

# Spatiotemporal dynamics of tumor microenvironment remodeling

Kamil Lisek<sup>1\*, #</sup>, Ilan Theurillat<sup>1\*</sup>, Tancredi Massimo Pentimalli<sup>1,2\*</sup>, Svea Beier<sup>1</sup>, Daniel León-Periñán<sup>1</sup>, Anna Antonatou<sup>1</sup>, Serafima Dubnov<sup>3</sup>, Marion Müller<sup>4</sup>, Florian Hubl<sup>1</sup>, Artemis Xhuri<sup>4</sup>, Hanna Romanowicz<sup>5</sup>, Beata Smolarz<sup>5</sup>, Elodie Montaudon<sup>6</sup>, Sandra Raimundo<sup>7</sup>, Anca Margineanu<sup>7</sup>, Marie Schott<sup>1</sup>, Séverine Kunz<sup>8</sup>, Elisabetta Marangoni<sup>6</sup>, Nikos Karaïskos<sup>1</sup>, Mor Nitzan<sup>9,10,11</sup>, Walter Birchmeier<sup>4</sup>, and Nikolaus Rajewsky<sup>1, 2, 12,13, 14,15,16, #</sup>

1. Systems Biology of Gene Regulatory Elements, Berlin Institute for Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany
2. Charité, Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Berlin, Germany
3. The Edmond & Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
4. Signal Transduction in Development and Cancer, Max Delbrück Center for Molecular Medicine, Berlin, Germany
5. Department of Pathology, Polish Mother's Memorial Hospital Research Institute, Lodz, Poland
6. Laboratory of Preclinical Investigation, Translational Research Department, Institut Curie, Paris, France
7. Advanced Light Microscopy Facility, Max Delbrück Center for Molecular Medicine, Berlin, Germany
8. Technology Platform for Electron Microscopy, Max Delbrück Center for Molecular Medicine, Berlin, Germany
9. School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel
10. Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem, Israel
11. Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel
12. NeuroCure Cluster of Excellence, Berlin, Germany
13. German Center for Cardiovascular Research (DZHK), Site Berlin, Berlin, Germany
14. German Cancer Consortium (DKTK), Heidelberg, Germany
15. National Center for Tumor Diseases (NCT), Site Berlin, Berlin, Germany
16. Lead contact

\* These authors contributed equally to this work

# Corresponding authors. [rajewsky@mdc-berlin.de](mailto:rajewsky@mdc-berlin.de), [klisek@mdc-berlin.de](mailto:klisek@mdc-berlin.de)

## Abstract

During tumorigenesis, interactions between tumor and stromal cells progressively remodel the tumor microenvironment (TME) towards pro-tumoral functions. Understanding early TME remodeling dynamics is therefore crucial for developing interceptive therapies. However, clinical samples typically provide isolated, late tumorigenesis snapshots. To overcome this limitation, we generated triple-negative breast cancer mice that develop multifocal, asynchronous tumors along a continuous luminal-to-basal transdifferentiation trajectory. Ordering spatial transcriptomes from 100+ ducts along this trajectory reveals the spatiotemporal dynamics of TME remodeling and underlying molecular mechanisms. Cancer-associated myofibroblasts (myCAFs) emerge as key players in advanced tumors, where they orchestrate pro-invasive remodeling of the tumor-stromal interface. myCAFs are conserved in patient-derived xenograft models and steer tumor trajectories towards invasive phenotypes when co-injected with tumor cells in syngeneic mice. Our study shows that temporal ordering of spatially-resolved disease snapshots unravels some of the molecular “forces” that, starting from the cell-of-origin, propel cells/microenvironments along a disease trajectory.

## Introduction

Solid tumors are complex ecosystems where malignant cells are embedded in a specialized extracellular matrix (ECM) and interact with surrounding stromal cells. During tumorigenesis, cell–cell and cell–ECM interactions initiate and sustain the changes in cellular composition and molecular phenotypes that ultimately remodel homeostatic tissues towards a tumor-permissive microenvironment<sup>1,2</sup>. Given their critical role in tumor progression, interactions within the tumor microenvironment (TME) have emerged as an attractive therapeutic target across diverse tumor entities<sup>3–5</sup>.

While single-cell RNA sequencing has transformed our understanding of tumor and stromal molecular phenotypes<sup>6,7</sup>, the lack of spatial context hinders the systematic mapping and understanding of cellular interactions in the TME. Pioneering analyses based on H&E morphology<sup>8</sup> and classical antibody staining<sup>9,10</sup> – alongside recent spatial omics approaches – have revealed that specific TME architectures correlate with tumor progression<sup>11,12</sup> and response to therapy<sup>5,13,14</sup>. Very recently, high-resolution spatial transcriptomic (ST) technologies have enabled comprehensive gene expression profiling within intact tissue sections at single-cell and even subcellular resolution<sup>15</sup>. By capturing cell–cell and cell–ECM interactions in their native spatial context, these approaches have provided further insights into the mechanisms that stabilize multicellular niches or drive tumor progression<sup>16</sup>. It has been proposed that these and other spatially resolved data hold significant potential for informing mechanism-based, personalized therapeutic strategies in clinical oncology, including interception of disease trajectories<sup>2,17,18</sup>.

Nevertheless, spatial omics only provide a static snapshot, limiting insight into which interactions coordinate TME remodeling dynamics over time. Moreover, patient tissues are often late-stage and heterogeneous, reflecting variable genetics, disease trajectories, and treatment histories. In contrast, *in vitro* systems (*i.e.* organoids and 2D cultures) and genetically engineered mouse models enable studying tumorigenesis along genetically-defined trajectories. While *in vitro* systems may not fully recapitulate immune and stromal components of the TME<sup>19</sup>, mouse models can reveal the dynamic interplay between epithelial transformation and TME remodeling within native, immunocompetent tissues<sup>20</sup>.

Even in mouse models, collecting tumor samples to profile early tumorigenesis can be challenging. For example, if tumor development occurs as a single lesion and several weeks after oncogenic induction, early lesions – being particularly small – can be difficult to localize. Furthermore, numerous animals would need to be analyzed in order to reach adequate spatiotemporal resolution. Therefore, we set out to design a mouse model that would combine (1) a relation to an aggressive human cancer where the TME represents an attractive therapeutic target, such as triple-negative breast cancer (TNBC)<sup>21–26</sup>, (2) genetically-defined tumor development to facilitate ordering individual tumors along a shared trajectory<sup>20,27,28</sup>, (3) fast tumor development minimizing the impact of somatic mutations and (4) genesis of multiple asynchronous tumors within the same gland to facilitate high-resolution, spatiotemporal TME profiling.

We reasoned that such a model could be generated by introducing three oncogenic mutations, commonly altered in human TNBC<sup>29–31</sup>, activating p53, PI3K, and WNT pathways into luminal epithelial cells of the mammary gland. This combined oncogene activation

jumpstarted tumorigenesis, inducing multiple, spatially-distinct tumors within the same gland and enabling the parallel analysis of dozens of healthy ducts, early and advanced tumors and their microenvironments in individual samples. Using high-resolution, unbiased spatial transcriptomics (Open-ST<sup>32</sup>), we generated a unique dataset encompassing 100+ ducts, which we complemented with single-nucleus RNA sequencing from independent animals. Ordering individual ducts along a shared tumor progression trajectory revealed TME remodeling dynamics from tumor initiation to tumor invasion. At the tumor-stromal interface, stromal phenotypes were organized in multicellular niches, which were dynamically remodeled closely following changes in intraductal tumor phenotypes and the emergence of transient cell–cell interactions.

A specific population of cancer-associated myofibroblasts (myCAFs) tightly wrapped around advanced tumor lesions and orchestrated ECM remodeling at the invasive margin. Co-transplantation of tumor cells in syngeneic mice confirmed the ability of myCAFs to steer tumor trajectories towards invasive phenotypes. Finally, we validated myCAFs emergence in response to human TNBC cells with different mutation profiles using patient-derived xenografts (PDX) models, suggesting that dynamics we identified may have translational relevance for targeting the TME in human TNBC.

Altogether, our approach allows capturing and quantifying the remodeling of the TME from very early events (cell-of-origin) to later stages (invasive tumors) - in space & time. We show that these data provide a better identification and understanding of the underlying molecular mechanisms that drive tumorigenesis.

## Results

### A murine model to study TME remodeling dynamics in TNBC

To model TME dynamics in TNBC, we generated murine models conditionally activating and *Pik3ca*<sup>H1047R</sup> and *Trp53*<sup>R172H</sup> oncogenic alleles<sup>33,34</sup>, together with an oncogenic, stabilized form of  $\beta$ -catenin (*Ctnnb1* <sup>$\Delta$ ex3</sup>)<sup>35</sup> in the mammary epithelium. Oncogenes were activated in luminal epithelial cells using the Cre recombinase under control of the *Wap* promoter (WAPiCre)<sup>36</sup>.

Recombined cells were co-labeled with enhanced yellow fluorescent protein (YFP) via a *Rosa26*<sup>eYFP</sup> reporter allele<sup>37</sup>, allowing tumor cell tracing (**Fig1A**). Activation of single oncogenes led to tumor development in all animals between 25 and 41 postnatal weeks (**SFig1A**). Instead, control animals, which carried all oncogenic alleles but lacked WAPiCre, never developed tumors. In triple-mutants, combined oncogene activation anticipated tumor onset to five postnatal weeks (**Fig1B-C**). Tumors then grew rapidly across all mammary glands (**SFig1B**), reaching size limits defined in the animal protocol at 10 postnatal weeks. At study termination, tumors featured a TNBC immunohistochemistry profile, lacking estrogen, progesterone and HER2 receptors (**Fig1D**). Accordingly, transformed lesions exhibited basal features, namely high basal (Krt14) and low luminal (Krt8) markers. Instead, YFP detection was restricted to Krt8+ cells in 5-week-old mice (**SFig1C**), confirming tumor initiation in the luminal compartment and suggesting the acquisition of basal features with tumor progression. At study termination, whole-gland hematoxylin and eosin (H&E) staining revealed multiple ducts at various stages of transformation separated by the adipocyte-rich stroma typical of healthy glands (**Fig1E**). In the same gland, ‘healthy’ ducts lined by a thin layer of fibroblasts (1) could still be detected together with ‘hyperplastic’ ducts (2), ‘*in situ*’ keratinized ducts confined by a fibrotic layer (3) and ‘invasive’ lesions where nests of tumor cells infiltrated the fibrotic stroma (4) (**Fig1F**). Confirming rapid tumor progression, whole-gland immunofluorescence (IF) highlighted nests of tumor cells in the intramammary lymph node when included in the same tissue section (**SFig1D**).

In summary, we generated a fast, reproducible and aggressive model of TNBC to study the dynamics of tumor and microenvironment remodeling from healthy ducts to invasive lesions.

### Spatiotemporal study of TME remodelling at single cell resolution

To study tissue remodeling throughout tumorigenesis, we combined snRNAseq with high-resolution, genome-wide (polyA-based) spatial transcriptomics (Open-ST<sup>32</sup>) on 24 abdominal mammary gland samples across one control (T0) and three tumor timepoints (T1: 5–6 weeks, T2: 8–9 weeks and T3:10–11 weeks) (**Fig1G, SFig1E**).

Unbiased clustering of 58,742 single-nuclei revealed 15 cell types (**SFig2A-B**) spanning epithelial, stromal, and immune compartments (**Fig2A, STable1**). The epithelial compartment included luminal, hormone sensing and myoepithelial cells as well as transcriptionally distinct tumor cell states not detected in control samples. Stromal populations consisted of fibroblasts, pericytes, endothelial cells, and adipocytes. The immune compartment included both lymphoid (B and T cells) and myeloid lineages, including macrophages, monocytes, plasmacytoid and migratory dendritic cells. To investigate TME molecular organization, we utilized our ST data. Unbiased clustering of 587,288 pseudocells (i.e.  $\sim 120\mu\text{m}^2$  edge hexagonal bins, Methods) identified regions characterized by epithelial, stromal, and immune marker gene expression (**Fig2B, SFig2C**).

To map snRNAseq populations to their tissue locations, we leveraged Robust Cell Type Deconvolution<sup>38</sup> (**Fig2C**). Our integrated analysis captured the progressive remodeling of gland architecture along tumorigenesis (**Fig2D**). While control samples only featured healthy ducts surrounded by a thin layer of fibroblasts and an adipocyte-rich stroma, gene expression profiling confirmed that healthy ducts co-existed with multiple transformed ducts in tumor samples. From T1 to T3, transformed ducts increased in size and were surrounded by abundant fibroblasts and macrophages, which replaced adipocytes in the periductal stroma. Macrophages also accumulated in proximity and inside transformed ducts. Of note, none of the ducts in ST sample T1#3 showed evidence of tumor transformation, suggesting that tumor initiation may not be fully synchronized across all animals (**SFig2C**). Surprisingly, while lymphocytes composed ~30% of dissociated cells, they were rare in the periductal stroma. In fact, analysis of individual single-nuclei samples revealed that several immune and stromal populations were either always present or absent together (**SFig2D**). Compatible with intramammary lymph nodes being at times captured upon tissue dissociation, these immune and stromal populations mapped exclusively to the intramammary LN captured in ST sample T2#3 (**SFig2E**). Therefore, integrating ST and snRNA-seq data excluded potential confounding effects brought about by tissue dissociation and highlighted an immune desert TME.

Altogether, the combination of snRNA-seq and ST allowed to map the organization of cellular populations in the TME.

### Epithelial and stromal cells undergo extensive remodeling in the TME

To study the phenotypic remodeling of cellular populations in the TME, we proceeded to investigate epithelial and stromal transcriptional states captured by snRNAseq throughout tumorigenesis.

