genes across organoid condi-
713 tions at particular θ of the polar scatterplot.

714

715 To summarize the meaning of genes with particular profiles in the polar scatterplots, genes
716 were grouped by profiles at 30-degree intervals starting at 0 degrees, which denotes genes that
717 were unchanged in expression from T^0 to T^1 but demonstrated a significant increase in expression
718 following growth factor plus inhibitor treatment of the organoids (T^2). Genes from each group
719 were assessed for enrichment of Gene Ontology biological processes (GP) using the EnrichGo
720 function of the *clusterProfiler* package in R (90). Processes with $p < 0.01$ after BH correction are
721 included in **Tables S4, 5**. Values were transformed ($-\log_{10}$) and displayed alongside polar scatter-
722 plots. Genes within each profile group of the polar scatterplot and utilized for GO enrichment were
723 scaled and displayed as a line plot. A dotted line was added to these linear plots to denote the mean
724 scaled expression pattern across organoid treatment conditions for all genes within a particular
725 interval.

726

727 ***Breast Cancer Tissue Microarray.*** The construction of the TNBC tissue microarray (TMA) con-
728 taining duplicate cores from 50 patients was performed by the Histopathology and Tissue Shared
729 Resource (HTSR) at Georgetown University. Details of the methodology, patient consent, de-iden-
730 tification and demographics have been previously described (91). Multiplex staining of deparaf-
731 finized and rehydrated cores was performed through sequential rounds of staining, stripping and
732 re-probing with antibodies recognizing KRT14, Vimentin and E-cadherin and OPAL fluorescent
733 secondary antibodies **(Table S1)**. The TMA cores were imaged with a Vectra3 Multi-Spectral

734 Imaging microscope (Perkin Elmer) and analyzed with inForm software (Perkin Elmer). Tumor
735 and stroma were segmented based on E-cadherin expression (tumor). Cores with folded sections
736 or no clear tumor-stomal interface by manual inspection were excluded from analysis. Cores were
737 classified "high invasion" if invasive strands were detected in >50% of the tumor-stomal interface
738 by manual inspection. The KRT14 and Vimentin signals were not visible when classifying inva-
739 sion status.
740
741 ***Breast Cancer Patient Gene Expression Analysis.*** The mRNA expression data for 996 patients in
742 the Breast Invasive Carcinoma (The Cancer Genome Atlas (TCGA), PanCancer Atlas) was ana-
743 lyzed and downloaded from the cBioPortal (92, 93)**.** Tumor subtype specific gene expression val-
744 ues are Z-scores relative to all samples. Correlative expression values are Log2 transformed.
745
746 ***Statistical Methods***. Data was analyzed with Prism 9.0.2 (Graphpad) and *DESeq2, GSEA* and
747 *ClusterProfiler* packages in R. Data with a normal distribution, as determined by Shapiro-Wilk
748 test, was analyzed by two tailed Student's t-test. Data that did not pass a normality test were ana-
749 lyzed by Mann-Whitney U test or Kruskal Wallis test with Dunn's Multiple comparison test. Dif-
750 ferentially expressed genes with $p < 0.05$ after Benjamini-Hochberg correction were identified us-
751 ing the *DESeq2* package in R. Fisher's exact tests to determine the enrichment of Gene Ontology
752 biological processes (BP) were performed using the EnrichGo function of the *ClusterProfiler*
753 package in R. The specific statistical tests are indicated in the figure legends. Sample numbers are
754 defined and indicated in the figure legends or on the figure panels. All experiments include a min-
755 imum of 2 biological replicates from separate biological preparations.
756

Data Availability. The RNA sequencing data generated in this study are publicly available in Gene Expression Omnibus (GEO) at GSE188300, GSE188312 and GSE188323. All other data are available within the article and its supplementary data files. Details of code used for analysis is available from the corresponding author upon reasonable request.

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

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

There are no potential conflicts of interest.

AUTHORS' CONTRIBUTIONS

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

ACKNOWLEDGEMENTS

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

REFERENCES

1. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochim Biophys Acta. 2010;1805(1):105-17. Epub 2009/11/26. doi: 10.1016/j.bbcan.2009.11.002. PubMed PMID: 19931353; PMCID: PMC2814927.
2. Beca F, Polyak K. Intratumor Heterogeneity in Breast Cancer. Adv Exp Med Biol. 2016;882:169-89. Epub 2016/03/19. doi: 10.1007/978-3-319-22909-6_7. PubMed PMID: 26987535.
3. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell. 2015;27(1):15-26. doi: 10.1016/j.ccell.2014.12.001. PubMed PMID: 25584892.
4. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell. 2020;37(4):471-84. Epub 2020/04/15. doi: 10.1016/j.ccell.2020.03.007. PubMed PMID: 32289271; PMCID: PMC7181408.
5. Ganesh K, Massagué J. Targeting metastatic cancer. Nat Med. 2021;27(1):34-44. Epub 2021/01/15. doi: 10.1038/s41591-020-01195-4. PubMed PMID: 33442008; PMCID: PMC7895475.
6. Calbo J, van Montfort E, Proost N, van Drunen E, Beverloo HB, Meuwissen R, Berns A. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. Cancer Cell. 2011;19(2):244-56. Epub 2011/02/15. doi: 10.1016/j.ccr.2010.12.021. PubMed PMID: 21316603.
7. Dongre A, Rashidian M, Reinhardt F, Bagnato A, Keckesova Z, Ploegh HL, Weinberg RA. Epithelial-to-Mesenchymal Transition Contributes to Immunosuppression in Breast Carcinomas. Cancer Res. 2017;77(15):3982-9. Epub 2017/04/22. doi: 10.1158/0008-5472.Can-16-3292. PubMed PMID: 28428275; PMCID: PMC5541771.
8. Celià-Terrassa T, Meca-Cortés O, Mateo F, Martínez de Paz A, Rubio N, Arnal-Estapé A, Ell BJ, Bermudo R, Díaz A, Guerra-Rebollo M, Lozano JJ, Estarás C, Ulloa C, Álvarez-Simón D, Milà J, Vilella R, Paciucci R, Martínez-Balbás M, de Herreros AG, Gomis RR, Kang Y, Blanco J, Fernández PL, Thomson TM. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. J Clin Invest. 2012;122(5):1849-68. Epub 2012/04/17. doi: 10.1172/jci59218. PubMed PMID: 22505459; PMCID: PMC3366719.
9. Marusyk A, Tabassum DP, Altrock PM, Almendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature. 2014;514(7520):54-8. doi: 10.1038/nature13556. PubMed PMID: 25079331; PMCID: 4184961.
10. Zhang M, Tsimelzon A, Chang CH, Fan C, Wolff A, Perou CM, Hilsenbeck SG, Rosen JM. Intratumoral heterogeneity in a Trp53-null mouse model of human breast cancer. Cancer Discov. 2015;5(5):520-33. doi: 10.1158/2159-8290.CD-14-1101. PubMed PMID: 25735774; PMCID: 4420701.
11. Neelakantan D, Zhou H, Oliphant MUJ, Zhang X, Simon LM, Henke DM, Shaw CA, Wu MF, Hilsenbeck SG, White LD, Lewis MT, Ford HL. EMT cells increase breast cancer metastasis via paracrine GLI activation in neighbouring tumour cells. Nat Commun. 2017;8:15773. doi: 10.1038/ncomms15773. PubMed PMID: 28604738; PMCID: 5472791.
12. Dongre A, Rashidian M, Eaton EN, Reinhardt F, Thiru P, Zagorulya M, Nepal S, Banaz T, Martner A, Spranger S, Weinberg RA. Direct and Indirect Regulators of Epithelial-Mesenchymal Transition-Mediated Immunosuppression in Breast Carcinomas. Cancer Discov.

827 2021;11(5):1286-305. Epub 2020/12/18. doi: 10.1158/2159-8290.Cd-20-0603. PubMed PMID:
828 33328216; PMCID: PMC8432413.
829 13. Sharif GM, Campbell MJ, Nasir A, Sengupta S, Graham GT, Kushner MH, Kietzman
830 WB, Schmidt MO, Pearson GW, Loudig O, Fineberg S, Wellstein A, Riegel AT. An AIB1 Isof-
831 orm Alters Enhancer Access and Enables Progression of Early-Stage Triple-Negative Breast
832 Cancer. Cancer Res. 2021;81(16):4230-41. Epub 2021/06/18. doi: 10.1158/0008-5472.CAN-20-
833 3625. PubMed PMID: 34135000; PMCID: PMC8373795.
834 14. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. Semin
835 Cell Dev Biol. 2017;64:79-89. doi: 10.1016/j.semcdb.2016.08.028. PubMed PMID: 27582427;
836 PMCID: 5330947.
837 15. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and can-
838 cer. Nat Rev Mol Cell Biol. 2009;10(7):445-57. Epub 2009/06/24. doi: 10.1038/nrm2720. Pub-
839 Med PMID: 19546857.
840 16. Nguyen-Ngoc KV, Cheung KJ, Brenot A, Shamir ER, Gray RS, Hines WC, Yaswen P,
841 Werb Z, Ewald AJ. ECM microenvironment regulates collective migration and local dissemina-
842 tion in normal and malignant mammary epithelium. Proc Natl Acad Sci U S A.
843 2012;109(39):E2595-604. doi: 10.1073/pnas.1212834109. PubMed PMID: 22923691; PMCID:
844 3465416.
845 17. Bronsert P, Enderle-Ammour K, Bader M, Timme S, Kuehs M, Csanadi A, Kayser G,
846 Kohler I, Bausch D, Hoeppner J, Hopt UT, Keck T, Stickeler E, Passlick B, Schilling O, Reiss
847 CP, Vashist Y, Brabletz T, Berger J, Lotz J, Olesch J, Werner M, Wellner UF. Cancer cell inva-
848 sion and EMT marker expression: a three-dimensional study of the human cancer-host interface.
849 J Pathol. 2014;234(3):410-22. doi: 10.1002/path.4416. PubMed PMID: 25081610.
850 18. Cheung KJ, Padmanaban V, Silvestri V, Schipper K, Cohen JD, Fairchild AN, Gorin
851 MA, Verdone JE, Pienta KJ, Bader JS, Ewald AJ. Polyclonal breast cancer metastases arise from
852 collective dissemination of keratin 14-expressing tumor cell clusters. Proc Natl Acad Sci U S A.
853 2016;113(7):E854-63. doi: 10.1073/pnas.1508541113. PubMed PMID: 26831077; PMCID:
854 4763783.
855 19. Pearson GW. Control of Invasion by Epithelial-to-Mesenchymal Transition Programs
856 during Metastasis. J Clin Med. 2019;8(5). doi: 10.3390/jcm8050646. PubMed PMID: 31083398;
857 PMCID: 6572027.
858 20. Cheung KJ, Gabrielson E, Werb Z, Ewald AJ. Collective invasion in breast cancer re-
859 quires a conserved basal epithelial program. Cell. 2013;155(7):1639-51. doi:
860 10.1016/j.cell.2013.11.029. PubMed PMID: 24332913; PMCID: 3941206.
861 21. Westcott JM, Prechtl AM, Maine EA, Dang TT, Esparza MA, Sun H, Zhou Y, Xie Y,
862 Pearson GW. An epigenetically distinct breast cancer cell subpopulation promotes collective in-
863 vasion. J Clin Invest. 2015;125(5):1927-43. doi: 10.1172/JCI77767. PubMed PMID: 25844900.
864 22. Kim YH, Choi YW, Lee J, Soh EY, Kim JH, Park TJ. Senescent tumor cells lead the col-
865 lective invasion in thyroid cancer. Nat Commun. 2017;8:15208. doi: 10.1038/ncomms15208.
866 PubMed PMID: 28489070; PMCID: 5436223.
867 23. Summerbell ER, Mouw JK, Bell JSK, Knippler CM, Pedro B, Arnst JL, Khatib TO,
868 Commander R, Barwick BG, Konen J, Dwivedi B, Seby S, Kowalski J, Vertino PM, Marcus AI.
869 Epigenetically heterogeneous tumor cells direct collective invasion through filopodia-driven fi-
870 bronectin micropatterning. Sci Adv. 2020;6(30):eaaz6197. Epub 2020/08/25. doi: 10.1126/sci-
871 adv.aaz6197. PubMed PMID: 32832657; PMCID: PMC7439406.