Unbiased clustering of 5,428 epithelial cells identified 12 populations, which we annotated based on their enrichment in CTRL vs mutant samples (**Fig3A**), the expression of canonical marker genes (**Fig3B**, **STable2**) and reference-based label transfer from a murine healthy and transformed mammary gland atlas<sup>39</sup> (**SFig3A**). While CTRL samples featured hormone-sensing (*Esr1+* *Pgr+* *ErbB2+*), luminal (*Elf5+*), and myoepithelial cells (*Acta2+*), multiple populations were only detected in mutant samples despite data integration (**SFig3B**). *Wap* expression, which drives oncogene activation, was specific to *Csn2+* alveolar luminal cells (**SFig3C**). This population, which we annotated as ‘cell-of-origin’, was already present but rare in CTRL samples and expanded during tumorigenesis. Again, label transfer scores and *Krt8* expression confirmed tumor initiation in the luminal compartment. Instead, further tumor populations showed decreasing levels of luminal and increasing level of basal markers (i.e. *Krt14* and *Trp63*). Despite their shared genotype, basal-like tumor populations acquired diverse phenotypes, distinguished by keratinization, Wnt signaling and EMT (epithelial-to-mesenchymal transition) signatures (**Fig3C**). Pseudotime analysis starting from the cell-of-origin (Methods) ordered tumor cells along continuous trajectory and revealed divergent endpoints, which corresponded to the activation of different molecular pathways in basal-like tumor cells (**Fig3D**). Consistent with such luminal-to-basal (L2B)<sup>40,41</sup> transition, IF showed widespread *Krt8+*/YFP+ luminal tumor cells in hyperplastic lesions, while YFP+/Trp63+/Krt8- basal tumor cells characterized invasive lesions (**Fig3E**, **SFig3D**). In contrast to hormone-sensing cells, which persisted as a stable, non-transformed

population, myoepithelial cells were remodeled in mutant samples. Compared to myoepithelial cells isolated from CTRL samples, tumor-associated basal cells (TABACs) upregulated *Trp73* expression and sonic Hedgehog pathway activity (**Fig3F**), downregulated *Acta2* while retaining the expression of canonical basal markers (i.e. *Trp63*) (**SFig3E**). Consistent with non-recombined, tumor-reprogrammed myoepithelial cells, IF identified a monolayer of YFP-/Trp73+ cells surrounding tumor lesions (**Fig3G, SFig3F**). Altogether, oncogene activation gradually transformed luminal alveolar cells into divergent basal-like tumor states and led to myoepithelial cell remodeling into TABACs.

Besides epithelial cells, we sought to identify which stromal phenotypes emerged in the TME (**SFig3G**). Compatible with their tissue-resident role, two fibroblast ('Fibroblasts 1' and 'Fibroblasts 2') and one macrophage population ('Macrophages/MΦ healthy') were enriched in CTRL samples (**Fig3H**). Conversely, fibroblasts in mutant samples expressed cancer-associated signatures<sup>42</sup> (**SFig3H**), supporting their annotation as inflammatory cancer-associated fibroblasts (iCAFs) and myofibroblastic CAFs (myCAFs). Similarly, two macrophage populations featured the activation of a tumor-associated macrophage (TAM) signature. In our model, *Chl1*, *Ncam1*, *Gpnmb* emerged as specific markers for iCAFs, myCAFs and TAMs, respectively (**Fig3I, SFig3I, STable3**). A population of *Mrc1*+, tumor-enriched macrophages ('Macrophages tumor') did not express TAM signatures, likely representing macrophages being recruited but not yet reprogrammed in the TME.

Overall, single-cell analysis captured the progressive remodeling and the acquisition of diverse molecular phenotypes in transformed glands. While oncogenes drive the remodeling of luminal epithelial cells, the emergence of diverse tumor (**Fig3C**) and stromal states (**Fig3J**) suggests a role for local interactions in shaping molecular phenotypes in the TME.

### **Tumor and stromal states are spatially organized in dynamic multicellular niches**

To investigate the organization of tumor and stromal phenotypes in the TME, we mapped single-nuclei populations in space (**SFig4A**).

Unbiased clustering of average RCTD deconvolution scores in cellular neighborhoods (**Fig4A**) identified 13 multicellular niches, ranging from tissue-resident to tumor-specific microenvironments (**Fig4B, SFig4B**). As a positive control, hormone sensing, luminal and myoepithelial cells colocalized in a specific niche, which we annotated as 'healthy duct'. Similarly, endothelial cells and pericytes colocalized in the 'blood vessel' niche. Vascular cells and adipocytes were enriched in the 'adipocyte stroma' and -together with fibroblasts 1- in the 'healthy periductal stroma' niches. Instead, Fibroblasts 2 were particularly abundant in the 'fibrous stroma' niche. Spatial mapping (**Fig4C, SFig4C**) confirmed the alignment of multicellular niches with healthy ducts, blood vessels, adipocyte stroma evident from tissue histology (**Fig2B**) and associated the 'fibrous stroma' niche with the gland capsule captured in T1#2 (**SFig4C**). Integrating single-nuclei and ST revealed how tumor-remodeled fibroblast and myeloid populations were closely associated in space, with iCAFs and tumor macrophages colocalizing in the 'iCAF/macrophage niche' and myCAFs and TAMs 1 in the 'myCAF/TAMs1 niche'. Instead, TAMs 2 were enriched in a specific 'TAM 2 niche'. As expected, immune populations mostly colocalized with lymph node-related stromal cell types to the 'lymph node niche'. Of note, monocytes, B and dendritic cells also colocalized to discrete 'immune foci' infrequently detected in tumor samples. Furthermore, the spatial

organization of tumor states closely followed their pseudotime. Early states including cell-of-origin and transdifferentiating tumor cells were enriched in the ‘Luminal tumor niche’, while late, basal phenotypes (i.e. keratinizing, Krt63+ and Wnt basal tumor populations) in the ‘Basal-like tumor niche’ mapped to ducts whose enlarged lumen was filled by keratin pearls. Instead, EMT-like tumor cells did not localize to the basal-like tumor niche but rather with TABACs, myCAFs and TAMs 1 in the myCAFs/TAMs 1 niche, compatible with a role for the local microenvironment in promoting invasive tumor phenotypes.

To understand the dynamics of niche remodeling during tumorigenesis, we mapped tumor pseudotime scores in space (Methods). As a control, healthy ducts and the periductal stroma featured near-zero pseudotime, while a range of positive pseudotime scores was detected across transformed ducts (**Fig4D**, **SFig4D**). We thus proceeded to segment (**Fig4E**, **SFig4E**) 111 individual ducts and sort them along their average pseudotime scores. The independent mapping of tumor pseudotime and multicellular niches highlighted a continuum of duct transformation (**Fig4F**), partitioned in 5 stages: ‘Stage 0’ (Average duct pseudotime from 0,0 to 0,15), I (0,16-0,31), II (0,32-0,47), III (0,48-0,62) and IV (0,63-0,78).

In summary, spatial mapping of epithelial and stromal phenotypes revealed their colocalization in multicellular niches. Pseudotime scoring and duct segmentation then revealed progressive niche remodeling along tumor progression.

### **Spatiotemporal dynamics of tumor-stromal interface remodeling**

Having ordered 100+ ducts along a shared tumorigenesis trajectory, we proceeded to investigate how the duct-stromal interface (**Fig5A**) was remodeled throughout tumor progression.

Supporting the robustness of spatial mapping, cell state dynamics (**SFig5A**) closely followed marker gene expression levels (**SFig5B**). Epithelial state transitions in the duct core followed tumor pseudotime defined at the single-cell level (**Fig5B**). Namely, luminal and myoepithelial populations characterized Stage 0 ducts (i.e. healthy), while cell-of-origin cells peaked upon tumor initiation (Stage I). With tumor progression, cell-of-origin cells were rapidly replaced by transdifferentiating tumor phenotypes first (Stages II-III), and by keratinizing and EMT phenotypes later (Stages III-IV). TABACs were detected from tumor initiation throughout tumor progression (Stages I-III) but were absent in Stage IV ducts, suggesting a loss of basement membrane integrity in advanced tumors. We then analyzed how stromal phenotypes were remodeled in response to intraductal epithelial changes (**Fig5C**). While healthy ducts were surrounded by fibroblasts and adipocytes (Stage 0), the interface was rapidly populated by iCAFs and macrophages upon tumor initiation (Stage I) and by myCAFs and TAMs 1 with tumor progression (Stages II-IV). Instead, TAMs 2 were detected at the tumor-stromal interface only transiently in Stage III tumors. In advanced tumors, however, we observed TAMs 2 accumulation in the duct core. Furthermore, iCAFs and macrophages remained abundant, but were predominantly located in the periductal stroma distal to the myCAF–TAM interface (**Fig5D**). We thus sought to quantify distances of stromal phenotypes to the duct center along tumor progression (Methods). This confirmed how stromal remodeling occurred along a reproducible spatiotemporal gradient: first iCAFs and macrophages were recruited and then populated the distal periductal stroma as myCAFs and TAMs 1 replaced them at the duct border with tumor progression (**Fig5E**). On the other

hand, the distance of TAMs 2 to the duct border increased with tumor progression, both towards the duct center and in the distal periductal stroma, compatible with their localization to TAM 2 niches. To confirm the remodeling of stromal states, we analyzed their dynamics in individual cellular neighborhoods using phase diagrams (Methods). Paired abundances at the tumor-stromal interface revealed the progressive transition from a macrophage-high/iCAF-high state in tumor initiation (Stage I) towards a myCAF/TAM rich state during tumor progression (Stages II-III) (**Fig5F**). Of note, such macrophage-high, fibroblast-high ‘hot-fibrosis’ state differed from the macrophage-low, fibroblast-low ‘healing’ state detected at the interface of healthy ducts (**SFig5C**). Finally, IF staining for population-specific markers confirmed the periductal stromal architecture in an intermediate tumor where Ncam1<sup>+</sup> myCAFs lined the tumor border, while Chl1<sup>+</sup> iCAFs were distributed more broadly in the periductal stroma (**Fig5G**, **SFig5D**). At the same time, Mrc1<sup>+</sup> macrophages were closely associated with iCAFs (**SFig5E**), while Gpnmb<sup>+</sup> TAMs accumulated in the duct center. Furthermore, electron microscopy confirmed the ultrastructural remodeling of the duct-stromal interface during tumorigenesis. While a clearly defined basement membrane surrounded stratified, cuboidal epithelial cells in a healthy duct, the basement membrane was no longer detected in an advanced tumor (**SFig5F**). Instead, elongated fibroblasts and numerous myeloid cells surrounded tumor cells with irregular nuclear morphology and keratin deposition, compatible with basement membrane disruption at the tumor–stroma interface.

Patient-derived xenograft (PDX) models complement mouse models by preserving key molecular and histological features of human tumors, offering a direct link to clinical heterogeneity and therapeutic response. We therefore investigated whether tumor cells from TNBC patients would elicit a similar stromal reaction in a cohort of PDX models<sup>43</sup> including but not limited to lines harboring the same genetic alterations (**STable4**). IF staining revealed that Ncam1<sup>+</sup> myCAFs tightly wrapped around PDX tumors, while Chl1<sup>+</sup> iCAFs and Mrc1<sup>+</sup> macrophages colocalized in the peritumoral stroma (**Fig5G**). Of note, TAMs 2 were rare in PDX models compared to primary tumors. With the exception of TABACs, which formed a thin layer between tumors and myCAFs in primary lesions (**Fig5H**, **SFig5G**), the remodeling of the tumor-stromal interface was shared between primary mouse and PDX tumors regardless of tumor genetics.

Overall, we identified and validated the remodeling of stromal populations at the tumor-stromal interface (**Fig 5I**) - this prompted us to investigate which communication events underlie such stromal dynamics.

### Coordinated and transient signaling axes orchestrate early interface remodeling

To identify which cell–cell interactions orchestrate tumor–stromal interface remodeling, we quantified the co-expression of known ligand-receptor (LR) pairs in cellular neighborhoods<sup>44</sup> (**Fig6A**, Methods).

Focusing on 390 secreted interactions, we first identified 49 stage-specific ligands enriched at the duct-stromal interface (**SFig6A**). We then quantified LR expression in single-nuclei populations to identify potential senders and receivers of such interactions and reconstruct stage-specific communication networks. Upon tumor initiation (Stage I), our integrated communication analysis revealed the coordinated induction of *Tgfb2*, *Csf1* and *Ccl6* and *Ccl8* signaling at the tumor-stromal interface (**Fig6B**). Compatible with early remodeling,

ligand activity peaked (Stage I) and quickly dropped following tumor initiation (Stage II-IV). In our single-nuclei data, *Tgfb2* ligand expression was specific to the cell-of-origin, while Tgfb receptor (*Tgfb1*, *Tgfb2*) expression was detected across all stromal populations (**SFig6B**). Instead, *Csf1* signaling was specific to periductal fibroblasts and iCAFs (*Csf1+*) and macrophages (*Csf1r+*). In turn, macrophages acted both as senders and receivers of *Ccl6* and 8 signaling, suggesting a positive feedback loop reinforcing macrophage recruitment<sup>42</sup> (**Fig6C**). After peaking at the early tumor interface (Stage I), *Tgfb2* activity decreased with tumor progression (Stages II-III) to increase again in Stage IV. In fact, *Tgfb2* expression was downregulated in intermediate tumor states but then detected again in EMT tumor cells. Furthermore, myCAFs also acted as sources of TGF- $\beta$  at the advanced tumor-stromal interface, where TAMs 1 emerged as the main TGF- $\beta$  recipients (**SFig6C**). To assess intracellular responses to TGF- $\beta$  signaling<sup>46</sup>, we quantified TGF- $\beta$  pathway activity in single-nuclei populations (Methods). Gene set expression increased from fibroblasts to iCAFs and myCAFs, from tumor macrophages to TAMs and from healthy myoepithelial cells to TABACs (**Fig6D**).