24. Konen J, Summerbell E, Dwivedi B, Galior K, Hou Y, Rusnak L, Chen A, Saltz J, Zhou W, Boise LH, Vertino P, Cooper L, Salaita K, Kowalski J, Marcus AI. Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion. Nat Commun. 2017;8:15078. doi: 10.1038/ncomms15078. PubMed PMID: 28497793; PMCID: 5437311.

25. Westcott JM, Camacho S, Nasir A, Huysman ME, Rahhal R, Dang TT, Riegel AT, Brekken RA, Pearson GW. ΔNp63-Regulated Epithelial-to-Mesenchymal Transition State Heterogeneity Confers a Leader–Follower Relationship That Drives Collective Invasion. Cancer Research. 2020;80(18):3933-44. doi: 10.1158/0008-5472.Can-20-0014.

26. Tabassum DP, Polyak K. Tumorigenesis: it takes a village. Nat Rev Cancer. 2015;15(8):473-83. doi: 10.1038/nrc3971. PubMed PMID: 26156638.

27. Celia-Terrassa T, Kang Y. Distinctive properties of metastasis-initiating cells. Genes Dev. 2016;30(8):892-908. doi: 10.1101/gad.277681.116. PubMed PMID: 27083997; PMCID: 4840296.

28. Maine EA, Westcott JM, Prechtl AM, Dang TT, Whitehurst AW, Pearson GW. The cancer-testis antigens SPANX-A/C/D and CTAG2 promote breast cancer invasion. Oncotarget. 2016;7(12):14708-26. doi: 10.18632/oncotarget.7408. PubMed PMID: 26895102.

29. Padmanaban V, Tsehay Y, Cheung KJ, Ewald AJ, Bader JS. Between-tumor and within-tumor heterogeneity in invasive potential. PLoS Comput Biol. 2020;16(1):e1007464. Epub 2020/01/22. doi: 10.1371/journal.pcbi.1007464. PubMed PMID: 31961880; PMCID: PMC6994152.

30. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. Nat Commun. 2020;11(1):5120. Epub 2020/10/11. doi: 10.1038/s41467-020-18794-x. PubMed PMID: 33037194; PMCID: PMC7547708.

31. Dang TT, Esparza MA, Maine EA, Westcott JM, Pearson GW. DeltaNp63alpha Promotes Breast Cancer Cell Motility through the Selective Activation of Components of the Epithelial-to-Mesenchymal Transition Program. Cancer Res. 2015;75(18):3925-35. doi: 10.1158/0008-5472.CAN-14-3363. PubMed PMID: 26292362; PMCID: 4573836.

32. Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. British Journal of Cancer. 2005;93(9):1046-52. doi: 10.1038/sj.bjc.6602787.

33. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World J Clin Oncol. 2014;5(3):412-24. Epub 2014/08/13. doi: 10.5306/wjco.v5.i3.412. PubMed PMID: 25114856; PMCID: PMC4127612.

34. Green JE, Shibata MA, Yoshidome K, Liu ML, Jorcyk C, Anver MR, Wigginton J, Wiltrout R, Shibata E, Kaczmarczyk S, Wang W, Liu ZY, Calvo A, Couldrey C. The C3(1)/SV40 T-antigen transgenic mouse model of mammary cancer: ductal epithelial cell targeting with multistage progression to carcinoma. Oncogene. 2000;19(8):1020-7. doi: 10.1038/sj.onc.1203280. PubMed PMID: 10713685.

35. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S, Backlund MG, Yin Y, Khramtsov AI, Bastein R, Quackenbush J, Glazer RI, Brown PH, Green JE, Kopelovich L, Furth PA, Palazzo JP, Olopade OI, Bernard PS, Churchill GA, Van Dyke T, Perou CM. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007;8(5):R76. Epub 2007/05/12. doi: 10.1186/gb-2007-8-5-r76. PubMed PMID: 17493263; PMCID: PMC1929138.

36. Hwang PY, Brenot A, King AC, Longmore GD, George SC. Randomly Distributed K14(+) Breast Tumor Cells Polarize to the Leading Edge and Guide Collective Migration in Response to Chemical and Mechanical Environmental Cues. Cancer Res. 2019;79(8):1899-912. Epub 2019/03/14. doi: 10.1158/0008-5472.Can-18-2828. PubMed PMID: 30862718; PMCID: PMC6467777.
37. Bergholtz H, Lien TG, Swanson DM, Frigessi A, Bathen TF, Borgen E, Børresen-Dale AL, Engebråten O, Garred Ø, Geisler J, Geitvik GA, Hartmann-Johnsen OJ, Hofvind S, Kristensen VN, Langerød A, Lingjærde OC, Mælandsmo GM, Naume B, Russnes H, Sauer T, Schlichting E, Skjerven HK, Daidone MG, Tost J, Wärnberg F, Sørlie T, Oslo Breast Cancer Research C. Contrasting DCIS and invasive breast cancer by subtype suggests basal-like DCIS as distinct lesions. npj Breast Cancer. 2020;6(1):26. doi: 10.1038/s41523-020-0167-x.
38. Christin JR, Wang C, Chung CY, Liu Y, Dravis C, Tang W, Oktay MH, Wahl GM, Guo W. Stem Cell Determinant SOX9 Promotes Lineage Plasticity and Progression in Basal-like Breast Cancer. Cell Rep. 2020;31(10):107742. Epub 2020/06/11. doi: 10.1016/j.celrep.2020.107742. PubMed PMID: 32521267; PMCID: PMC7658810.
39. Altschuler SJ, Wu LF. Cellular heterogeneity: do differences make a difference? Cell. 2010;141(4):559-63. Epub 2010/05/19. doi: 10.1016/j.cell.2010.04.033. PubMed PMID: 20478246; PMCID: PMC2918286.
40. Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. Nature Reviews Cancer. 2003;3(11):859-68. doi: 10.1038/nrc1209.
41. Bakiri L, Macho-Maschler S, Custic I, Niemiec J, Guío-Carrión A, Hasenfuss SC, Eger A, Müller M, Beug H, Wagner EF. Fra-1/AP-1 induces EMT in mammary epithelial cells by modulating Zeb1/2 and TGFβ expression. Cell Death Differ. 2015;22(2):336-50. Epub 2014/10/10. doi: 10.1038/cdd.2014.157. PubMed PMID: 25301070.
42. Liu ML, Von Lintig FC, Liyanage M, Shibata MA, Jorcyk CL, Ried T, Boss GR, Green JE. Amplification of Ki-ras and elevation of MAP kinase activity during mammary tumor progression in C3(1)/SV40 Tag transgenic mice. Oncogene. 1998;17(18):2403-11. Epub 1998/11/12. doi: 10.1038/sj.onc.1202456. PubMed PMID: 9811472.
43. Albeck John G, Mills Gordon B, Brugge Joan S. Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals. Molecular Cell. 2013;49(2):249-61. doi: https://doi.org/10.1016/j.molcel.2012.11.002.
44. Siersbæk R, Madsen JGS, Javierre BM, Nielsen R, Bagge EK, Cairns J, Wingett SW, Traynor S, Spivakov M, Fraser P, Mandrup S. Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation. Molecular Cell. 2017;66(3):420-35.e5. doi: https://doi.org/10.1016/j.molcel.2017.04.010.
45. Oft M, Peli J, Rudaz C, Schwarz H, Beug H, Reichmann E. TGF-beta1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. Genes Dev. 1996;10(19):2462-77. Epub 1996/10/01. doi: 10.1101/gad.10.19.2462. PubMed PMID: 8843198.
46. Derynck R, Muthusamy BP, Saeteurn KY. Signaling pathway cooperation in TGF-β-induced epithelial–mesenchymal transition. Current Opinion in Cell Biology. 2014;31:56-66. doi: https://doi.org/10.1016/j.ceb.2014.09.001.
47. Parvani JG, Galliher-Beckley AJ, Schiemann BJ, Schiemann WP. Targeted inactivation of β1 integrin induces β3 integrin switching, which drives breast cancer metastasis by TGF-β. Mol Biol Cell. 2013;24(21):3449-59. Epub 2013/09/06. doi: 10.1091/mbc.E12-10-0776. PubMed PMID: 24006485; PMCID: PMC3814150.