Therefore, TGF- $\beta$  signaling was upregulated in all interface populations as they transitioned from healthy to the tumor-specific phenotypes, supporting the central role of TGF- $\beta$  in the remodeling of cellular phenotypes at the tumor-stromal interface.

### **Myofibroblasts orchestrate pro-invasive ECM remodeling at the tumor-stromal interface**

Besides secreted signals, cell-ECM interactions play a central role in TME remodeling<sup>47</sup>. We thus extended our spatial communication analysis to ECM ligand-receptor pairs<sup>44</sup> (**Fig6E**).

This revealed dynamic changes in 50 ECM interactions at the tumor-stromal interface (**SFig6D**). In particular, Tenascin-X (*Tnxb*) activity, which is linked with tumor restraining properties<sup>48</sup>, peaked upon tumor initiation (Stage I) and declined sharply with tumor progression (Stages III-IV), while Tenascin-C (*Tnc*) - its tumor-specific counterpart<sup>49,50</sup> - linearly increased together with *Col1a1* and *Lamc2* signaling, peaking at the interface of advanced tumors (Stage IV) (**Fig6F**). While *Col1a1* was expressed by all fibroblast states -although at highest levels in myCAFs-, we observed a shift in tenascin expression from *Tnxb* in fibroblasts and iCAFs to *Tnc* in myCAFs. EMT-like tumor cells also featured *Tnc* expression and emerged as the only population expressing *Lamc2*, a recently reported marker of invasive tumor cells<sup>51</sup> (**Fig6G**). To assess cellular contributions to ECM remodeling, we quantified 'ECM organization' pathway activity in snRNAseq populations. Signature scores increased at the transition from fibroblasts to iCAFs and peaked in myCAFs, supporting their central role in ECM remodeling (**Fig6H**). Among tumor phenotypes, EMT-like tumor cells displayed the highest pathway scores, compatible with their active role in shaping the ECM in advanced tumors. In contrast, macrophages—including TAMs—contributed minimally to this process. In the TME, ECM organization was spatially restricted to the tumor-stromal interface (**Fig6I**, **SFig6E**), which was validated by second harmonic generation (SHG) imaging, a label-free method to quantify fibrillar collagen in tissue sections<sup>52</sup>. While collagen was limited to a thin layer around healthy ducts in CTRL glands, SHG revealed a marked increase in collagen deposition around transformed ducts and within the dysplastic stroma in a T3 gland (**Fig6J**). IF staining of the consecutive section then confirmed the colocalization of *Ncam1+* myCAFs

to regions of collagen deposition (**Fig6K**, **SFig6F**). Notably, IF also highlighted infiltrating Lamc2+/YFP+ tumor cells within a dense network of Ncam1+ cells (**Fig6L**, **SFig6G-H**), confirming the tight spatial association of myCAFs with invasive, EMT-like tumor cells. Furthermore, IF highlighted Tnc deposition at the tumor-stromal interface in both primary (**SFig6I-J**) and PDX tumors (**Fig6M**, **SFig6K**), revealing a consistent ECM remodeling phenotype across genotypes.

Overall, while in the tumor initiation TGF- $\beta$  emerged as a main transforming factor of TME, myCAFs emerged as the principal drivers of ECM remodeling at the tumor–stroma interface, depositing a collagen- and Tnc-rich matrix that encapsulates invasive tumor cells (**Fig6N**).

### Myofibroblasts promote tumor growth and invasion

The role of myCAFs in cancer remains debated, with studies supporting both tumor-suppressive and tumor-promoting functions<sup>53,54</sup>. Given their colocalization with invasive tumor cells at the tumor-stromal interface in our model, we set out to experimentally test their impact on tumor growth.

To do so, we first established and characterized 3D organoid cultures from healthy and tumor-bearing mice, as well as fibroblast cultures from matched tissues. Organoids from healthy tissue formed polarized acinar structures composed of a single layer of Trp63+ basal cells encasing Krt8+ luminal cells, organized around a central lumen (**SFig7A**). In contrast, tumor-derived organoids displayed a disrupted architecture, with loss of mature luminal cell identity, diffused Trp63 expression, and protrusions extending into the surrounding matrix (**SFig7B**). Furthermore, electron microscopy confirmed disrupted tumor organoid architecture at the ultrastructure level (**SFig7C**). Consistent with the presence of iCAFs and myCAFs *in vivo*, fibroblast cultures from tumor-bearing mice included distinct Ncam1+ and Ncam1- subpopulations (**SFig7D**) and featured a heterogeneous morphology (**SFig7E**). For functional *in vivo* assessment, we orthotopically transplanted YFP+ luminal progenitor tumor cells<sup>55</sup> (**SFig7F**) into the mammary fat pad of immunocompetent, syngeneic recipient mice either alone (Group A), together with Ncam1- fibroblasts from healthy controls (Group B), or tumor-derived Ncam1+ myCAFs (Group C). Tumors in Groups A and B either failed to develop palpable lesions or formed only small, slowly growing masses. In contrast, all Group C mice developed significantly larger and heavier tumors (**Fig7A**, **SFig7G**). Histological analysis revealed that tumors from Groups A and B resembled early-stage disease, whereas Group C tumors displayed a disorganized architecture with nests of epithelial cells invading the surrounding stroma (**Fig7B**). To investigate the impact of myCAFs on TME molecular architecture, we performed Open-ST on a selected tumor from each group. While iCAFs were abundant in the peritumoral stroma in Group A and, to a lesser extent, Group B tumors, they were rare in Group C (**Fig7C**). Instead, myCAFs featured the opposite pattern: they were rare in Group A, limited to a peritumoral ring in Group B, while they were abundant and intermixed with tumor cells in Group C tumors. Of note, B cells were abundant in the myCAF-poor periductal stroma of Group A tumors but were absent in Group B and C (**SFig7H**), suggesting immunosuppressive effects of myCAFs on the TME of the injected mice. Similar to their localization in primary tumors, macrophages were abundant in the peritumoral stroma of Groups A and B (**SFig7I**), while TAMs 1 occupied the tumor-stromal interface and TAMs 2 formed hotspots in the tumor core across all groups, being particularly abundant in Group C (**SFig7J**). Tumor cells featured late pseudotime in all groups (**SFig7K**),

however Groups A and B featured abundant Keratinizing and Wnt<sup>+</sup> phenotypes (**SFig7L**), while tumor cells in Group C upregulated EMT signatures (**Fig7D**). ECM organization pathway activity was also increased in Group C and colocalized with myCAFs in space (**Fig7E**). While a thick collagen layer surrounded tumors in groups A-B, SHG revealed dense and disorganized collagen deposition inside Group C tumor (**Fig7F**). Finally, IF staining identified abundant Lamc2<sup>+</sup>/YFP<sup>+</sup> tumor cells invading a Tnc-rich stroma in Group C tumors (**Fig7G**, **SFig7M**).

In summary, our results demonstrate myCAFs ability to promote aggressive tumor growth, ECM remodeling and tumor invasion driven by EMT-like tumor cell states in an immunocompetent, syngeneic transplant setting.

## Discussion

Understanding TME remodeling from tumor initiation throughout progression is essential for early disease intervention or interception, however clinical samples typically provide isolated and heterogeneous snapshots of the tumorigenesis process. In our model, the simultaneous activation of three oncogenes—distinct from the stepwise accumulation of driver mutations in sporadic tumors—accelerated tumor development along a genetically-defined trajectory. Following oncogene activation, mammary epithelial cells underwent a consistent luminal-to-basal (L2B) transdifferentiation, in line with previous reports<sup>40,41,56,57</sup>. Besides L2B transdifferentiation, further trajectories could ensue as tumors in our model develop in a native, immunocompetent setting. For example, specific cellular phenotypes and multicellular niches may rapidly disappear from the TME through shedding or immune clearance, escaping detection in our 7-day sampling interval. However, we think these trajectories are rare or do not contribute to later stages - we have surveyed more than a hundred of progressions in ducts and only once saw a special case (of a duct completely covered with macrophages- probably an isolated event where the immune system was capable of killing the tumor). In other words, almost always, tumor development featured an early onset and rapid progression with conserved morphologies and reproducible molecular profiles across replicates and ducts, supporting the existence of a single, homogenous early trajectory. Hence, we used L2B transdifferentiation as a molecular clock to temporally order 100+ tumorigenesis snapshots and capture TME remodeling dynamics.

These data revealed that TME dynamics closely follow intraductal tumor phenotypes. Following tumor initiation, resident populations surrounding healthy mammary ducts underwent rapid remodeling: myoepithelial cells were reprogrammed to TABACs and periductal fibroblasts were replaced by iCAFs (early) and myCAFs (late). Mapping cell-cell interactions dynamics identified which signaling axes closely followed, hence likely drive, these events. Tumor-derived TGF- $\beta$ <sup>58</sup> emerged as the initial trigger of TME remodeling - among the hundreds of pathways analyzed- because (1) TGF- $\beta$  was upregulated by tumor cells soon after recombination (cell-of-origin), (2) TGF- $\beta$  receptors were expressed by resident fibroblasts and myoepithelial cells, (3) intracellular targets of TGF- $\beta$  signaling were upregulated in all remodeled populations at the tumor-stromal interface. While the concurrent activation of three oncogenes in our model confounds the identification of intracellular signals controlling TGF- $\beta$  upregulation, single-mutant mouse models may help disentangle the contribution of individual oncogenic pathways to epithelial and stromal remodeling. Besides resident populations, numerous macrophages were recruited to the tumor-stromal interface. iCAF-derived Csf1 likely represents the first trigger of macrophage recruitment, then amplified by the release of macrophage-derived chemokines and followed by TGF- $\beta$ -driven reprogramming to TAMs. As such early signaling axes were rapidly switched off upon tumor progression, profiling of advanced tumors may fail to identify the initial triggers of TME remodeling, which however represent attractive therapeutic targets. While the dynamics of duct-interface populations presented here support a sequential model of remodeling within the same lineage, transdifferentiation from distinct lineages cannot be excluded in the absence of lineage tracing data. For example, iCAFs preceded the appearance of myCAFs at the tumor-stromal interface and while TGF- $\beta$  has been shown to drive iCAF-to-myCAF transitions<sup>59–61</sup>, pericytes could also serve as a source of myCAF-like precursors<sup>62</sup>.

With tumor progression, basal-like tumor cells adopted divergent transcriptomic phenotypes despite sharing the same oncogenic drivers. While somatic mutations may be acquired by tumor cells during progression, several lines of evidence suggest that their contribution to tumor heterogeneity in our model is limited: (1) the short 28-day timeframe likely precludes genetic evolution, (2) prior mouse model studies have shown that phenotypic diversity can emerge from genetically-defined oncogenic inputs alone<sup>20,27,41</sup>, (3) tumor states were reproducibly organized in specific multicellular niches. Thus, our data indicate an active role of local microenvironments in instructing the fate of tumor cells. Consistently, keratinizing cells were confined to the ductal core in close proximity to Wnt-high populations, in line with the role of overactive Wnt signaling in driving squamous differentiation<sup>63,64</sup>. In contrast, EMT-like cells infiltrated the tumor–stroma interface and colocalized with myCAFs.

myCAFs emerged as the main ECM organizers at the duct stromal interface, contributing to collagen and TNC-rich deposition. These represent potent outside-in cues activating intracellular pathways that can promote tumor cell EMT and invasion<sup>65</sup>. When co-transplanted with tumor cells into syngeneic, immunocompetent mice, Ncam1<sup>+</sup> myCAFs were sufficient to shift tumor phenotypes away from keratinization towards EMT phenotypes. Beyond ECM proteins, matrix stiffening also plays a role in the emergence of aggressive phenotypes through YAP activation<sup>66</sup>. In our model, EMT-like tumor cells expressed *Nuak1*, a target and strong regulator of YAP signaling<sup>67</sup>. Therefore, future studies will be needed to disentangle the contributions of ECM proteins, matrix biomechanics and other myCAF-derived factors to EMT induction in this context. In addition, tumor-infiltrating lymphocytes were rarely detected in our model and transplant experiments. In line with the recently reported immunosuppressive role of Ncam1<sup>+</sup> myCAFs<sup>68–70</sup>, group A tumors –where myCAFs were scarce– represented the only exception, featuring abundant B cell infiltration in the peritumoral stroma.

Together, targeting myCAFs could represent an attractive therapeutic strategy to simultaneously unleash anti-tumoral immune responses and restrain tumor invasion in our model. Supporting their translational relevance, myCAFs were conserved across species and genetic backgrounds. In PDX models, the injection of human TNBC cells with different mutational backgrounds induced the emergence of Ncam1<sup>+</sup> myCAFs and TNC-rich ECM. This underlies the potential of TME targeting as a shared therapeutic strategy across TNBC oncogenic drivers. Spatial analysis of rare, early-stage and treatment-naïve TNBC specimens will be ultimately needed to assess the conservation of such pro-tumorigenic TME across TNBC molecular subtypes.

Beyond fibroblasts, myoepithelial cells and macrophages also emerged as key components of the tumor–stroma interface. The functional role of myoepithelial cells in cancer remains controversial, with evidence supporting both tumor-suppressive and pro-tumor functions<sup>71,72</sup>. In our model, TABACs may arise from Shh-induced myoepithelial remodeling and may participate in ECM organization. As for Ncam1<sup>+</sup> myCAFs, the specific expression of surface markers by TABACs (*i.e.* p73) and TAMs (*i.e.* Gpnmb), may be useful for isolating, characterizing and evaluating the functional impact of remodeled populations on tumor progression in syngeneic mouse models. Furthermore, injecting tumor cells in syngeneic mouse models limits tumorigenesis to a single lesion, offering the possibility for a longer follow up time within tumor size limits defined in the animal protocol. It would also allow for tumor and stromal manipulation before injection, enabling for example genetic

manipulations, lineage tracing and CRISPR screens. Furthermore, matched organoid cultures derived from primary, transplant and PDX tumors could be used to screen *in vitro* the impact of specific tumor-stromal interactions before their *in vivo* testing.

In summary, we have shown that temporal ordering of spatially resolved snapshots of a developing disease allows identifying some of the molecular “forces” that propel cells/microenvironments along a disease trajectory. We hope that this approach will be of value for the study of the largely unexplored universe of the early onset of diseases in general.

## Literature

1. De Visser, K. E. & Joyce, J. A. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell* **41**, 374–403 (2023).
2. Swanton, C. *et al.* Embracing cancer complexity: Hallmarks of systemic disease. *Cell* **187**, 1589–1616 (2024).
3. Bejarano, L., Jordão, M. J. C. & Joyce, J. A. Therapeutic Targeting of the Tumor Microenvironment. *Cancer Discov.* **11**, 933–959 (2021).
4. Kalluri, R. The biology and function of fibroblasts in cancer. *Nat. Rev. Cancer* **16**, 582–598 (2016).
5. Chen, Y., McAndrews, K. M. & Kalluri, R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat. Rev. Clin. Oncol.* **18**, 792–804 (2021).
6. Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat. Commun.* **8**, 15081 (2017).
7. Wagner, J. *et al.* A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell* **177**, 1330–1345.e18 (2019).
8. Yuan, Y. Spatial Heterogeneity in the Tumor Microenvironment. *Cold Spring Harb. Perspect. Med.* **6**, a026583 (2016).
9. Dawson, C. A. & Visvader, J. E. The Cellular Organization of the Mammary Gland: Insights From Microscopy. *J. Mammary Gland Biol. Neoplasia* **26**, 71–85 (2021).
10. Valenti, G. *et al.* Cancer Stem Cells Regulate Cancer-Associated Fibroblasts via Activation of Hedgehog Signaling in Mammary Gland Tumors. *Cancer Res.* **77**, 2134–2147 (2017).
11. Karimi, E. *et al.* Single-cell spatial immune landscapes of primary and metastatic brain tumours. *Nature* **614**, 555–563 (2023).
12. Sorin, M. *et al.* Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature* **614**, 548–554 (2023).
13. Li, Z. *et al.* Presence of onco-fetal neighborhoods in hepatocellular carcinoma is

- associated with relapse and response to immunotherapy. *Nat. Cancer* **5**, 167–186 (2024).
14. Hwang, W. L. *et al.* Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. *Nat. Genet.* **54**, 1178–1191 (2022).
  15. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nat. Methods* **19**, 534–546 (2022).
  16. Ji, A. L. *et al.* Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell* **182**, 497–514.e22 (2020).
  17. Pentimalli, T. M. *et al.* Combining spatial transcriptomics and ECM imaging in 3D for mapping cellular interactions in the tumor microenvironment. *Cell Syst.* **16**, 101261 (2025).
  18. Rajewsky, N. *et al.* LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature* **587**, 377–386 (2020).
  19. Kim, S.-Y., Van De Wetering, M., Clevers, H. & Sanders, K. The future of tumor organoids in precision therapy. *Trends Cancer* S2405803325000731 (2025)  
doi:10.1016/j.trecan.2025.03.005.
  20. Hill, W., Caswell, D. R. & Swanton, C. Capturing cancer evolution using genetically engineered mouse models (GEMMs). *Trends Cell Biol.* **31**, 1007–1018 (2021).
  21. Bianchini, G., De Angelis, C., Licata, L. & Gianni, L. Treatment landscape of triple-negative breast cancer — expanded options, evolving needs. *Nat. Rev. Clin. Oncol.* **19**, 91–113 (2022).
  22. Keren, L., Bosse, M., Marquez, D. A. *et al.* structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* **174**, 1373–1387.e19 (2018).
  23. Jackson, H., Fischer, J., Zanutelli, V. *et al.* The single-cell pathology landscape of breast cancer. *Nat. Genet.* **52**, 986–995 (2020).
  24. Wu, S. Z. *et al.* A single-cell and spatially resolved atlas of human breast cancers. *Nat.*

- Genet.* **53**, 1334–1347 (2021).
25. Wang, X. *et al.* Spatial transcriptomics reveals substantial heterogeneity in triple-negative breast cancer with potential clinical implications. *Nat. Commun.* **15**, 10232 (2024).
  26. Nolan, E., Lindeman, G. J. & Visvader, J. E. Deciphering breast cancer: from biology to the clinic. *Cell* **186**, 1708–1728 (2023).
  27. Wellenstein, M. D. *et al.* Loss of p53 triggers WNT-dependent systemic inflammation to drive breast cancer metastasis. *Nature* **572**, 538–542 (2019).
  28. Hollern, D. P., Swiatnicki, M. R. & Andrechek, E. R. Histological subtypes of mouse mammary tumors reveal conserved relationships to human cancers. *PLOS Genet.* **14**, e1007135 (2018).
  29. Costa, R. L. B., Han, H. S. & Gradishar, W. J. Targeting the PI3K/AKT/mTOR pathway in triple-negative breast cancer: a review. *Breast Cancer Res. Treat.* **169**, 397–406 (2018).
  30. Thennavan, A. *et al.* Molecular analysis of TCGA breast cancer histologic types. *Cell Genomics* **1**, 100067 (2021).
  31. Pohl, S.-G. *et al.* Wnt signaling in triple-negative breast cancer. *Oncogenesis* **6**, e310–e310 (2017).
  32. Schott, M. *et al.* Open-ST: High-resolution spatial transcriptomics in 3D. *Cell* **187**, 3953–3972.e26 (2024).
  33. Olive, K. P. *et al.* Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell* **119**, 847–860 (2004).
  34. Adams, J. R. *et al.* Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. *Cancer Res.* **71**, 2706–2717 (2011).
  35. Harada, N. *et al.* Intestinal polyposis in mice with a dominant stable mutation of the  $\beta$ -catenin gene. *EMBO J.* **18**, 5931–5942 (1999).
  36. Wintermantel, T. M., Mayer, A. K., Schütz, G. & Greiner, E. F. Targeting mammary epithelial cells using a bacterial artificial chromosome. *Genes. N. Y. N 2000* **33**, 125–30 (2002).

37. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 1–8 (2001).
38. Cable, D. M. *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* (2021) doi:10.1038/S41587-021-00830-W.
39. Bach, K. *et al.* Time-resolved single-cell analysis of Brca1 associated mammary tumourigenesis reveals aberrant differentiation of luminal progenitors. *Nat. Commun.* **12**, (2021).
40. Molyneux, G. *et al.* BRCA1 Basal-like Breast Cancers Originate from Luminal Epithelial Progenitors and Not from Basal Stem Cells. *Cell Stem Cell* **7**, 403–417 (2010).
41. Quinn, H. M. *et al.* YAP and  $\beta$ -Catenin cooperate to drive oncogenesis in basal breast cancer. *Cancer Res.* **81**, 2116–2127 (2021).
42. Mayer, S. *et al.* The tumor microenvironment shows a hierarchy of cell-cell interactions dominated by fibroblasts. *Nat. Commun.* **14**, 5810 (2023).
43. Coussy, F. *et al.* A large collection of integrated genomically characterized patient-derived xenografts highlighting the heterogeneity of triple-negative breast cancer. *Int. J. Cancer* **145**, 1902–1912 (2019).
44. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell–cell communication from single-cell transcriptomics. *Nat. Protoc.* **20**, 180–219 (2025).
45. Adler, M. *et al.* Principles of Cell Circuits for Tissue Repair and Fibrosis. *iScience* **23**, 100841 (2020).
46. Massagué, J. & Sheppard, D. TGF- $\beta$  signaling in health and disease. *Cell* **186**, 4007–4037 (2023).
47. Naba, A. Mechanisms of assembly and remodelling of the extracellular matrix. *Nat. Rev. Mol. Cell Biol.* **25**, 865–885 (2024).
48. Tucker, R. P. & Degen, M. Revisiting the Tenascins: Exploitable as Cancer Targets? *Front. Oncol.* **12**, 908247 (2022).
49. Lowy, C. & Oskarsson, T. Tenascin C in metastasis: A view from the invasive front. *Cell Adhes. Migr.* (2015) doi:10.1080/19336918.2015.1008331.

50. Oskarsson, T. *et al.* Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat. Med.* **17**, 867–874 (2011).
51. Tong, D. *et al.* LAMC2 promotes EGFR cell membrane localization and acts as a novel biomarker for tyrosine kinase inhibitors (TKIs) sensitivity in lung cancer. *Cancer Gene Ther.* **30**, 1498–1512 (2023).
52. Chen, X., Nadiarynkh, O., Plotnikov, S. & Campagnola, P. J. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* **7**, 654–669 (2012).
53. Arpinati, L. & Scherz-Shouval, R. From gatekeepers to providers: Regulation of immune functions by cancer-associated fibroblasts. *Trends Cancer* **9**, 421–443 (2023).
54. Croizer, H. *et al.* Deciphering the spatial landscape and plasticity of immunosuppressive fibroblasts in breast cancer. *Nat. Commun.* **15**, 2806 (2024).
55. Rosenbluth, J. M. *et al.* Organoid cultures from normal and cancer-prone human breast tissues preserve complex epithelial lineages. *Nat. Commun.* **11**, 1711 (2020).
56. Keymeulen, A. V. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193 (2011).
57. Valdés-Mora, F. *et al.* Single-cell transcriptomics reveals involution mimicry during the specification of the basal breast cancer subtype. *Cell Rep.* **35**, 108945 (2021).
58. Massagué, J. TGF $\beta$  in Cancer. *Cell* vol. 134 215–230 (2008).
59. Biffi, G. *et al.* IL1-Induced JAK/STAT Signaling Is Antagonized by TGF $\beta$  to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. *Cancer Discov.* **9**, 282–301 (2019).
60. Tsoumakidou, M. The advent of immune stimulating CAFs in cancer. *Nat. Rev. Cancer* **23**, 258–269 (2023).
61. Öhlund, D., Elyada, E. & Tuveson, D. Fibroblast heterogeneity in the cancer wound. *J. Exp. Med.* **211**, 1503–1523 (2014).
62. Hosaka, K., Yang, Y., Seki, T. *et al.* Pericyte–fibroblast transition promotes tumor growth and metastasis. *PNAS* **113**, E5618–E5627 (2016).
63. Miyoshi, K. *et al.* Activation of different Wnt/b-catenin signaling components in mammary

- epithelium induces transdifferentiation and the formation of pilar tumors. *Oncogene* **21**, 5548–5556 (2002).
64. Miyoshi, K. & Hennighausen, L.  $\beta$ -Catenin: a transforming actor on many stages. *Breast Cancer Res.* **5**, 63 (2003).
  65. Provenzano, P., Inman, D., Eliceiri, K. & al, et. Collagen density promotes mammary tumor initiation and progression. *BMC Med.* **4**, 38 (2006).
  66. Azzolin, L. *et al.* YAP/TAZ incorporation in the  $\beta$ -catenin destruction complex orchestrates the Wnt response. *Cell* **158**, 157–70 (2014).
  67. Zhang, T. *et al.* NUA1 promotes organ fibrosis via YAP and TGF- $\beta$ /SMAD signaling. *Sci. Transl. Med.* **14**, eaaz4028 (2022).
  68. Thinyakul, C. *et al.* Hippo pathway in cancer cells induces NCAM1+ $\alpha$ SMA+ fibroblasts to modulate tumor microenvironment. *Commun. Biol.* **7**, 1343 (2024).
  69. Monteran, L. & Erez, N. The Dark Side of Fibroblasts: Cancer-Associated Fibroblasts as Mediators of Immunosuppression in the Tumor Microenvironment. *Front. Immunol.* **10**, (2019).
  70. Ye, J. *et al.* Senescent CAFs Mediate Immunosuppression and Drive Breast Cancer Progression. *Cancer Discov.* **14**, 1302–1323 (2024).
  71. Shams, S. *et al.* The dual role of myoepithelial cells in breast cancer: antagonists and promoters. *Cancer Cell Int.* **22**, 157 (2022).
  72. Allinen, M. *et al.* Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **6**, 17–32 (2004).
  73. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
  74. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using scDbIFinder. *F1000Research* **10**, 979 (2022).
  75. Sztanka-Toth, T. R., Jens, M., Karaikos, N. & Rajewsky, N. Spacemake: processing and analysis of large-scale spatial transcriptomics data. *GigaScience* **11**, (2022).

76. Setty, M. *et al.* Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).
77. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
78. Mayer, S. *et al.* The tumor microenvironment shows a hierarchy of cell-cell interactions dominated by fibroblasts. *Nat. Commun.* **14**, 5810 (2023).
79. Hahsler, M., Piekenbrock, M. & Doran, D. **dbscan**: Fast Density-Based Clustering with *R. J. Stat. Softw.* **91**, (2019).