964 48. Blee AM, Huang H. ERG-Mediated Cell Invasion: A Link between Development and
965 Tumorigenesis. Medical Epigenetics. 2015;3(2-3):19-29. doi: 10.1159/000440978.
966 49. Huang P, Chen A, He W, Li Z, Zhang G, Liu Z, Liu G, Liu X, He S, Xiao G, Huang F,
967 Stenvang J, Brünner N, Hong A, Wang J. BMP-2 induces EMT and breast cancer stemness
968 through Rb and CD44. Cell Death Discovery. 2017;3(1):17039. doi: 10.1038/cddiscov-
969 ery.2017.39.
970 50. González-González L, Alonso J. Periostin: A Matricellular Protein With Multiple Func-
971 tions in Cancer Development and Progression. Front Oncol. 2018;8:225. Epub 2018/06/28. doi:
972 10.3389/fonc.2018.00225. PubMed PMID: 29946533; PMCID: PMC6005831.
973 51. Ma B, Ran R, Liao H-Y, Zhang H-H. The paradoxical role of matrix metalloproteinase-
974 11 in cancer. Biomedicine & Pharmacotherapy. 2021;141:111899. doi:
975 https://doi.org/10.1016/j.biopha.2021.111899.
976 52. Wu S, Chen M, Huang J, Zhang F, Lv Z, Jia Y, Cui Y-Z, Sun L-Z, Wang Y, Tang Y,
977 Verhoeft KR, Li Y, Qin Y, Lin X, Guan X-Y, Lam K-O. ORAI2 Promotes Gastric Cancer Tu-
978 morigenicity and Metastasis through PI3K/Akt Signaling and MAPK-Dependent Focal Adhesion
979 Disassembly. Cancer Research. 2021;81(4):986. doi: 10.1158/0008-5472.CAN-20-0049.
980 53. Holland SJ, Powell MJ, Franci C, Chan EW, Friera AM, Atchison RE, McLaughlin J,
981 Swift SE, Pali ES, Yam G, Wong S, Lasaga J, Shen MR, Yu S, Xu W, Hitoshi Y, Bogenberger J,
982 Nör JE, Payan DG, Lorens JB. Multiple roles for the receptor tyrosine kinase axl in tumor for-
983 mation. Cancer Res. 2005;65(20):9294-303. Epub 2005/10/19. doi: 10.1158/0008-5472.Can-05-
984 0993. PubMed PMID: 16230391.
985 54. Zecchini S, Bombardelli L, Decio A, Bianchi M, Mazzarol G, Sanguineti F, Aletti G,
986 Maddaluno L, Berezin V, Bock E, Casadio C, Viale G, Colombo N, Giavazzi R, Cavallaro U.
987 The adhesion molecule NCAM promotes ovarian cancer progression via FGFR signalling.
988 EMBO Molecular Medicine. 2011;3(8):480-94. doi: https://doi.org/10.1002/emmm.201100152.
989 55. Black SA, Nelson AC, Gurule NJ, Futscher BW, Lyons TR. Semaphorin 7a exerts plei-
990 otropic effects to promote breast tumor progression. Oncogene. 2016;35(39):5170-8. Epub
991 2016/04/12. doi: 10.1038/onc.2016.49. PubMed PMID: 27065336; PMCID: PMC5720143.
992 56. Adorno-Cruz V, Hoffmann AD, Liu X, Dashzeveg NK, Taftaf R, Wray B, Keri RA, Liu
993 H. ITGA2 promotes expression of ACLY and CCND1 in enhancing breast cancer stemness and
994 metastasis. Genes & Diseases. 2021;8(4):493-508. doi: https://doi.org/10.1016/j.gen-
995 dis.2020.01.015.
996 57. Sayan AE, Stanford R, Vickery R, Grigorenko E, Diesch J, Kulbicki K, Edwards R, Pal
997 R, Greaves P, Jariel-Encontre I, Piechaczyk M, Kriajevska M, Mellon JK, Dhillon AS,
998 Tulchinsky E. Fra-1 controls motility of bladder cancer cells via transcriptional upregulation of
999 the receptor tyrosine kinase AXL. Oncogene. 2012;31(12):1493-503. Epub 2011/08/09. doi:
1000 10.1038/onc.2011.336. PubMed PMID: 21822309.
1001 58. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Can-
1002 cer. Cell. 2015;163(4):1011-25. doi: 10.1016/j.cell.2015.10.025. PubMed PMID: 26544944;
1003 PMCID: 4695400.
1004 59. Chen F, Zhang Y, Parra E, Rodriguez J, Behrens C, Akbani R, Lu Y, Kurie JM, Gibbons
1005 DL, Mills GB, Wistuba II, Creighton CJ. Multiplatform-based molecular subtypes of non-small-
1006 cell lung cancer. Oncogene. 2017;36(10):1384-93. doi: 10.1038/onc.2016.303.
1007 60. McConkey DJ, Choi W. Molecular Subtypes of Bladder Cancer. Current Oncology Re-
1008 ports. 2018;20(10):77. doi: 10.1007/s11912-018-0727-5.

1009 61. Collisson EA, Bailey P, Chang DK, Biankin AV. Molecular subtypes of pancreatic can-
1010 cer. Nature Reviews Gastroenterology & Hepatology. 2019;16(4):207-20. doi: 10.1038/s41575-
1011 019-0109-y.
1012 62. Katsuno Y, Derynck R. Epithelial plasticity, epithelial-mesenchymal transition, and the
1013 TGF-β family. Developmental Cell. 2021;56(6):726-46. doi:
1014 https://doi.org/10.1016/j.devcel.2021.02.028.
1015 63. van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, Burdziak C, Moon KR,
1016 Chaffer CL, Pattabiraman D, Bierie B, Mazutis L, Wolf G, Krishnaswamy S, Pe'er D. Recover-
1017 ing Gene Interactions from Single-Cell Data Using Data Diffusion. Cell. 2018;174(3):716-
1018 29.e27. doi: https://doi.org/10.1016/j.cell.2018.05.061.
1019 64. Lambert AW, Weinberg RA. Linking EMT programmes to normal and neoplastic epithe-
1020 lial stem cells. Nat Rev Cancer. 2021;21(5):325-38. Epub 2021/02/07. doi: 10.1038/s41568-021-
1021 00332-6. PubMed PMID: 33547455.
1022 65. Chakrabarti R, Hwang J, Andres Blanco M, Wei Y, Lukačišin M, Romano RA, Smalley
1023 K, Liu S, Yang Q, Ibrahim T, Mercatali L, Amadori D, Haffty BG, Sinha S, Kang Y. Elf5 inhib-
1024 its the epithelial-mesenchymal transition in mammary gland development and breast cancer me-
1025 tastasis by transcriptionally repressing Snail2. Nat Cell Biol. 2012;14(11):1212-22. Epub
1026 2012/10/23. doi: 10.1038/ncb2607. PubMed PMID: 23086238; PMCID: PMC3500637.
1027 66. Lehmann W, Mossmann D, Kleemann J, Mock K, Meisinger C, Brummer T, Herr R,
1028 Brabletz S, Stemmler MP, Brabletz T. ZEB1 turns into a transcriptional activator by interacting
1029 with YAP1 in aggressive cancer types. Nature Communications. 2016;7(1):10498. doi:
1030 10.1038/ncomms10498.
1031 67. Doehn U, Hauge C, Frank SR, Jensen CJ, Duda K, Nielsen JV, Cohen MS, Johansen JV,
1032 Winther BR, Lund LR, Winther O, Taunton J, Hansen SH, Frödin M. RSK Is a Principal Effec-
1033 tor of the RAS-ERK Pathway for Eliciting a Coordinate Promotile/Invasive Gene Program and
1034 Phenotype in Epithelial Cells. Molecular Cell. 2009;35(4):511-22. doi:
1035 https://doi.org/10.1016/j.molcel.2009.08.002.
1036 68. Shin S, Dimitri CA, Yoon S-O, Dowdle W, Blenis J. ERK2 but Not ERK1 Induces Epi-
1037 thelial-to-Mesenchymal Transformation via DEF Motif-Dependent Signaling Events. Molecular
1038 Cell. 2010;38(1):114-27. doi: https://doi.org/10.1016/j.molcel.2010.02.020.
1039 69. Massagué J. TGFβ signalling in context. Nat Rev Mol Cell Biol. 2012;13(10):616-30.
1040 Epub 2012/09/21. doi: 10.1038/nrm3434. PubMed PMID: 22992590; PMCID: PMC4027049.
1041 70. David CJ, Massagué J. Contextual determinants of TGFβ action in development, immun-
1042 ity and cancer. Nat Rev Mol Cell Biol. 2018;19(7):419-35. Epub 2018/04/13. doi:
1043 10.1038/s41580-018-0007-0. PubMed PMID: 29643418; PMCID: PMC7457231.
1044 71. Lee MK, Pardoux C, Hall MC, Lee PS, Warburton D, Qing J, Smith SM, Derynck R.
1045 TGF-β activates Erk MAP kinase signalling through direct phosphorylation of ShcA. The
1046 EMBO Journal. 2007;26(17):3957-67. doi: https://doi.org/10.1038/sj.emboj.7601818.
1047 72. Zhao Y, Ma J, Fan Y, Wang Z, Tian R, Ji W, Zhang F, Niu R. TGF-β transactivates
1048 EGFR and facilitates breast cancer migration and invasion through canonical Smad3 and
1049 ERK/Sp1 signaling pathways. Mol Oncol. 2018;12(3):305-21. Epub 2017/12/08. doi:
1050 10.1002/1878-0261.12162. PubMed PMID: 29215776; PMCID: PMC5830653.
1051 73. Blanchette F, Rivard N, Rudd P, Grondin F, Attisano L, Dubois CM. Cross-talk between
1052 the p42/p44 MAP Kinase and Smad Pathways in Transforming Growth Factor β1-in-
1053 duced Furin Gene Transactivation *. Journal of Biological Chemistry. 2001;276(36):33986-94.
1054 doi: 10.1074/jbc.M100093200.