## Author contribution

K.L. generated and characterized the mouse model, derived and characterized *in vitro* lines derived from the model, supervised the injection experiment. I.T. performed snRNA-seq, spatial transcriptomics experiments, initial processing and computational analyses of the sequencing data. T.M.P. performed computational analyses. S.B. and F.H provided support for immunofluorescence experiments and organoid characterization. D.L.P., A.A. and N.K. provided support for computational analyses. M.M and A.X. provided support for the mouse model characterization, IFs, mouse colony handling and genotyping. H.R. and B.S. performed histology characterization. S.R. and A.M. performed imaging of IF and second harmonic generation imaging experiments. J.L., M.S. and A.B. provided technical support for snRNA-seq and spatial-transcriptomics experiments. S.D. performed spatial transcriptomics analyses under the supervision of M.N., S.K. performed and interpreted electron microscopy experiments. Eli.M., Elo. M., generated the xenograft sample cohort. W.B. provided mouse model supervision. K.L., W.B. and N.R. provided funding acquisition. K.L., I.T., T.M.P., and N.R. conceptualized and designed the project. N.R. supervised the project. K.L., I.T., T.M.P. and N.R. wrote the manuscript with input from all authors.

## Acknowledgements

K.L. was financially supported by the Deutsche Krebshilfe (DKH), grant number: 70113738. I.T. was the recipient of an EMBO long term post-doctoral fellowship. T.M.P. is financially supported by the Berlin School of Integrative Oncology through the GSSP program of the German Academy of Exchange Service (DAAD) and by the Add-on Fellowship of the Joachim Herz Foundation. N.R. thanks the Deutsche Forschungsgemeinschaft (DFG) grant number RA 838/5-1, Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) grant numbers 81Z0100105 and 81X2100155, the Chan Zuckerberg Foundation (CZI)/Seed Network, and the NeuroCure/Cluster of Excellenz in the neurosciences at the Charité – Universitätsmedizin Berlin. We thank Jeannine Wilder, Tatiana Borodina and the genomics facility of MDC for sequencing the samples. We thank Caroline Braeuning and the BIMS FACS facility for cell sorting. We thank Diana Behrens and the EPO GmbH team for

coordinating and performing the injection experiment. We thank Giuseppe Macino, Elena Splendiani and Jan Licha for helping with Open-ST protocol optimization during their stay in the Rajewsky lab. We thank Takako Sasaki for sharing the laminin gamma (Lamc2) antibody. We thank Stefan Florian for insightful discussions and expertise during initial assessments of samples histology. The authors thank the whole Rajewsky lab for discussions and constructive feedback on the manuscript.

## Resource availability

### Lead contact

Further information and requests for resources and reagents may be directed to Nikolaus Rajewsky (rajewsky[at]mdc-berlin.de).

### Data and code availability

Open-ST RNA-seq and microscopy raw and processed data have been deposited at GEO and will be publicly available as of the date of peer-reviewed publication.

All original code will be publicly available as of the date of publication.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## Methods details

### Mouse Strains

All animal experiments were conducted in accordance with European, national, and MDC regulations. WAP-iCre (B6129-Tg(Wap-cre)11738Mam/J Stock No: #:003552), Ctnnb1 ex3 (Ctnnb1tm1Mmt), YFP (B6.129X1-Gt(ROSA) 26Sortm1(EYFP)Cos/J, Stock No: 006148) mice were crossed into FVB/N background and previously described<sup>41</sup>. 129S-Trp53<sup>tm2Tyj/J</sup> (Stock No: 008652) and FVB.129S6-Gt(ROSA)26Sor<sup>tm1</sup> (Pik3ca<sup>H1047R</sup>)Egan/J (Stock No: 016977) mice were purchased from Jackson Laboratories and crossed into FVB/N mouse background for more than 10 generations. Animal experiments were approved by the ethical board Landesamt für Gesundheit und Soziales (LaGeSo), Berlin (Approval G#0213.18). Tumor dynamics analysis is described in the result section of the manuscript, in short: tumor mutant mice developed palpable tumors from the 8th week onward. This occurred without additional stimulation of the promoter with pregnancy, induction of the promoter occurred around the 5th week of life of mice. Tumors were developed selectively in all 10 mammary glands of each mouse carrying the oncogenic genotype (GOF mutation in Trp53, Ctnnb1, Pik3ca and YFP with promoter WAPiCre), no other organs were affected with tumor formation, no other phenotype apart from mammary gland tumors growth was observed. The tumors were harvested between 6 and 11 weeks of mouse age and reached a maximum size of 1 cm<sup>3</sup>. For genotyping, ear biopsies were digested in lysis buffer (100 mM Tris, pH 8.0–8.5, 10 mM EDTA, pH 8.0, 0.2% SDS, and 200 mM NaCl) with 1:30 10 mg/ml Proteinase K (Roche, Cat. #03 115 879 001). Lysates were then diluted 1:20 in nuclease-free water, and PCR was carried out using Taq DNA polymerase (Invitrogen). The following primers were used for genotyping:

### List of Genotyping primers

| <i>Target/Primer Name</i> | <i>Sequence:</i> |
|---------------------------|------------------|
|---------------------------|------------------|

|                     |   |                                  |
|---------------------|---|----------------------------------|
| LSL in p53 wildtype | - | TTACACATCCAGCCTCTGTGG            |
| LSL in p53 mutant   | - | AGCTAGCCACCATGGCTTGAGTAA GTCTGCA |
| LSL in p53 Rev      | - | CTTGGAGACATAGCCACACTG            |
| PI3KCA CDS 562 FW   | - | CAC TTG GGG TGA AGG GAC TC       |
| PI3KCA BGH Rev      | - | TAG AAG GCA CAG TCG AGG          |
| WAPiCre Fw          | - | GAA AAG CAC CAG GAG AAG TCA C    |
| WAPiCre Rev         | - | GAC ACA GCA TTG GAG TCA GAA G    |
| YFP Fw              | - | TGCAGTGCTTCGCCCCGCTACC           |
| YFP Rev             | - | CCGTCGCCGATGGGGGTGTTC            |
| Bcat 482 Fw         | - | AGAATCACGGTGACCTGGGTTAAA         |
| Bcat 1051 Rev       | - | CATTCATAAAGGACTTGGGAGGTGT        |

## Histology and Immunostaining

Mammary glands were fixed overnight at 4°C in 4% formaldehyde (Roth), then dehydrated and embedded in paraffin. For staining, tissue sections were deparaffinized, rehydrated, stained with Hematoxylin and Eosin (H&E), dehydrated, and mounted. Images were acquired with a bright-field Zeiss microscope. For immunostaining, antigen retrieval was performed on paraffin-embedded sections by boiling in Tris-EDTA buffer (10 mmol/L Tris, 1 mmol/L EDTA, pH 9.0) for 20 minutes. Cryosections were directly frozen in OCT after tissue was excised from the mice. Before the staining sections were fixed with 4% methanol-free paraformaldehyde for 15 minutes.

All samples were blocked with 10% horse serum for 1 hour at room temperature before overnight incubation with primary antibodies at 4°C (see STable5 for antibody list). For fluorescent staining, samples were incubated with secondary antibodies and DAPI for 1 hour at room temperature before mounting with Immu-Mount (Thermo Scientific).

For immunohistochemistry, after antigen retrieval, slides were washed, incubated with 3% H<sub>2</sub>O<sub>2</sub> for 15 minutes at room temperature, washed three times for 5 minutes in PBS, and incubated with primary antibodies overnight at 4°C. Horseradish peroxidase-labeled secondary antibodies were diluted and applied on slides for 1 hour at room temperature. Slides were then washed, developed using DAB chromogenic substrate (DAKO), dehydrated, and mounted.

## Histology and Immunohistochemistry (IHC)

Brightfield imaging of hematoxylin and eosin (H&E) stained paraffin sections and chromogenic immunohistochemistry (IHC) was performed using the 3DHISTECH Panoramic SCAN II slide scanner (20× Plan-Apochromat objective, NA 0.8). Full mammary gland sections were scanned at 20× magnification under brightfield illumination. Stitching and image rendering were performed using CaseViewer software. For IHC, diaminobenzidine (DAB) chromogen was used, and scanned slides were evaluated uniformly with matched exposure settings. All scans included full-gland imaging for downstream analysis of lesion heterogeneity and tumor progression stages.

## Immunofluorescence imaging

Fluorescently stained cryosections were imaged using both widefield and confocal platforms. Widefield fluorescence imaging was conducted on a Leica Thunder Imager 3D Tissue system using a DFC9000 GTC sCMOS camera and a 20× HC PL APO dry objective (NA 0.75). Images were acquired at 20× magnification and stitched using Leica LAS X software when required for whole-gland visualization.

High-resolution confocal imaging was carried out using a Leica Stellaris 5 system equipped with a White Light Laser and HyD detectors. Sections were scanned using a 20× HC PL APO CS2 glycerol objective (NA 0.75) with 1024×1024 pixel resolution and sequential laser scanning. Z-stacks were collected when necessary for visualization of spatial organization within tumor-stromal niches. All image acquisitions were performed using standardized laser settings and exposure times to allow quantitative comparisons across conditions.

### **Isolation of mammary gland cells**

Tumors were minced and digested in DMEM/F12 HAM (Invitrogen) supplemented with 5% FBS (Invitrogen), 5 mg/mL insulin (Sigma-Aldrich), 0.5 mg/mL hydrocortisone (Sigma-Aldrich), and 10 ng/mL EGF (Sigma-Aldrich), forming a digestion medium containing 300 U/mL collagenase type III (Worthington), 100 U/mL hyaluronidase (Worthington), and 20 mg/mL Liberase (Sigma). Digestion was carried out at 37°C for 1.5 hours with shaking.

The resulting organoids were resuspended in 0.25% trypsin-EDTA (Invitrogen) at 37°C for 1 minute, then dissociated in digestion medium containing 2 mg/mL Dispase (Invitrogen) and 0.1 mg/mL DNase I (Worthington) at 37°C for 45 minutes while shaking. Samples were filtered using 40 µm cell strainers (BD Biosciences) and incubated with 0.8% NH<sub>4</sub>Cl solution on ice for 3 minutes for red blood cell lysis. Lysis was stopped by washing in 30 mL Dulbecco's Phosphate-Buffered Saline (DPBS) without Ca<sup>2+</sup> and Mg<sup>2+</sup> (Gibco, catalog no. 14190169). The resulting pellets were used for downstream applications as outlined below.

### **Organotypic Stem Cell-Enriched 3D Cultures**

Organotypic stem cell-enriched cultures were generated using previously published protocols, with minor modifications. Single cells from digested mammary glands were resuspended and plated on collagen I-coated plates (50µg/mL) in stem cell-enriching medium MEBM (Lonza Cat. #CC-3151), supplemented with 2% B27 (Invitrogen, Cat. #17504044), 20 ng/mL bFGF (Invitrogen, Cat. #13256029), 20 ng/mL EGF (Sigma, Cat. #SRP3196-500 µg), 4µg/mL heparin (Sigma, Cat. #H3149), 5µg/mL insulin (Sigma, Cat. #I0516-5ml), 0.5µg/mL hydrocortisone (Sigma, Cat. #H0888-1G), and 1X Gentamicin (Sigma, Cat. #G1397-100ML) for 16–18 hours.

Cells were washed twice with DPBS, treated with 0.25% trypsin-EDTA for 30 seconds, and then incubated with TrypLE (Thermo Fisher#12604021) at RT for 5–7 minutes until detachment. Cells were resuspended in stem cell-enriching medium, counted, and seeded in 25µL droplets containing 95% reduced growth factor Matrigel at a density of 100 cells/µL. Plates were carefully inverted, and Matrigel was allowed to solidify at 37°C for 45 minutes to 1 hour.

Then, 0.5 mL of stem cell medium was added per well of 24-well plates containing a 25 µL droplet. Medium was changed every second day. For secondary sphere formation, spheres

were dissociated for 10 minutes in TrypLE, and cells were filtered using 0.45µm filters. Single-cell suspensions were re-seeded as described for primary cells. All treatments with verteporfin were performed under minimal lighting. Tumor cell were enriched and isolated by sorting them for the presence of YFP using Fluorescence -activated cell sorting (FACS)

Fibroblast cultures were isolated the same way as epithelial compartment but were grown in DMEM, FBS 10%, P/S 1% on uncoated plastic dishes. After plating cells were detached, FACS sorted for the absence of CD45 to exclude all of possible contamination with immune cells and YFP to exclude the chance of tumor cells growth.

### **Fluorescence -activated cell sorting (FACS )**

Single-cell suspensions from dissociated tumors were resuspended at 10,000 cells/µl and incubated with conjugated primary antibodies at 4°C for 15 minutes. Cells were then washed three times in DPBS and subsequently incubated with 7-AAD (5 µl per 10<sup>6</sup> cells) at room temperature for 5 minutes to stain dead/dying cells. Cells were sorted using the FACS Aria II or III (BD Biosciences) or analyzed using the LSRFortessa (BD Biosciences). Compensation and unstained controls were carried out before FACS experiment. Data is shown for three independent biological experiments from cell lines derived from three different animals, that were analyzed using FlowJo Analysis Software.

### **Patient-derived xenograft models**

Patient-derived xenograft (PDX) models were established from patients with triple-negative breast cancer with their informed consent as described previously (PDX paper of Marangoni). The experimental protocol and animal housing were in accordance with institutional guidelines as proposed by the French Ethics Committee (agreement no. B75-05-18). PDXs were obtained from patients with their informed written consent.

### **Cell grafting**

Engraftment experiments were conducted at EPO GmbH, Berlin-Buch, and performed according to the German Animal Protection Law with approval from the responsible authorities (LAGeSo, approval number E0023-23). The in vivo procedures were consistent and in compliance with the UKCCCR guidelines.