74. Hough C, Radu M, Doré JJE. TGF-Beta Induced Erk Phosphorylation of Smad Linker Region Regulates Smad Signaling. PLOS ONE. 2012;7(8):e42513. doi: 10.1371/journal.pone.0042513.
75. Clement JH, Raida M, Sänger J, Bicknell R, Liu J, Naumann A, Geyer A, Waldau A, Hortschansky P, Schmidt A, Höffken K, Wölft S, Harris AL. Bone morphogenetic protein 2 (BMP-2) induces in vitro invasion and in vivo hormone independent growth of breast carcinoma cells. Int J Oncol. 2005;27(2):401-7. Epub 2005/07/13. PubMed PMID: 16010421.
76. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25. doi: 10.1016/j.cels.2015.12.004. PubMed PMID: 26771021.
77. Huang P, Chen A, He W, Li Z, Zhang G, Liu Z, Liu G, Liu X, He S, Xiao G, Huang F, Stenvang J, Brünner N, Hong A, Wang J. BMP-2 induces EMT and breast cancer stemness through Rb and CD44. Cell Death Discov. 2017;3:17039. Epub 2017/07/21. doi: 10.1038/cddiscovery.2017.39. PubMed PMID: 28725489; PMCID: PMC5511860.
78. Cheng YX, Xiao L, Yang YL, Liu XD, Zhou XR, Bu ZF, Cao PC, Wang DK. Collagen type VIII alpha 2 chain (COL8A2), an important component of the basement membrane of the corneal endothelium, facilitates the malignant development of glioblastoma cells via inducing EMT. J Bioenerg Biomembr. 2021;53(1):49-59. Epub 2021/01/07. doi: 10.1007/s10863-020-09865-1. PubMed PMID: 33405048.
79. Zhao J, Ou B, Han D, Wang P, Zong Y, Zhu C, Liu D, Zheng M, Sun J, Feng H, Lu A. Tumor-derived CXCL5 promotes human colorectal cancer metastasis through activation of the ERK/Elk-1/Snail and AKT/GSK3β/β-catenin pathways. Mol Cancer. 2017;16(1):70. Epub 2017/03/31. doi: 10.1186/s12943-017-0629-4. PubMed PMID: 28356111; PMCID: PMC5372323.
80. Susek KH, Karvouni M, Alici E, Lundqvist A. The Role of CXC Chemokine Receptors 1-4 on Immune Cells in the Tumor Microenvironment. Front Immunol. 2018;9:2159. Epub 2018/10/16. doi: 10.3389/fimmu.2018.02159. PubMed PMID: 30319622; PMCID: PMC6167945.
81. Roberts PJ, Usary JE, Darr DB, Dillon PM, Pfefferle AD, Whittle MC, Duncan JS, Johnson SM, Combest AJ, Jin J, Zamboni WC, Johnson GL, Perou CM, Sharpless NE. Combined PI3K/mTOR and MEK inhibition provides broad antitumor activity in faithful murine cancer models. Clin Cancer Res. 2012;18(19):5290-303. doi: 10.1158/1078-0432.CCR-12-0563. PubMed PMID: 22872574; PMCID: 3715399.
82. Usary J, Zhao W, Darr D, Roberts PJ, Liu M, Balletta L, Karginova O, Jordan J, Combest A, Bridges A, Prat A, Cheang MC, Herschkowitz JI, Rosen JM, Zamboni W, Sharpless NE, Perou CM. Predicting drug responsiveness in human cancers using genetically engineered mice. Clin Cancer Res. 2013;19(17):4889-99. doi: 10.1158/1078-0432.CCR-13-0522. PubMed PMID: 23780888; PMCID: 3778918.
83. Reguart N, Remon J. Common EGFR-mutated subgroups (Del19/L858R) in advanced non-small-cell lung cancer: chasing better outcomes with tyrosine kinase inhibitors. Future Oncology. 2015;11(8):1245-57. doi: 10.2217/fon.15.15. PubMed PMID: 25629371.
84. Li S, Liu M, Do MH, Chou C, Stamatiades EG, Nixon BG, Shi W, Zhang X, Li P, Gao S, Capistrano KJ, Xu H, Cheung N-KV, Li MO. Cancer immunotherapy via targeted TGF-β signalling blockade in TH cells. Nature. 2020;587(7832):121-5. doi: 10.1038/s41586-020-2850-3.

85. Kanda T, Sullivan KF, Wahl GM. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. Curr Biol. 1998;8(7):377-85. Epub 1998/05/16. doi: 10.1016/s0960-9822(98)70156-3. PubMed PMID: 9545195.

86. Dang TT, Prechtl AM, Pearson GW. Breast cancer subtype-specific interactions with the microenvironment dictate mechanisms of invasion. Cancer Res. 2011;71(21):6857-66. Epub 2011/09/13. doi: 10.1158/0008-5472.Can-11-1818. PubMed PMID: 21908556; PMCID: PMC3206184.

87. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300. doi: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

88. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.

89. Rauch A, Haakonsson AK, Madsen JGS, Larsen M, Forss I, Madsen MR, Van Hauwaert EL, Wiwie C, Jespersen NZ, Tencerova M, Nielsen R, Larsen BD, Röttger R, Baumbach J, Scheele C, Kassem M, Mandrup S. Osteogenesis depends on commissioning of a network of stem cell transcription factors that act as repressors of adipogenesis. Nature Genetics. 2019;51(4):716-27. doi: 10.1038/s41588-019-0359-1.

90. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology. 2012;16(5):284-7. doi: 10.1089/omi.2011.0118.

91. Fernandez AI, Geng X, Chaldekas K, Harris B, Duttargi A, Berry VL, Berry DL, Mahajan A, Cavalli LR, Győrffy B, Tan M, Riggins RB. The orphan nuclear receptor estrogen-related receptor beta (ERRβ) in triple-negative breast cancer. Breast Cancer Research and Treatment. 2020;179(3):585-604. doi: 10.1007/s10549-019-05485-5.

92. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-4. doi: 10.1158/2159-8290.CD-12-0095. PubMed PMID: 22588877; PMCID: 3956037.

93. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, Omberg L, Wolf DM, Shriver CD, Thorsson V, Hu H. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell. 2018;173(2):400-16.e11. Epub 2018/04/07. doi: 10.1016/j.cell.2018.02.052. PubMed PMID: 29625055; PMCID: PMC6066282.

Figure S1

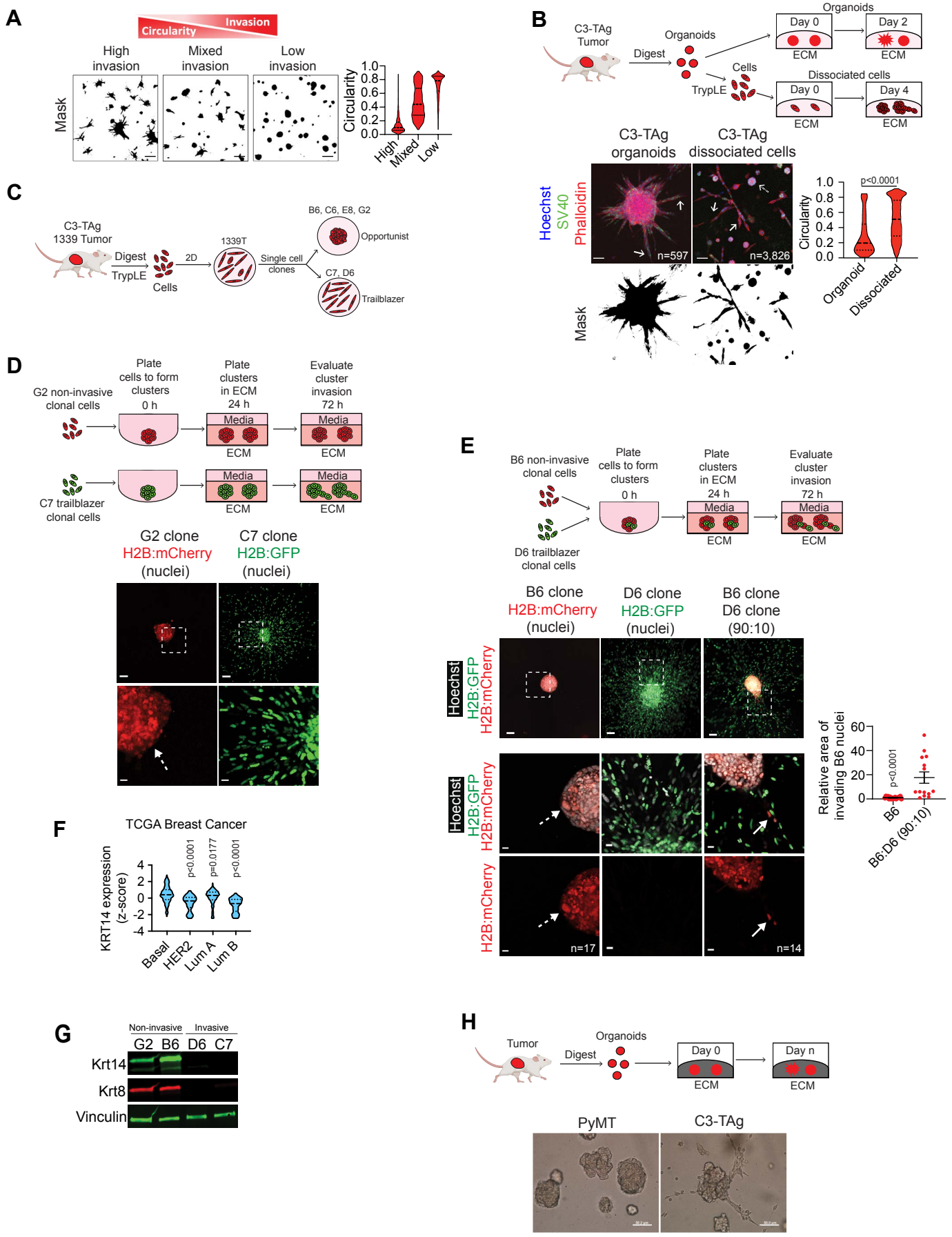

A
Circularity
Invasion
High invasion
Mixed invasion
Low invasion
Mask
Circularity
1.0
0.8
0.6
0.4
0.2
0.0
High
Mixed
Low
B
C3-TAg Tumor
Digest
Organoids
Cells
TrypLE
Organoids
Day 0
Day 2
ECM
ECM
Dissociated cells
Day 0
Day 4
ECM
ECM
C3-TAg organoids
C3-TAg dissociated cells
Hoechst
SV40
Phalloidin
n=597
n=3,826
Mask
Circularity
p<0.0001
Organoid
Dissociated
C
C3-TAg 1339 Tumor
Digest
TrypLE
Cells
2D
1339T
Single cell clones
B6, C6, E8, G2
Opportunist
C7, D6
Trailblazer
D
G2 non-invasive clonal cells
Plate cells to form clusters 0 h
Plate clusters in ECM 24 h
Evaluate cluster invasion 72 h
Media
ECM
C7 trailblazer clonal cells
G2 clone
H2B:mCherry
(nuclei)
C7 clone
H2B:GFP
(nuclei)
E
B6 non-invasive clonal cells
D6 trailblazer clonal cells
Plate cells to form clusters 0 h
Plate clusters in ECM 24 h
Evaluate cluster invasion 72 h
Media
ECM
B6 clone
H2B:mCherry
(nuclei)
D6 clone
H2B:GFP
(nuclei)
B6 clone
D6 clone
(90:10)
Hoechst
H2B:GFP
H2B:mCherry
H2B:mCherry
n=17
n=14
Relative area of invading B6 nuclei
60
40
20
0
p<0.0001
B6
B6:D6 (90:10)
F
TCGA Breast Cancer
KRT14 expression (z-score)
4
2
0
-2
-4
p<0.0001
p=0.0177
p<0.0001
Basal
HER2
Lum A
Lum B
G
Non-invasive
Invasive
G2
B6
D6
C7
Krt14
Krt8
Vinculin
H
Tumor
Digest
Organoids
Day 0
Day n
ECM
ECM
PyMT
C3-TAg

Figure S1, continued

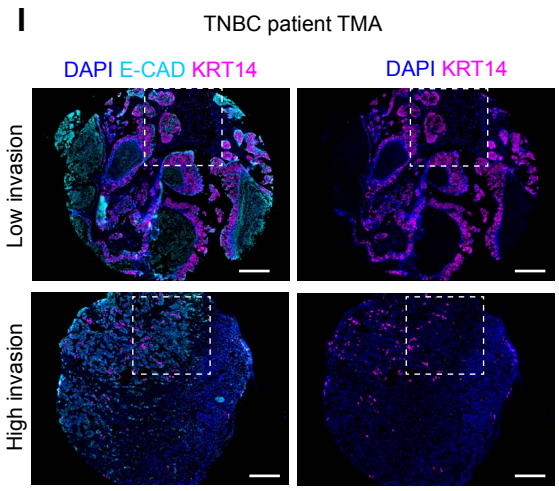


Figure S1. A. Representative masked images of C3-TAg organoids showing how reduced organoid invasion is represented by increased organoid circularity. Scale bar 50 µm. **B.** C3-TAg organoids and dissociated cells from the same tumors were grown in ECM for 2 and 4 days respectively to evaluate invasion. Solid arrows indicate representative invasive cells. Dashed arrow indicates a representative non-invasive spheroid. Violin plot shows the quantification of circularity (Mann Whitney test, n=organoids (C3-TAg organoids) or spheroids (dissociated cells), r=2). Scale bars, 50 µm. **C.** Graphic shows the methodology for deriving clonal trailblazer and opportunist cell lines. **D.** Representative images of non-invasive (G2) and invasive (C7) clonal cell lines. Scale bars, 100 µm (top), 20 µm (bottom, inset). **E.** Invasive D6-H2B:GFP cells promoted the opportunistic invasion of non-invasive B6-H2B:mCherry cells. Top row, 2x2 tiled LED fluorescence images. Inset images are confocal sections of the ROIs indicated by the dashed line boxes. Graph shows the quantification of B6-H2B:mCherry cell invasion (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, top row 100 µm; inset images 20 µm. **F.** KRT14 mRNA is more highly expressed in the basal-like subtype compared to other types of breast cancer in the TCGA. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to the basal-like tumors. **G.** Immunoblots showing the expression of Krt14 and Krt8 in non-invasive (G2, B6) and invasive (D6, C7) C3-TAg clonal cell lines (n=2). **H.** Representative images from 2 independent experiments show that C3-TAg organoids are intrinsically more invasive than PyMT organoids in 100% Cultrex ECM. **I.** Full TNBC patient tumors images with dashed boxes showing the ROIs from **1H**. Scale bars, 200 µm.

Figure S2

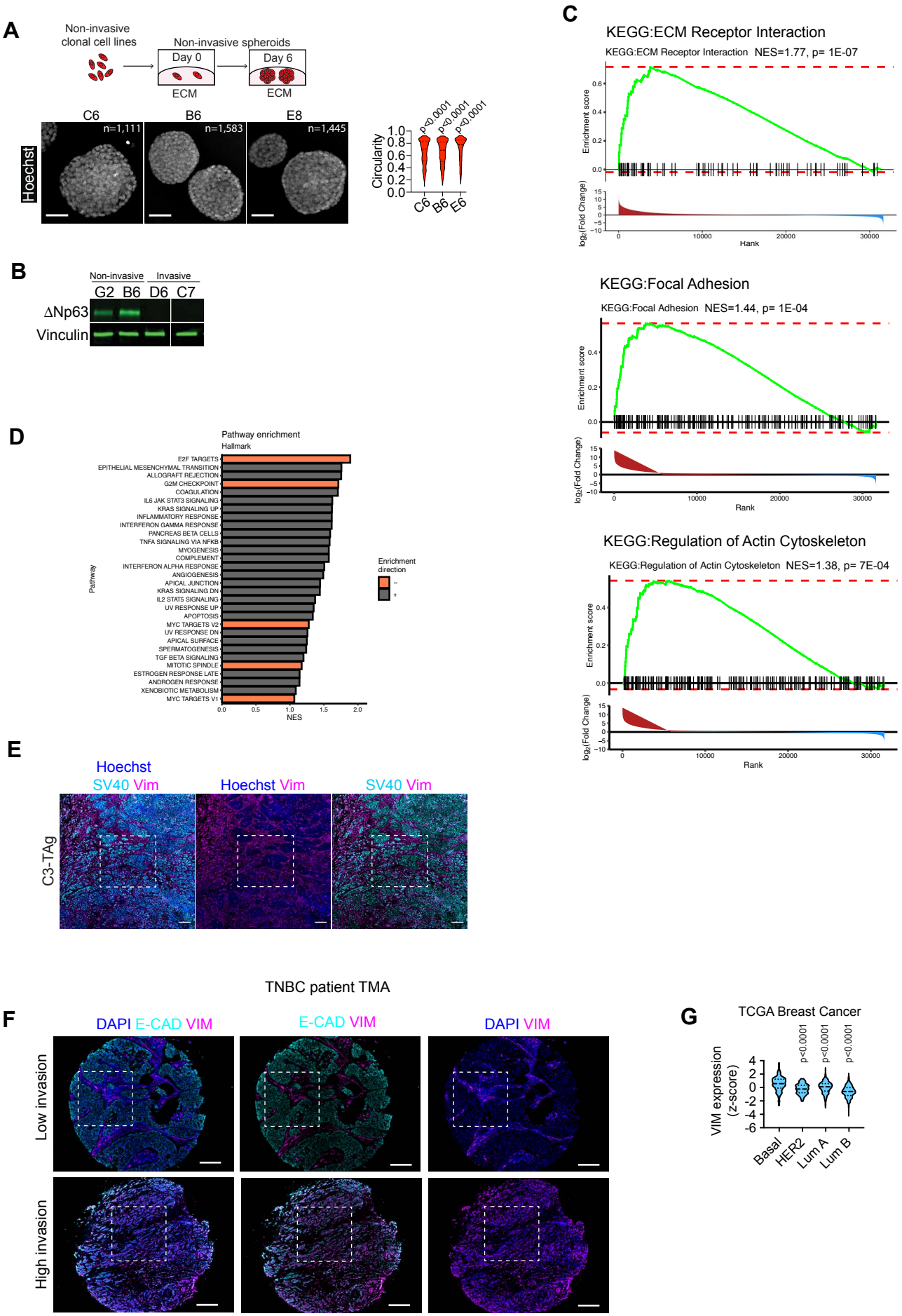

A
Non-invasive clonal cell lines
Non-invasive spheroids
Day 0
Day 6
ECM
ECM
C6
B6
E8
n=1,111
n=1,583
n=1,445
Hoechst
Circularity
p<0.0001
p<0.0001
p<0.0001
B
Non-invasive
Invasive
G2 B6 D6 C7
ΔNp63
Vinculin
C
KEGG:ECM Receptor Interaction
KEGG:ECM Receptor Interaction NES=1.77, p= 1E-07
Enrichment score
log2(Fold Change)
Rank
KEGG:Focal Adhesion
KEGG:Focal Adhesion NES=1.44, p= 1E-04
KEGG:Regulation of Actin Cytoskeleton
KEGG:Regulation of Actin Cytoskeleton NES=1.38, p= 7E-04
D
Pathway enrichment
Hallmark
Pathway
E2F TARGETS
EPITHELIAL MESENCHYMAL TRANSITION
ALLOGRAFT REJECTION
G2M CHECKPOINT
COAGULATION
IL6 JAK STAT3 SIGNALING
KRAS SIGNALING UP
INFLAMMATORY RESPONSE
INTERFERON GAMMA RESPONSE
PANCREAS BETA CELLS
TNFA SIGNALING VIA NFKB
MYOGENESIS
COMPLEMENT
INTERFERON ALPHA RESPONSE
ANGIOGENESIS
APICAL JUNCTION
KRAS SIGNALING DN
IL2 STAT5 SIGNALING
UV RESPONSE UP
APOPTOSIS
MYC TARGETS V2
UV RESPONSE DN
APICAL SURFACE
SPERMATOGENESIS
TGF BETA SIGNALING
MITOTIC SPINDLE
ESTROGEN RESPONSE LATE
ANDROGEN RESPONSE
XENOBIOTIC METABOLISM
MYC TARGETS V1
NES
Enrichment direction
E
Hoechst
SV40 Vim
Hoechst Vim
SV40 Vim
C3-TAg
TNBC patient TMA
F
DAPI E-CAD VIM
E-CAD VIM
DAPI VIM
Low invasion
High invasion
G
TCGA Breast Cancer
VIM expression (z-score)
p<0.0001
p<0.0001
p<0.0001
Basal
HER2
Lum A
Lum B

Figure S2. A. The indicated non-invasive opportunist C3-TAg clonal cell lines were grown in ECM for at least 6 days. Violin plot shows the quantification of spheroid circularity (Mann Whitney test, n=spheroids, r=2). Scale bars, 50 µm. **B.** Immunoblots showing the expression of ΔN-p63 in non-invasive (G2, B6) and invasive (D6, C7) C3-TAg clonal cell lines (n=2). Lysates are from the same gel. **C.** GSEA was performed on RNA-seq data from C3-TAg organoids and non-invasive C3-TAg clones to identify pathway enrichment in C3-TAg organoids. Plots show the GSEA for the indicated KEGG pathways. **D.** Graph shows Hallmark pathways with NES ≥1. **E**. Overview image of a C3-TAg tumor with dashed boxes showing the Tumor-ECM interface ROIs from **2G**. Scale bars 100 µm. **F.** Full TNBC patient tumor images with dashed boxes showing the ROIs from **2H**. Scale bars, 200 µm. **G.** VIM mRNA is more highly expressed in the basal-like subtype compared to other types of breast cancer in the TCGA. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to basal-like tumors.