Ncam1 was used as a marker of Cancer Associated Fibroblasts (myCAFs) to sort these cells by FACSs (FAB7820P, R&D Systems). Ncam1<sup>high</sup> population of CAFs were derived from transformed glands of T2 from tumor lines (P3WeY). Ncam1<sup>low</sup> Normal Fibroblasts (NFs) were derived from healthy ctrl mammary glands (WY genotype). Cells were screened for the absence of Immune cell marker Ptprc/Cd45 (25-0451-81, Invitrogen) and lack of YFP expression (it would indicate the contamination of fibroblasts with tumor cells).

Accordingly tumor cells were sorted for the presence of YFP as explained method section: mammary gland cells isolation.

2 biological replicates for each of the lines, organoids, Normal fibroblasts and CAFs were used for the injection experiments. First tumor organoids were dissociated to single cells in Tryple for 30min in 37°C and filtered through Flow me 40um mesh (FlowMi)Cell strainer,

then counted in Trypan blue. Accordingly, Fibroblasts cultures were detached by Trypsin and counted in Trypan blue.

2 x 10<sup>6</sup> tumor cells were injected into fat pads into immunocompetent, 7 week old, healthy mice either: alone (Group A), with 2 x 10<sup>6</sup> of normal fibroblasts (NF) (Group B) or with 2 x 10<sup>6</sup> of CAFs (Group C). Every injection was supplemented by equal volume of BME (Cultrex UltiMatrix RGF BME, R&D systems). Every injection consisted of cells resuspended in 20 ul of Medium that was additionally supplemented with 20ul of BME. Addition of BME aimed to help the cells to recover after dissociation procedure.

Each of 5 mice which we used per group was injected into two mammary fat pads making 2 experimental points per mouse. After injection mice were routinely (every 2 days) monitored for palpable tumor appearance and weighted. After the largest tumors reached the size of 1,5cm all mice were sacrificed at the same times. Glands were collected either for paraffin or direct freezing in OCT.

### **Electron microscopy**

Mammary gland tissue, healthy and tumor as well as organoids were processed for ultrastructural analysis. All sample types were fixed by immersion in 2% (w/v) formaldehyde and 2.5% (v/v) glutaraldehyde in 0.1 M phosphate buffer for 2 hrs at room temperature followed by an overnight incubation at 4°C. Samples were post-fixed with 1% (v/v) osmium tetroxide (Sigma-Aldrich), dehydrated in a graded series of EtOH and embedded in PolyBed 812 resin (Polysciences). Ultrathin sections (60–80 nm) were poststained with 0.5% (w/v) uranyl acetate (Sigma-Aldrich) and 3% lead citrate (Leica microsystems). Sections were imaged using back-scattered electrons at 2 kV, 0.8nA with scanning electron microscope Helios 5CX (Thermo Fisher Scientific). Acquisition was performed with the retractable DBS detector in immersion mode and the MAPS software (Thermo Fisher Scientific).

### **snRNAseq data generation**

Healthy or tumor pieces from the abdominal mammary gland were snap-frozen in liquid nitrogen. For processing, half a gland was used for healthy and early tumors and only a piece was used for middle and late tumors. Frozen tumor pieces were kept on ice and chopped into pieces with a cold scalpel in 500 µl of lysis buffer (10 mM Tris-HCl PH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% NP-40, 1 mM DTT, 1u/µl Rnase Inhibitor). The tissue was further collected and processed on ice in 1.5 mL tubes for the rest of the procedure. First, 800 µl of lysis buffer was added, the tissue was crushed with a plastic pestle 10 to 15 times and incubated for 5 minutes with gentle resuspension from time to time using a 1000 µl pipette. Tissue pieces were then collected at the bottom of the tube with a fast spin down, the supernatant was strained in a new tube using a 100 µl filter (pluriselect) and spun down at 500G in a swing bucket rotor for 5 min at 4°C. The supernatant was removed and 1 ml of wash buffer (1% BSA, 1u/µl Rnase Inhibitor in BSA) was gently added without disturbing the nuclei pellet and incubated for 5 minutes. Nuclei were spun down at 500 G in a swing bucket rotor for 5 min at 4°C. An additional wash was performed in 1 ml of wash buffer and after a final centrifugation, nuclei were resuspended in 500 µl of wash buffer and filtered through a 40 µm strainer (pluriselect). DAPI was then added to a final concentration of 2 µM. In the case of T2 and T3 tumor samples that presented a higher concentration of debris, around 250,000 nuclei were sorted based on DAPI using a BD FACS Aria III with a 100 µm nozzle in

250 µl of sorting buffer (4% BSA, 4u/µl Rnase Inhibitor in BSA). Nuclei were then counted using a Neubauer counting chamber with a Keyence microscope and processed using the 10X Genomics 3' V3.1 kit following the manufacturer instruction with a target of 8,000 to 9,000 recovered nuclei. The only adjustment concerned the number of PCR cycles which was increased by one accounting for the usual lower concentration of RNA in nuclei in comparison to cells. Libraries were sequenced with the recommended read lengths.

### snRNAseq data processing

Fastq files were generated using Cellranger v7.1.0 ('Cellranger mkfastq') and aligned to the mm10 mouse genome ('Cellranger count') to generate digital gene expression matrices with default parameters. Filtered gene expression matrices were imported in R v4.4.3 for downstream processing using Seurat v5.1.0<sup>73</sup>. Single samples were merged into a single Seurat object and cells with less than 1000 genes and more than 50% of mitochondrial transcripts were removed. Doublets were identified using *scDbIFinder* v1.18.0<sup>74</sup> and cells with a doublet score higher than 0.5 were removed. Raw gene expression counts were log-normalized and scaled.

### Open-ST data generation

Healthy and tumorigenic mammary tissue were embedded into OCT before being frozen and kept at -80°C. 12 µm sections were cut from blocks in a cryostat and placed on flow-cell pieces prepared according to the Open-ST protocol. Samples were processed according to the Open-ST protocol<sup>32</sup>. Briefly, tissue sections were fixed for 30 min in pre-chilled methanol at -20°C and stained with hematoxylin and eosin (H&E). Images of the stained sections were acquired at 10X or 20X magnification on a Keyence microscope. The tissue on the flow cell was incubated with permeabilization solution for 30 min at 37°C then washed in RT buffer (SSIV 1X RT buffer, Rnase inhibitor 1U/µl) before being incubated overnight at 42°C in RT (SSIV 1X RT buffer, 5 mM DTT, BSA 0.187 mg/ml, 1 mM dNTP mix, 6.67 U/µl, Ribolock Rnase inhibitor 1U/µl). The next day exonuclease treatment (1X Exo1 buffer, Exo1 1U/µl) was performed for 45 min at 37°C followed by tissue removal (100 mM Tris-HCl PH 8.0, 200 mM NaCl, 2% SDS, 5 mM EDTA, Proteinase K 16 mU/µl) for 40 min at 37°C. Tissues were washed 3 times with ultrapure water, 3 times with 0.1 M NaOH (with 5 min incubation each time), 3 times with 0.1 M Tris PH 7.5 and 3 times with ultrapure water. Second strand synthesis (1X NEBuffer-2, 10 µM random primers, 1 mM dNTPs, Klenow exo (-) fragment enzyme 0.5 U/µl) was then performed for 2 h at 37°C. The flow-cell pieces were washed 3 times with ultrapure water, and the second strand product was eluted in 2 rounds of 100 µl 0.1M NaOH. The final 200 µl of second strand product was mixed with 28.6 µl of 1M Tris-HCl pH7.5. The product was purified using Ampure beads at a 1.8x ratio according to the manufacturer instructions and eluted in 82.5 µl of ultrapure water. 2.5 µl of purified product was used in a qPCR (1x blue S-Green qPCR mix + ROX, 1 µM forward primer, 1 µM reverse primer) reaction to determine the concentration. The threshold of 50% of the peak deltaRN was used to determine the number of cycles used in the subsequent PCR reaction using unique index barcodes for each sample. This number of cycles minus 5 was used in the PCR amplification of the cDNA where a 200 µl reaction was split into 4 50 µl tubes. The 200 µl PCR product was purified using Ampure XP beads at a 1X ratio according to the manufacturer instructions and eluted in 20 µl of ultrapure water. The size and concentration of the cDNA was checked on a BioAnalyzer or a Tapestation. The final size selection of

fragments between 350 and 1100 bp was performed using a BluePippin or Pippin HT instrument according to the manufacturer instructions. The final library sizes were assessed using Tapestation or Bio Analyzer and the concentration was measured using Qubit. If a small contamination peak was still observed below 300bp, an additional AmPure bead purification was performed at a ratio of 1X.

### **Open-ST data processing**

Raw BCL files were demultiplexed using BCL2FASTQ and mapped to the mm10 mouse genome using spacemake V0.7.9<sup>75</sup> with the default parameters of the Open-ST barcode flavor. For downstream analyses we used the 7  $\mu$ m hexagonal mesh. When not done automatically during the spacemake pipeline, stitching of the images and data matrix were done using the stitch.py function of the Open-ST package. Samples were filtered for unwanted tiles and saved as a 10X genomics matrix folder format to be converted to Seurat object in R. Samples were filtered according to sample-specific threshold accounting for sequencing depth (T0\_1:400, T0\_2:100, T0\_3:100, T1\_1:250, T1\_2:100, T1\_3:100, T2\_1:150, T2\_2:200, T2\_3:200, T3\_1:100, T3\_2:150, T3\_3:200 genes). Raw gene expression counts were log-normalized and scaled (SCTransform) using Seurat.

### **snRNAseq cell type identification**

The top 3'000 variable genes were used to compute 30 principal components and UMAP coordinates, while a subset of 2964 genes with standardized variance higher than 1 in both snRNAseq and Open-ST data was used to identify clusters (resolution= 1.5, dims=1:30). Cluster annotation were supported by marker gene expression (FindMarkers only.pos=T, min.pct= 0.2, min.diff.pct=0.1, max.cells.per.ident=500) and label transfer scores from a murine healthy and tumor mammary gland atlas<sup>39</sup>. For label transfer, the atlas was subsetting to 1000 cells per identity ('CellTypesFinal') before label transfer (FindTransferAnchors normalization.method='LogNormalize', reduction='pcaproject', dims=1:30, non.method='rann', eps=0.5).

### **snRNAseq epithelial state identification**

Epithelial cells (including 'Hormone sensing', 'Luminal', 'Myoepithelial' and 'Tumor') were integrated (IntegrateLayers method= CCAIntegration) and re-clustered. Doublets identified by means of marker gene expression ('Ptprc', 'Mrc1' and 'Meg3') and higher scDbtFinder scores were excluded. UMAP coordinates (reduction= 'epithelial.cca', dims=1:30) and clusters were recomputed (resolution=0.4). Cluster annotations were supported by marker gene expression (FindMarkers only.pos=T, min.pct= 0.2, min.diff.pct=0.1, max.cells.per.ident=500), label transfer scores and enrichment for control and tumor-bearing samples. UMAP 2D densities were computed over a regular grid using MASS v7.3-65 (kde2d n=100) for each condition after balancing the number of epithelial cells from both conditions (slice\_sample n=200) and then subtracted. Stromal densities were computed in a similar way, modified to account for the higher number of stromal cells (slice\_sample n=4000, kde2d n=1000).

### **snRNAseq tumor pseudotime ordering**

Tumor pseudotime was computed using Palantir v1.3.2.<sup>76</sup> Healthy (i.e. ‘Hormone sensing’, ‘Luminal’ and ‘Myoepithelial’) and proliferating populations were excluded. 10 components were used as diffusion map input (run\_diffusion\_maps n\_components=10, determine\_multiscale\_space). Pseudotime was computed starting from a Wap-expressing ‘Cell-of-origin’ cell from a T1 sample (run\_palantir start\_cell=‘T1\_Rep2-GTTACCCGTGTTTCGTA’, num\_waypoints=500) .

### **Gene set scoring in single nuclei and spatial data**

Gene set scores in individual cells were computed using AUCell v1.26.0<sup>77</sup>. For each cells in either single nuclei or spatial data, genes were ranked by their raw expression (AUCell\_buildRankings) and the gene set area under the curve score was computed (AUCell\_calcAUC). Scores were then added as a new assay and scaled (ScaleData). Stromal signatures were obtained from Mayer et. al., 2023<sup>78</sup> while gene sets were selected from the HALLMARK, REACTOME and KEGG databases using msigdb v7.5.1.

### **Open-ST cell type identification**

The top 2’000 variable genes were used to compute 10 principal components and identify clusters (resolution=0.8). Cluster annotations were supported by marker gene expression.

### **Open-ST Robust Cell Type Deconvolution**

Robust Cell Type Deconvolution<sup>38</sup> scores of snRNAseq cell states were computed using spacexr v2.2.1. To simultaneously score all samples in a group (run.RCTD), merged primary and EPO seurat objects were generated from 12 and 3 samples, respectively. Proliferating states (‘Immune proliferating’, ‘Fibroblast proliferating’, ‘Tumor proliferating’) and epithelial doublets were excluded from the single cell reference and a maximum of 500 cells per category was kept (Reference n\_max\_cells=500). Cells identified as ‘Nerve and ‘Muscle’ were excluded from the spatial data.

### **Open-ST niche identification**

To identify multicellular niches, we first identified the 30 closest spatial neighbours for each cell using dbSCAN v1.2.0<sup>79</sup> on the stitched x,y coordinates (kNN, k=30). Neighbours further than 50 µm were excluded (pixel-to-micron scaling factor=1/30) and the neighbourhood matrix was computed by averaging RCTD scores per cellular neighbourhood. Dimensionality reduction and clustering were then performed on the scaled neighbourhood matrix as described for snRNAseq (ScaleData, RunPCA, RunUMAP dims=1:6, FindNeighbours dim=1:6, FindClusters resolution=0.4). Niches were annotated analyzing the average RCTD scores per cluster and their spatial localization.