Figure S3

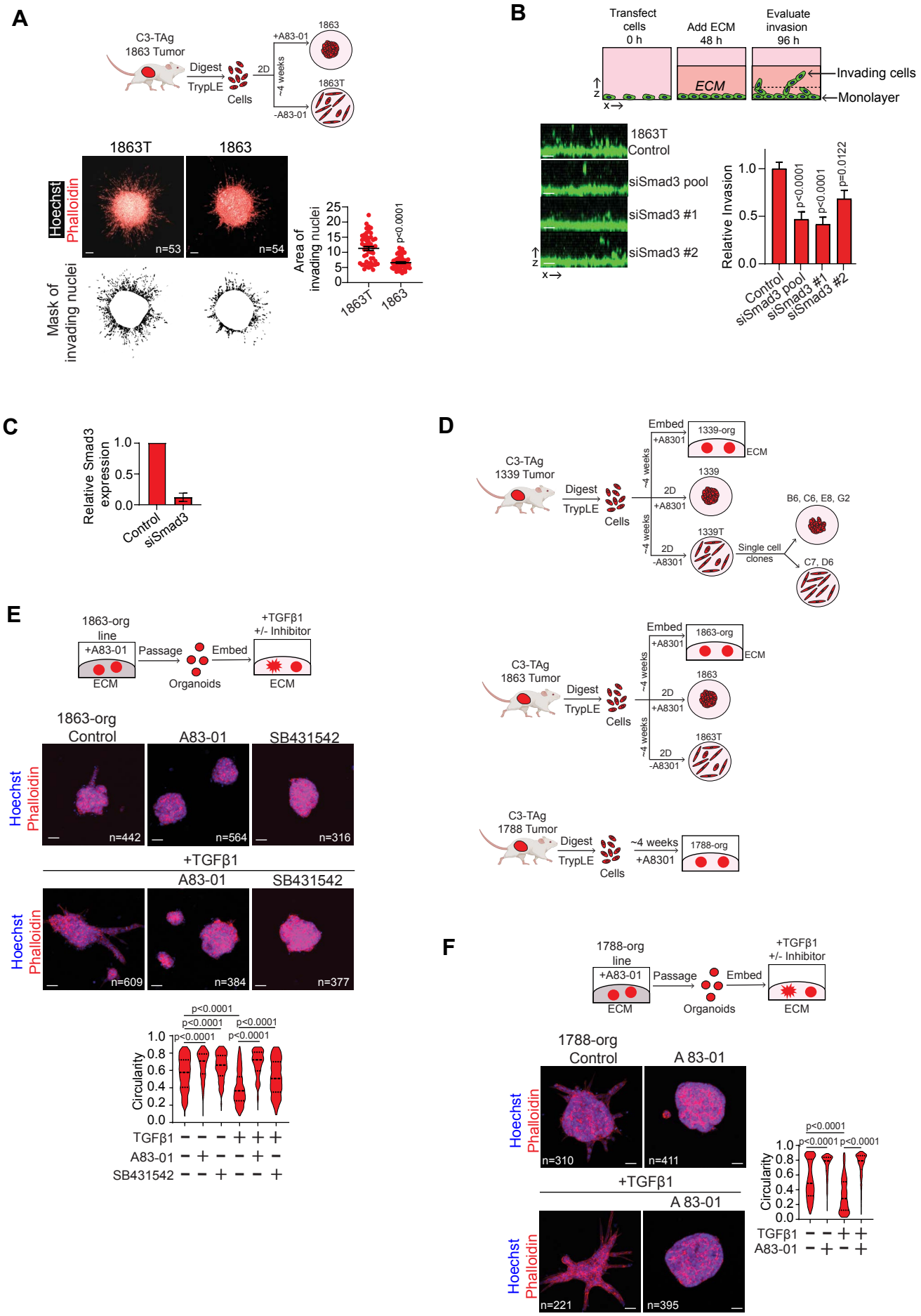

A
C3-TAg
1863 Tumor
Digest
TrypLE
Cells
2D
~4 weeks
+A83-01
-A83-01
1863
1863T
1863T
1863
Hoechst
Phalloidin
n=53
n=54
Mask of invading nuclei
Area of invading nuclei
p<0.0001
1863T
1863
B
Transfect cells 0 h
Add ECM 48 h
Evaluate invasion 96 h
ECM
Invading cells
Monolayer
1863T Control
siSmad3 pool
siSmad3 #1
siSmad3 #2
Relative Invasion
p<0.0001
p<0.0001
p=0.0122
Control
siSmad3 pool
siSmad3 #1
siSmad3 #2
C
Relative Smad3 expression
Control
siSmad3
D
C3-TAg
1339 Tumor
Digest
TrypLE
Cells
Embed
+A8301
1339-org
ECM
1339
2D
+A8301
1339T
2D
-A8301
Single cell clones
B6, C6, E8, G2
C7, D6
C3-TAg
1863 Tumor
1863-org
1863
1863T
C3-TAg
1788 Tumor
1788-org
E
1863-org line
+A83-01
ECM
Passage
Organoids
Embed
+TGFβ1
+/- Inhibitor
1863-org
Control
A83-01
SB431542
Hoechst
Phalloidin
n=442
n=564
n=316
+TGFβ1
n=609
n=384
n=377
p<0.0001
Circularity
TGFβ1
A83-01
SB431542
F
1788-org line
1788-org
Control
A 83-01
n=310
n=411
n=221
n=395
Circularity
TGFβ1
A83-01

Figure S3. A. 1863 cells were derived in the presence of A83-01 and were less invasive than 1863 cells, which were derived from the same tumor but in the presence of A83-01. Graph shows cluster invasion (mean±SEM, Mann Whitney test, n= clusters, r=2). Scale bars, 50µm. **B.** 1863T vertical invasion was suppressed by a Smad3 siRNA pool and individual Smad3 siRNAs (mean±SEM, unpaired Student's t test, n=3). Scale bars, 50 µm. **C.** Graph shows siRNA induced Smad3 depletion (mean±SEM, n=2). **D.** Graphics showing the process for deriving 1339-org, 1863-org and 1788-org lines and their relationship to other cell lines derived from the same tumors. **E, F.** Tgfbr1 inhibition suppresses intrinsic and extrinsic TGFβ1 induced invasion of multiple C3-TAg organoid lines (**E**, 1863-org; **F**, 1788-org). Organoids were treated with the Tgfbr1 inhibitors A83-01 (500 nM) and SB431542 (1 µM) and extrinsic TGFβ1 (2 ng/ml) as indicated. Violin plots show the quantification of organoid circularity (Mann Whitney test, n=organoids). Scale bars, 50 µm.

Figure S4

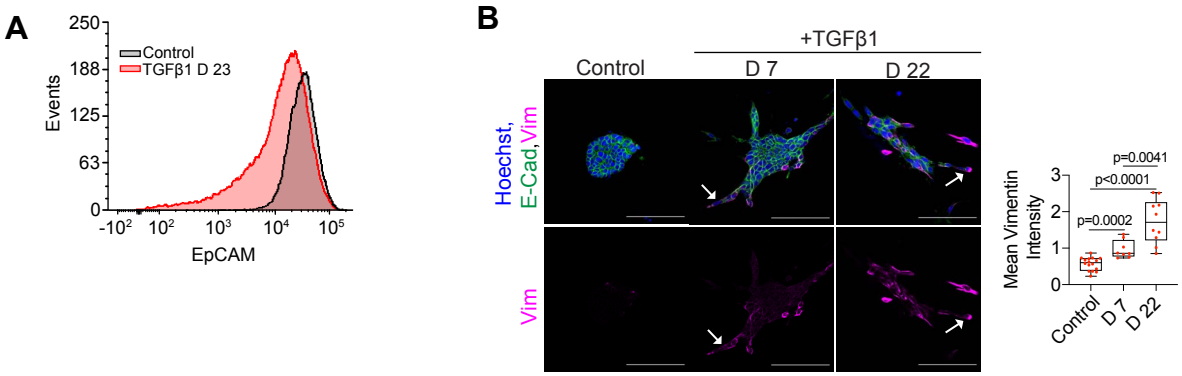

A
250
188
125
63
0
Events
Control
TGFβ1 D 23
-10²
10²
10³
10⁴
10⁵
EpCAM
B
+TGFβ1
Control
D 7
D 22
Hoechst,
E-Cad,Vim
Vim
Mean Vimentin
Intensity
3
2
1
0
p=0.0041
p<0.0001
p=0.0002
Control
D 7
D 22

Figure S4. A. FACS showing how TGFβ1 induces a shift in the distribution of EpCAM expressing 1339-org cells (representative of 2 independent experiments). **B.** TGFβ1 progressively increases Vimentin expression in 1339-org cells. Arrows indicate representative Vimentin expressing cells leading collective invasion. Graph shows Vimentin expression in 1339-org organoids (Student's t-test, n=images analyzed). Scale bars, 100 μm.