### **Open-ST pseudotime scoring**

Seurat integration pipeline was used to map tumor pseudotime scores in space. To avoid missing values, tumor pseudotime scores were set to 0 for non-tumor cells in the single cell reference.

To simultaneously score all samples in a group, merged primary and EPO seurat objects were generated from 12 and 3 samples, respectively. Gene expression data was subset to

shared 2,964 variable genes, normalized (SCTransform) and, following feature selection (SelectIntegrationFeatures), transfer anchors were identified between single nuclei and merged spatial objects (FindTransferAnchors, normalization.method = "SCT", nn.method = 'rann', eps = 0.5) to transfer pseudotime scores (TransferData prediction.assay = TRUE, weight.reduction = 'pcaproject', dims = 1:30, eps = 0.5).

### **Open-ST duct segmentation**

Duct segmentation was performed via manual lasso selection using plotly v4.10.4. Cells with a cumulative RCTD epithelial score higher than 0.5 were plotted interactively using shiny v1.9.1 to identify the barcodes of cells belonging to individual ducts. The duct-stromal interface was then defined as the cellular neighbours of duct cells as described for Open-ST niche identification.

### **Open-ST duct pseudotime staging**

Duct average tumor pseudotime scores computed across ductal cells were used to order individual ducts along tumor progression. Five stages were identified by binning average pseudotime scores in 0.15 intervals. RCTD scores, gene expression, ligand activity are then averaged per duct stage and scaled for visualization purposes.

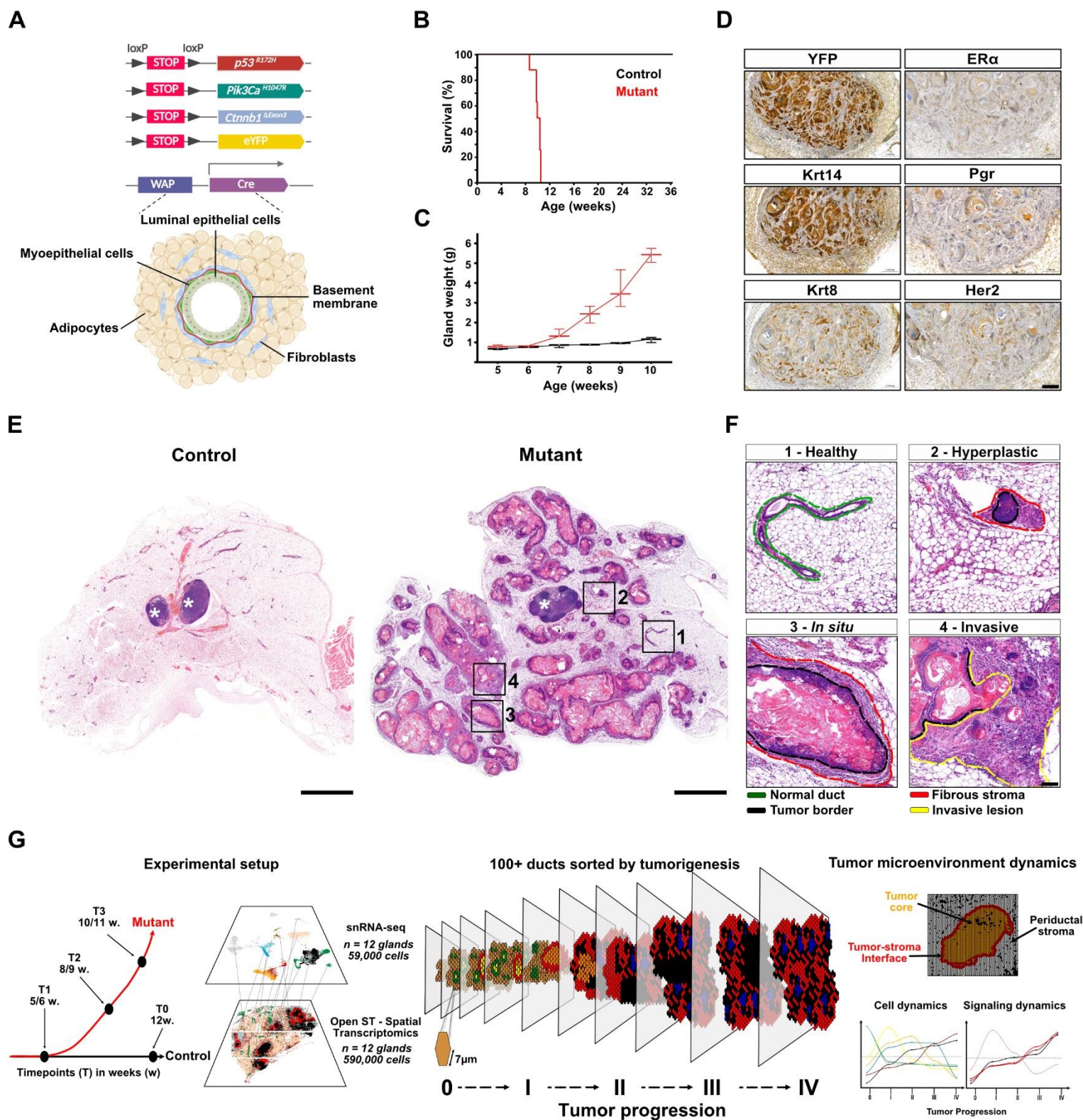
### **Open-ST analysis of stromal state distance to duct-stromal interface**

For each stromal cell (i.e. with RCTD scores for macrophage and fibroblast states higher than 0.5), the closest duct-stromal interface was identified and their distance computed. To normalize for different duct radii and shapes, the distance to the duct interface was divided by the distance between the duct interface cells and the duct center of mass (i.e. the average x,y coordinates of ductal cells). Distances were then log-normalized so that positive scores identify extra-ductal stromal cells and negative values intraductal ones. Stromal cells further than 300  $\mu\text{m}$  to any duct center were excluded.

### **Open-ST cell-cell communication analysis**

To investigate cell-cell communication in space, known mouse ligand-receptor pairs were obtained from CellChat v2.1.2<sup>44</sup>. For each ligand-receptor pair, an interaction score was computed in each cell as the geometric mean of receptor expression and average ligand expression in the 30 closest neighbors. Ligand activity scores were then computed by summing interaction scores of receptor-ligand pairs sharing the same ligand.

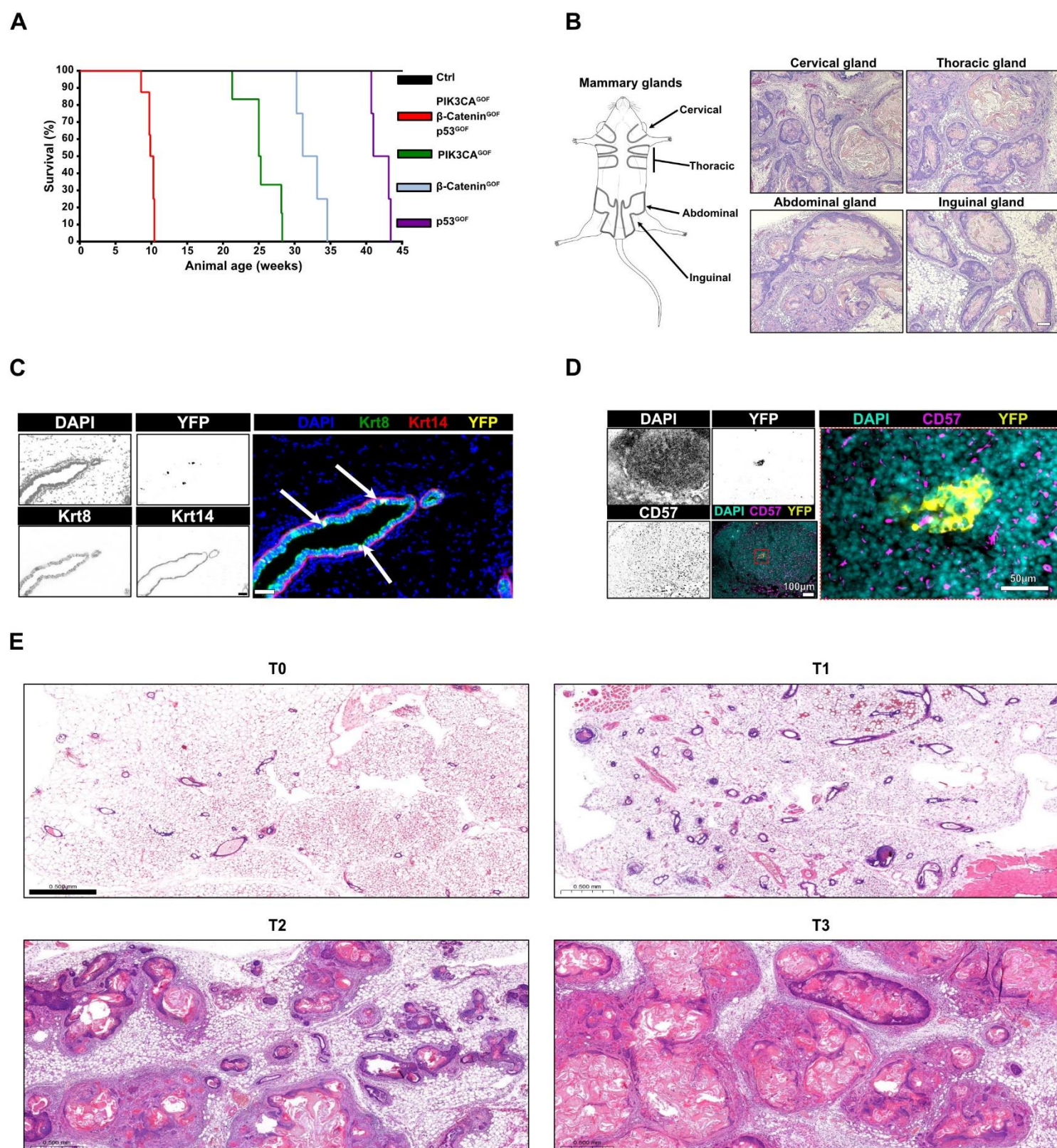
**Figure 1**



# Figure 1. Multiple asynchronous tumors arise in a triple-negative breast cancer (TNBC) model

- (A) The TNBC mouse model. Using a WAP-iCre system, cells in the luminal epithelium of the mammary glands can recombine and express three oncogenic drivers (mutant *Trp53*<sup>R172H</sup>, *Pik3ca*<sup>H1047R</sup>, and a stabilized form of  $\beta$ -catenin (*Ctnnb1* <sup>$\Delta$ ex3</sup>). Each of the pathways are frequently activated, sometimes jointly, in human patients. Recombined cells are tagged with YFP.
- (B) Tumors develop consistently across all mutant animals. Kaplan-Meier survival curves comparing mutant and control animals (all oncogenic transgenes but no WAP-iCre system). All tumor animals were euthanized due to tumor burden. *N* = 8 per group.
- (C) Tumor growth is detectable starting after 6 postnatal weeks. Combined weight of thoracic and abdominal mammary glands in mutant and control (WAP-iCre and YFP but no oncogenic transgenes) animals. *N* = 3 per timepoint.
- (D) Tumors are triple-negative and basal-like at study endpoint. Immunohistochemistry staining of mammary tumors at 9 postnatal weeks. Random example (*N* = 3). Scale bars: 200  $\mu$ m.
- (E) Multiple tumors arise in individual glands. Hematoxylin and eosin (H&E) staining of abdominal mammary glands from mutant and control animals at postnatal week 8. White stars: intramammary lymph nodes. Random example (*N* = 3). Scale bars: 1 cm.
- (F) Ducts in individual glands co-exist at different stages of transformation. Higher magnification H&E images from the mutant glands shown in (E). Scale bars: 100  $\mu$ m.
- (G) Framework to study tumor progression and TME remodeling: 1) Spatial Transcriptomics and snRNA-seq were performed at four stages (T0–T3) spanning healthy to advanced tumors. 2) Ducts were computationally segmented to resolve events at the duct–stromal interface. 3) 100+ ducts were computationally ordered along tumor progression. 4) Cellular, ligand–receptor, and ECM dynamics were analyzed across stages of tumor evolution.