Figure S5

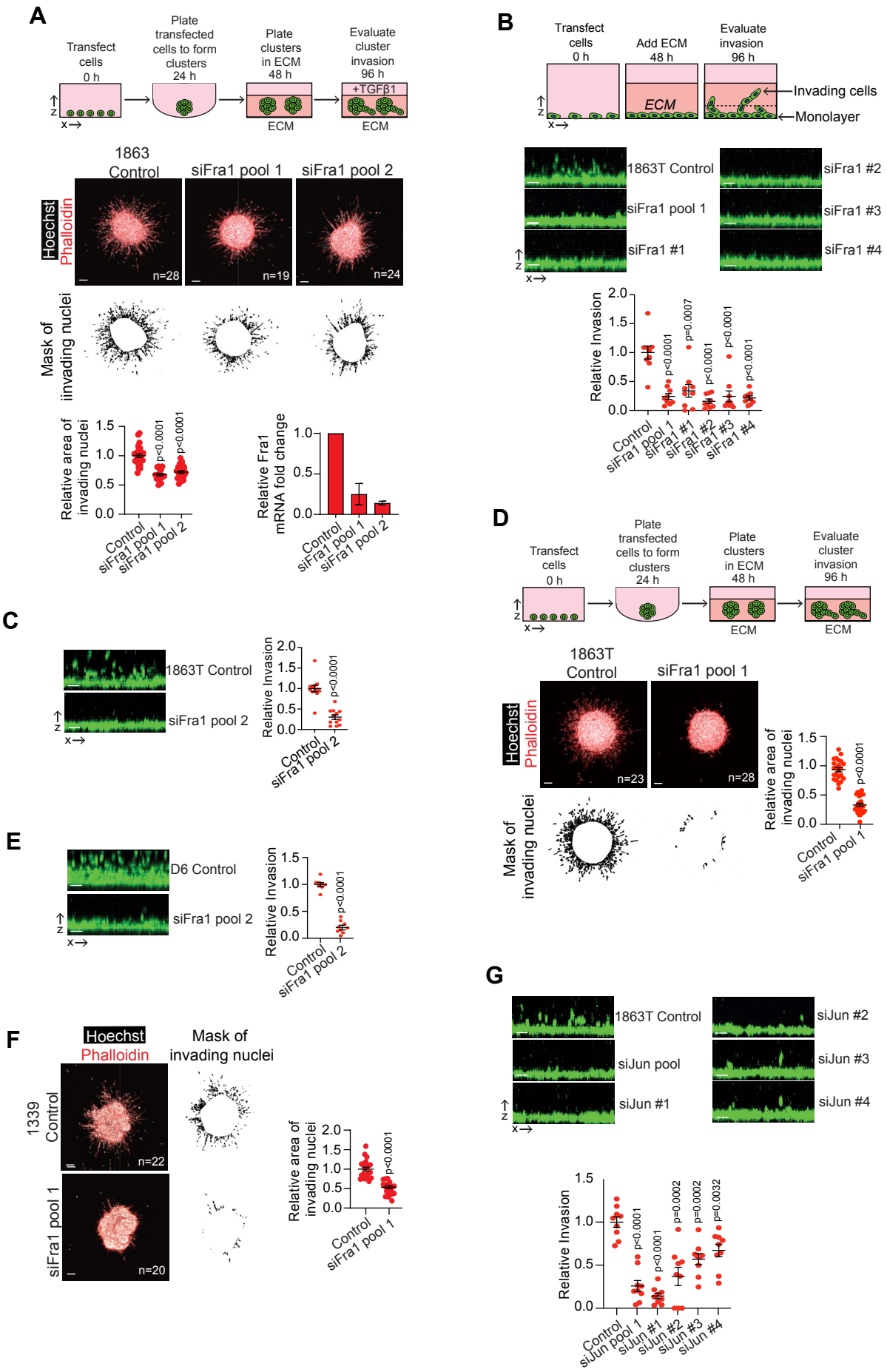

A
Transfect cells 0 h
Plate transfected cells to form clusters 24 h
Plate clusters in ECM 48 h
Evaluate cluster invasion 96 h
+TGFβ1
ECM
ECM
1863 Control
siFra1 pool 1
siFra1 pool 2
Hoechst
Phalloidin
n=28
n=19
n=24
Mask of iinvading nuclei
Relative area of invading nuclei
p<0.0001
p<0.0001
Control
siFra1 pool 1
siFra1 pool 2
Relative Fra1 mRNA fold change
B
Transfect cells 0 h
Add ECM 48 h
Evaluate invasion 96 h
ECM
Invading cells
Monolayer
1863T Control
siFra1 pool 1
siFra1 #1
siFra1 #2
siFra1 #3
siFra1 #4
Relative Invasion
p<0.0001
p=0.0007
p<0.0001
p<0.0001
p<0.0001
C
1863T Control
siFra1 pool 2
Relative Invasion
p<0.0001
D
Transfect cells 0 h
Plate transfected cells to form clusters 24 h
Plate clusters in ECM 48 h
Evaluate cluster invasion 96 h
ECM
ECM
1863T Control
siFra1 pool 1
Hoechst
Phalloidin
n=23
n=28
Mask of invading nuclei
Relative area of invading nuclei
p<0.0001
E
D6 Control
siFra1 pool 2
Relative Invasion
p<0.0001
F
Hoechst
Phalloidin
Mask of invading nuclei
1339 Control
siFra1 pool 1
n=22
n=20
Relative area of invading nuclei
p<0.0001
G
1863T Control
siJun pool
siJun #1
siJun #2
siJun #3
siJun #4
Relative Invasion
p<0.0001
p<0.0001
p=0.0002
p=0.0002
p=0.0032

Figure S5. A. Fra1 depletion by a second distinct Fra1 siRNA pool (pool 2) decreased the invasion of TGFβ1 treated 1863 clusters (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 µm. Bar graph shows the depletion of Fra1 expression by two different pools of Fra1 siRNA (mean±SEM, n=3 for pool 1 and n=2 for pool 2). **B.** 1863T vertical invasion was suppressed by four individual sequences from Fra1 siRNA pool 1 (mean±SEM, Mann Whitney test, n= 9 wells, r=3) Scale bars, 50 µm. **C.** Fra1 siRNA pool 2 reduced the vertical invasion of 1863T cells (mean±SEM, Mann Whitney test, n=12 wells, r=4). Scale bars, 50 µm. **D.** Fra1 depletion suppressed the invasion of 1863T clusters (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 µm. **E.** Fra1 siRNA pool 2 reduced the vertical invasion of D6 cells (mean±SEM, Mann Whitney test, n= 9 wells, r=3). Scale bars, 50 µm. **F.** Fra1 knockdown (pool 1) suppressed the invasion of TGFβ1 treated 1339 clusters (mean±SEM, Mann Whitney test, n=clusters, r=2). Scale bars, 50 µm. **G.** Jun depletion with 4 different siRNAs decreases the vertical invasion of 1863T cells (mean±SEM, Mann Whitney test, n=9 wells, r=2). Scale bars, 50 µm.

Figure S6

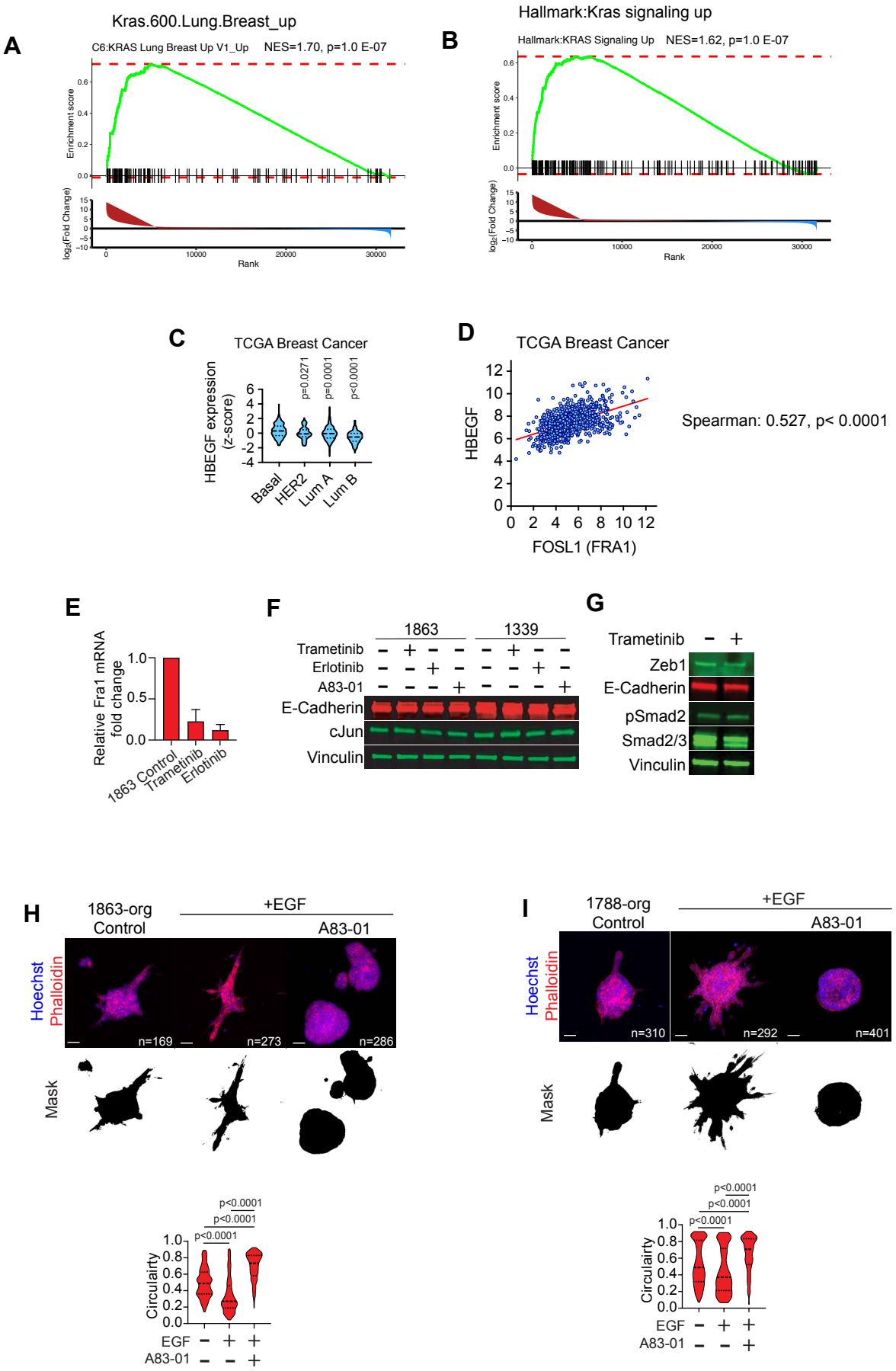

A
Kras.600.Lung.Breast_up
C6:KRAS Lung Breast Up V1_Up NES=1.70, p=1.0 E-07
Enrichment score
log2(Fold Change)
Rank
B
Hallmark:Kras signaling up
Hallmark:KRAS Signaling Up NES=1.62, p=1.0 E-07
Enrichment score
log2(Fold Change)
Rank
C
TCGA Breast Cancer
HBEGF expression (z-score)
p=0.0271
p=0.0001
p<0.0001
Basal
HER2
Lum A
Lum B
D
TCGA Breast Cancer
HBEGF
FOSL1 (FRA1)
Spearman: 0.527, p< 0.0001
E
Relative Fra1 mRNA fold change
1863 Control
Trametinib
Erlotinib
F
1863
1339
Trametinib
Erlotinib
A83-01
E-Cadherin
cJun
Vinculin
G
Trametinib
Zeb1
E-Cadherin
pSmad2
Smad2/3
Vinculin
H
1863-org
Control
+EGF
A83-01
Hoechst
Phalloidin
n=169
n=273
n=286
Mask
p<0.0001
p<0.0001
p<0.0001
Circulairty
EGF
A83-01
I
1788-org
Control
+EGF
A83-01
Hoechst
Phalloidin
n=310
n=292
n=401
Mask
p<0.0001
p<0.0001
p<0.0001
Circulairty
EGF
A83-01

Figure S6. A-B. GSEA was performed on RNA-seq data from C3-TAg organoids and non-invasive C3-TAg clones to identify pathway enrichment in C3-TAg organoids. Plots show the GSEA associated with Kras signaling. **C.** HBEGF mRNA is more highly expressed in the basal-like subtype compared to other types of breast cancer in the TCGA. Z-scores relative to all samples are shown (Kruskal Wallis test with Dunn's multiple comparisons test). P-values are in comparison to the basal-like tumors. **D.** FRA1 (FOSL1) correlates with HBEGF expression in breast cancer patient tumors. **E.** Trametinib (10 nM) and erlotinib (1 μM) reduced Fra1 expression in serum-starved 1863 cells (mean±SEM, n=2). **F.** Immunoblots showing that Jun expression was not perturbed by treatment with trametinib (10 nM), erlotinib (1 μM) or A83-01 (500 nM) in 1863 and 1339 cells. Representative of 2 (erlotinib, A83-01) and 3 (trametinib) independent experiments. **G.** Immunoblot showing that trametinib (10 nM) did not reduce Zeb1 expression or Smad2 phosphorylation (n=2). **H-I.** The Tgfbr1 inhibitor A83-01 suppressed the EGF induced invasion of 1863-org and 1788-org organoids. Violin plot shows the quantification of organoid circularity (Mann Whitney test, n=organoids, r=2).

Figure S7

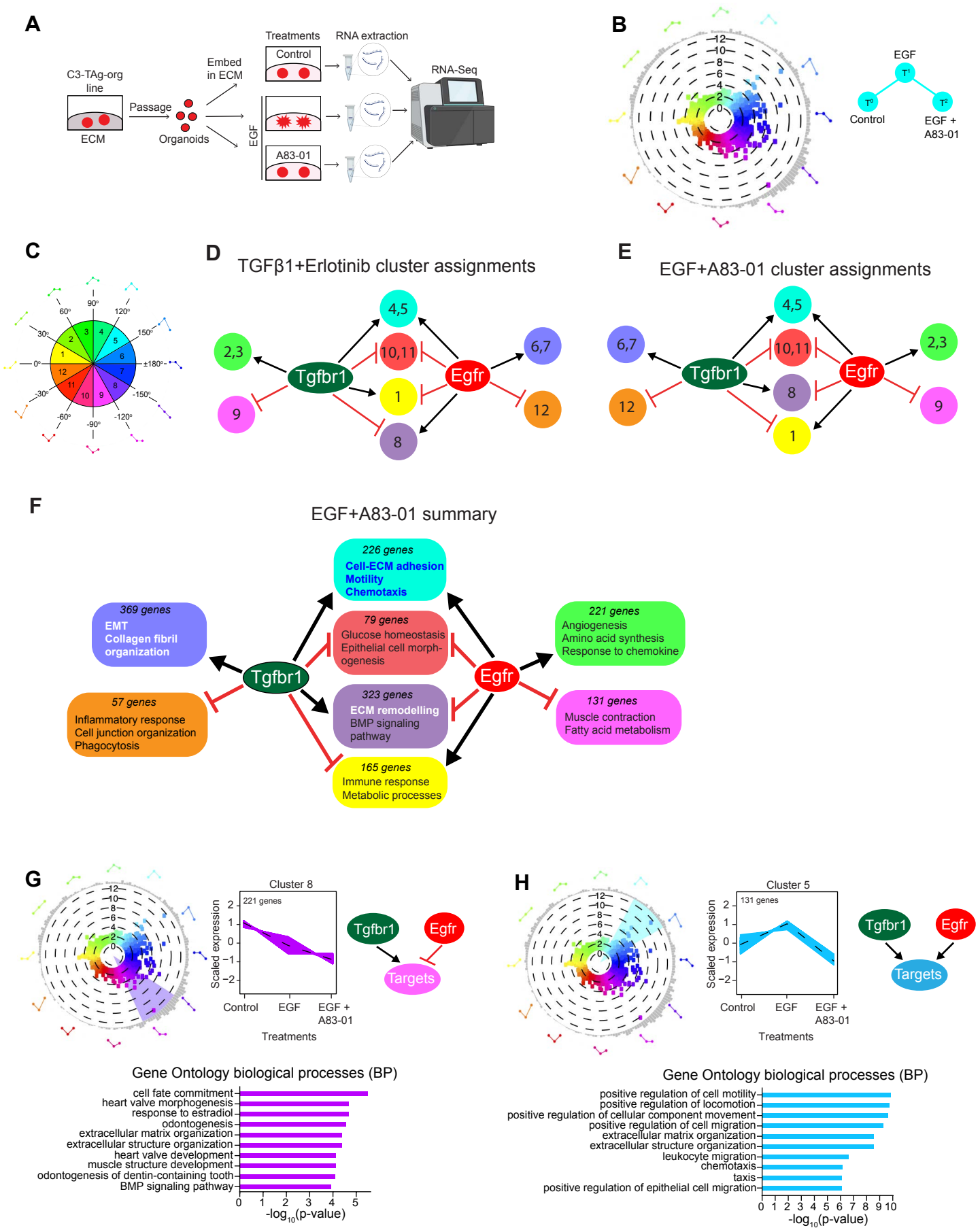

A
C3-TAg-org line
ECM
Passage
Organoids
Embed in ECM
Treatments
Control
EGF
A83-01
RNA extraction
RNA-Seq
B
EGF
Control
EGF + A83-01
C
D
TGFβ1+Erlotinib cluster assignments
Tgfbr1
Egfr
E
EGF+A83-01 cluster assignments
F
EGF+A83-01 summary
226 genes
Cell-ECM adhesion
Motility
Chemotaxis
369 genes
EMT
Collagen fibril organization
79 genes
Glucose homeostasis
Epithelial cell morph-ogenesis
221 genes
Angiogenesis
Amino acid synthesis
Response to chemokine
57 genes
Inflammatory response
Cell junction organization
Phagocytosis
323 genes
ECM remodelling
BMP signaling pathway
131 genes
Muscle contraction
Fatty acid metabolism
165 genes
Immune response
Metabolic processes
G
Cluster 8
221 genes
Scaled expression
Control
EGF
EGF + A83-01
Treatments
Targets
Gene Ontology biological processes (BP)
cell fate commitment
heart valve morphogenesis
response to estradiol
odontogenesis
extracellular matrix organization
extracellular structure organization
heart valve development
muscle structure development
odontogenesis of dentin-containing tooth
BMP signaling pathway
-log10(p-value)
H
Cluster 5
131 genes
positive regulation of cell motility
positive regulation of locomotion
positive regulation of cellular component movement
positive regulation of cell migration
extracellular matrix organization
extracellular structure organization
leukocyte migration
chemotaxis
taxis
positive regulation of epithelial cell migration

Figure S7. A. Overview of the approach to test the interactions between the Egfr and Tgfβ signaling pathways when cells were treated with EGF or EGF+A83-01. **B.** Polar plot organizing genes represented as individual dots based on their change in expression in response to treatment with EGF and co-treatment with EGF and A83-01. Genes that underwent ≥2 fold-change with a $p < 0.05$ after Benjamini-Hochberg correction in at least one condition relative to one of the other conditions are shown (3 total comparisons). The distance from the center of the plot represents the $\Delta\log_2$ value of each gene (increased difference in expression, further from the center). The bar plots on the surface of the circle show the number of genes at the degree point in the plot. Genes are grouped into 12 color-coded groups along 30 degree increments over the 360 degree plot. As an example, genes in group 5 were increased in expression by EGF and this increase in expression was inhibited by A83-01. In contrast, genes in group 1 were increased in expression by EGF, and this induction is further enhanced by treatment with A83-01. **C.** Plot shows the cluster locations along the 30 degree increments in the polar plots shown in **Fig. 6B** and **S6B**. **D.** Clusters assigned to different summary groups for TGFβ1 treated organoids shown in **Fig. 6B**. **E.** Clusters assigned to different summary groups for EGF treated organoids shown in **Fig. S6B**. **F.** Summary of the number of genes and represented biological processes associated with different modes of interaction between the Tgfbr1 and Egfr pathways in cells treated with stimulated EGF. Also see **Table S5** for the assignment of polar plot clusters and detailed biological processes used to define the summary integration groups. **G.** Egfr restricted the expression of a subset of genes induced by autocrine Tgfβ. Line plots show the expression of genes in Cluster 1. Dashed line indicates the average scaled expression of all Cluster 1 genes. Graph shows the top 10 biological processes (GO-BP) associated with Cluster 1 genes (Fisher's exact test). **H.** Tgfbr1 was necessary for the expression of a subset of genes induced by extrinsic EGF. Line plots show the expression of genes in Cluster 5. Dashed line indicates the average scaled expression of all Cluster 5 genes. Graph shows the top 10 biological processes (GO-BP) associated with Cluster 5 genes (Fisher's exact test).

Table S1. Description of GEM models, organoid/cell line models and reagents used in the study.

Table S2. RNA-Seq data (log2(CPM) genes x samples) from comparison between non-invasive clones and C3-TAg tumor organoids. Pathway enrichment analysis from comparison between non-invasive clones and C3-TAg tumor organoids

Table S3. Gene clusters and enriched biological processes from comparison between control and TGFβ1 treated C3-TAg 1339-org line organoids

Table S4. Gene clusters and enriched biological processes from comparison between control (T^0), TGFβ1 (T^1) and TGFβ1 + Erlotinib (T^2) treated C3-TAg 1339-org line organoids.

Table S5. Gene clusters and enriched biological processes from comparison between control (T^0), EGF (T^1) and EGF + A8301 (T^2) treated C3-TAg 1339-org line organoids.

Video S1. Live imaging of C3-TAg derived tumor organoids in organotypic culture (GFP, green, tumor cell). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with invasive heterogeneity in C3-TAg organoids shown in **Fig. 1A, S1B**.

Video S2. Live imaging of C3-TAg derived tumor dissociated organoids in organotypic culture. Bright field images were acquired for 5 h at 30 min intervals. The video corresponds with the intratumor invasive heterogeneity in C3-TAg tumor dissociated organoids shown in **Fig. S1B**.

Video S3. Live imaging of homogenous clusters of C3-TAg tumor derived non-invasive clonal cells B6 (H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with non-invasive B6 clusters shown in **Fig. 1B, S1E**.

Video S4. Live imaging of homogenous clusters of C3-TAg derived tumor cells. Bright field images were acquired for 14 h at 30 min intervals. The video corresponds with invasive C3-TAg tumor cells clusters shown in **Fig. 1B**.

Video S5. Live imaging of heterogeneous clusters of C3-TAg derived tumor cells and non-invasive clonal cells B6 (H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with opportunistic invasion of B6 cells induced by C3-TAg tumor cells as shown in **Fig. 1B**.

Video S6. Live imaging of heterogeneous clusters of C3-TAg derived invasive (D6) and non-invasive (B6) clonal cells (H2B: GFP, green, nuclei; H2B:mCherry, red, nuclei). Fluorescent and bright field images were acquired for 14 h at 30 min intervals. The video corresponds with opportunistic invasion of B6 cells induced by D6 cells as shown in **Fig. S1E**.

Video S7. Live imaging of PyMT derived tumor organoids in organotypic culture. Bright field images were acquired for 18 h at 30 min intervals. The video corresponds with less invasive PyMT organoids as shown in **Fig. 1F**.

Video S8. Live imaging of C3-TAg derived tumor organoids in organotypic culture. Bright field images were acquired for 18 h at 30 min intervals. The video corresponds with more highly invasive C3-TAg organoids as shown in **Fig. 1F**.

Video S9. Live imaging of C3-TAg derived 1339-org line organoids in serum starved media (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with the baseline invasion of C3-TAg derived 1339-org line organoids as shown in **Fig. 3C, S3C**.

Video S10. Live imaging of C3-TAg derived 1339-org line organoids in serum starved media + 500 nM A830-01 (Tgfbr1 inhibitor) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with suppression of 1339-org line organoid invasion by inhibiting autocrine TGFβR1 activity as shown in **Fig. S3C**.

Video S11. Live imaging of C3-TAg derived 1339-org line organoids in presence of exogenous TGFβ1 (48 h) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with TGFβ1 induced invasion in 1339-org line organoids as shown in **Fig. 3C, S3C**.

Video S12. Live imaging of C3-TAg derived 1339-org line organoids in presence of persistent exogenous TGFβ1 (Day 16) (H2B:GFP, green, nuclei). Fluorescent and bright field images were acquired for 18 h at 30 min intervals. The video corresponds with progressively increased TGFβ1 induced invasion in 1339-org line organoids as shown in **Fig. 3C**.