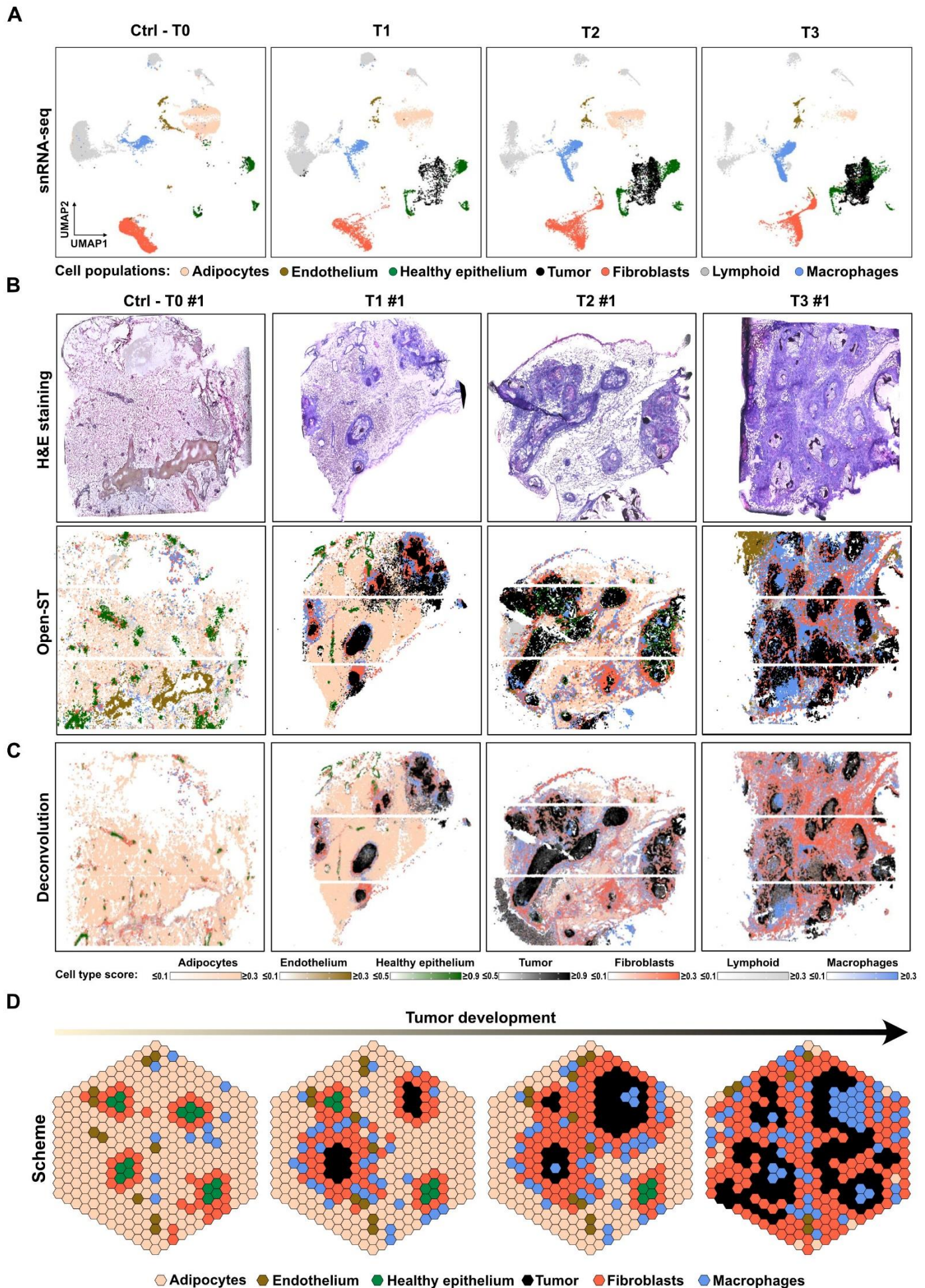
**SFigure 1**



## SFigure 1

- (A) Tumor development is delayed in single-mutant animals. Kaplan–Meier survival curves comparing tumor onset in animals carrying individual oncogenic alleles to triple-mutant and healthy control animals (N=8).
- (B) Tumors arise across all mammary glands. Left: murine mammary gland anatomy. Right: H&E staining of cervical, thoracic, abdominal and inguinal glands from a mutant animal at postnatal week 8. Random example (N=3). Scale bar: 100  $\mu$ m.
- (C) Oncogene activation (YFP<sup>+</sup>) is restricted to the luminal compartment (Krt8<sup>+</sup>/Krt14<sup>+</sup>) in early-stage glands. Single channel (left) and merged (right) immunofluorescence (IF) images of mammary glands from a mutant animal at postnatal week 5. Arrows: YFP<sup>+</sup>/Krt8<sup>+</sup> cells. Random example (N = 3). Scale bars: 50  $\mu$ m.
- (D) Tumor cells (YFP<sup>+</sup>) are detected in intramammary lymph nodes at study endpoint. Single-channel and merged overview (left) and merged zoom-in (right) IF images of lymph node tissue from mutant animals at 10 postnatal weeks. Random example (N = 3). Scale bars: 100  $\mu$ m (overview), 50  $\mu$ m (zoom-in).
- (E) Histology across study timepoints. H&E staining of mammary glands collected at T0 (healthy), T1 (5–6 weeks), T2 (8–9 weeks), and T3 (10–11 weeks). Scale bars: 500  $\mu$ m. Random example (N = 10).

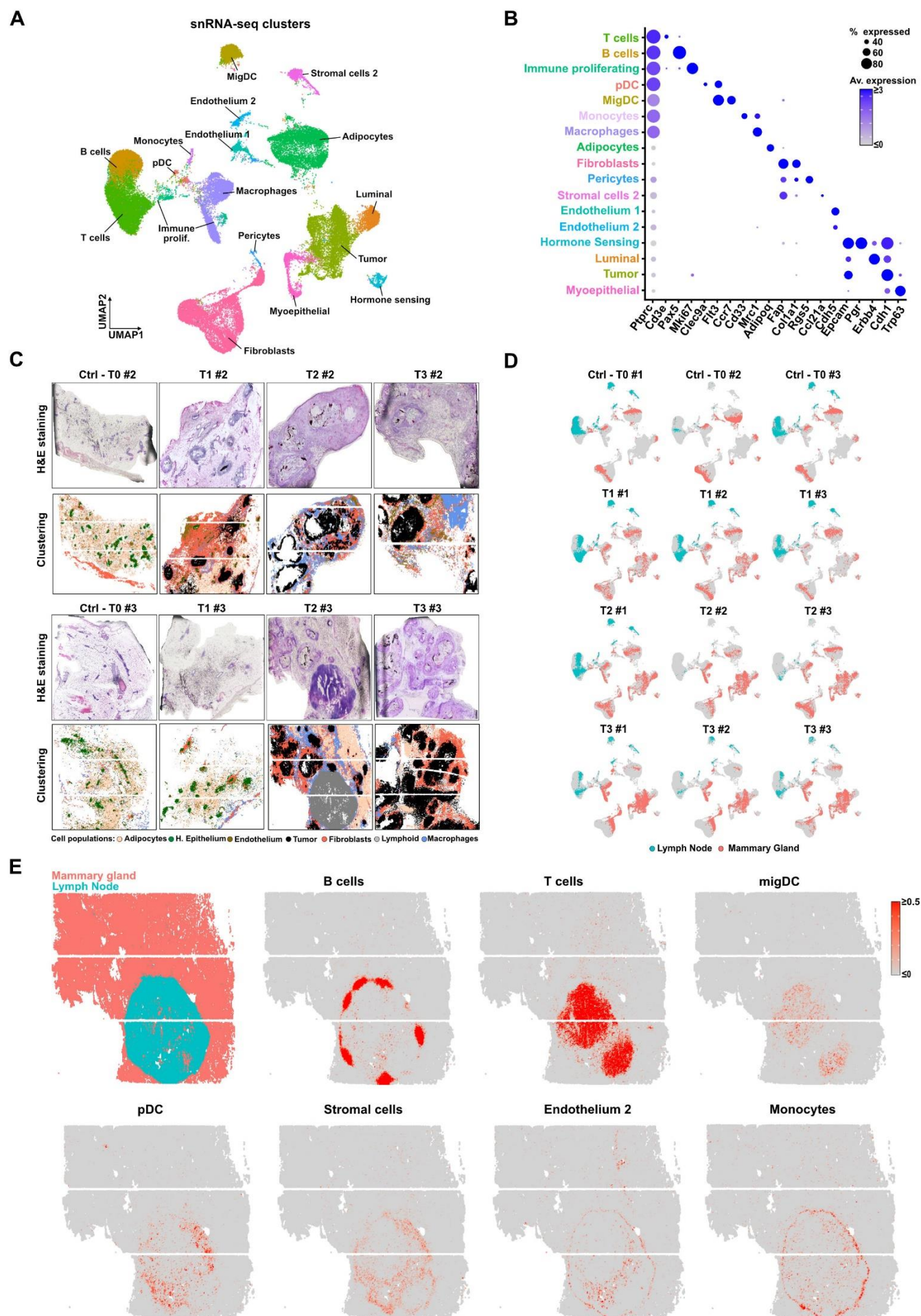
**Figure 2**



## Figure 2. Spatiotemporal quantification of tumorigenesis at single cell resolution

- (A) Epithelial and stromal cell types are remodeled in the tumor microenvironment (TME). Uniform manifold approximation and projection (UMAP) of single-nucleus RNA-seq data isolated from mammary glands along tumor progression (Timepoints T0–T3). RNA from  $N = 58,742$  nuclei was sequenced (12 samples, 3 biological replicates/timepoint, Methods). Cell types in SFig2A.
- (B) Spatial transcriptomics captures the progressive remodeling of TME molecular architecture. Top: H&E images of tissue slices from glands acquired for spatial transcriptomics. Bottom: Clustering of the genome-wide, high-resolution spatial transcriptomics data (Open-ST). RNA from  $N = 587,288$  pseudocells was sequenced (12 samples, 3 biological replicates/timepoint, Methods). Additional replicates in SFig2C. Scale bars: 500  $\mu\text{m}$ .
- (C) Single-nucleus RNA-seq mapped to spatial transcriptomics. As in (B) but colored by Robust Cell Type Deconvolution scores (RCTD) for selected cell types captured by single-nucleus RNA sequencing. RCTD is a supervised learning approach to decompose cell type mixtures in single pseudocells leveraging profiles learned from single-nuclei RNAseq after correcting for platform-specific effects. Each cellular population gets a probabilistic score from 0 to 1 in each spatial location and scores in each spatial location sum up to 1.
- (D) Schematic representation of epithelial and stromal organization during tumor progression.

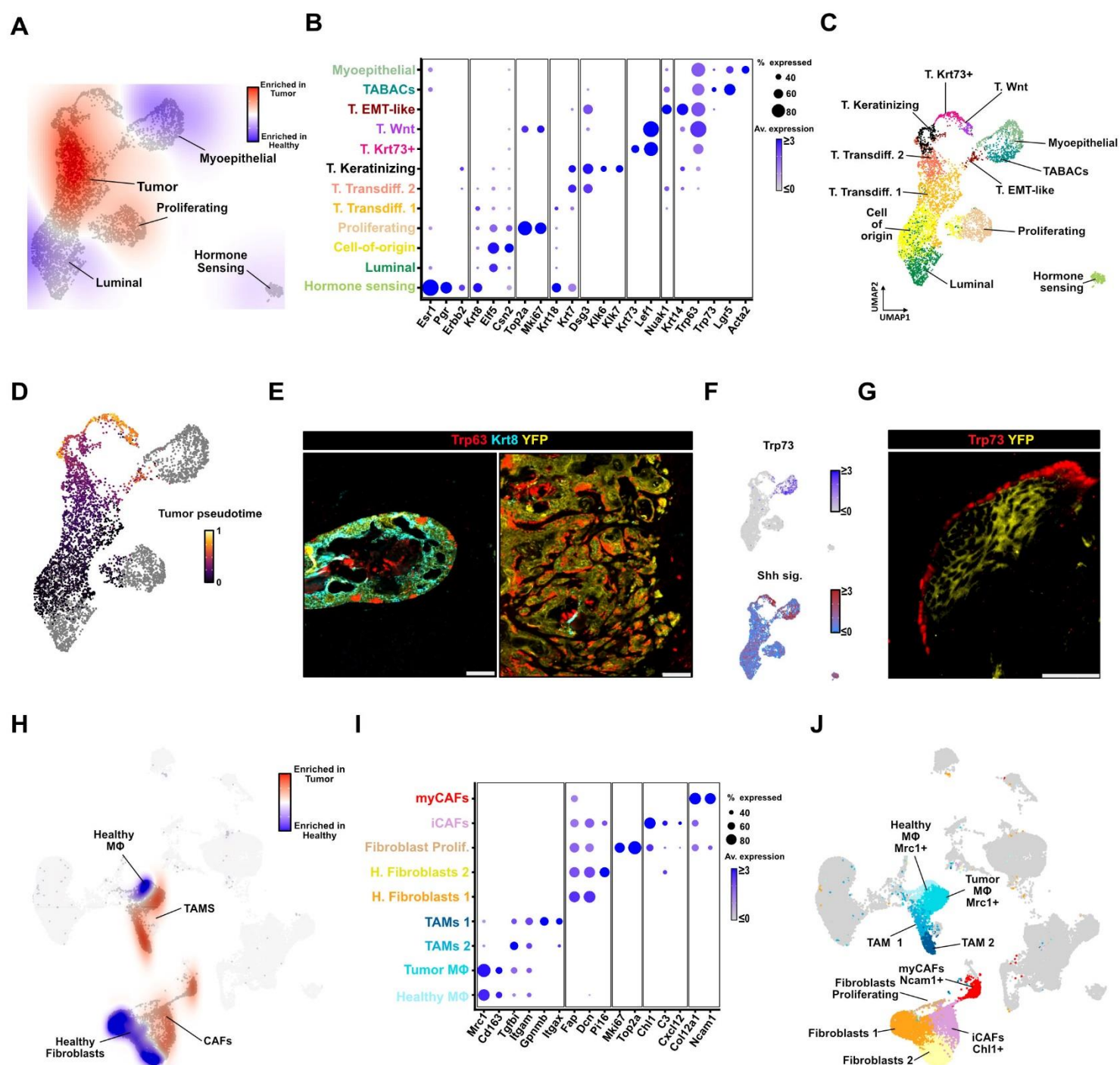
## SFigure 2



## SFigure 2

- (A) Uniform manifold approximation and projection (UMAP) plot of healthy and transformed mammary gland cell types. RNA from  $N = 58,742$  nuclei was sequenced (12 samples, 3 biological replicates/timepoint, Methods).
- (B) Marker gene expression across cell types. Dot size: percentage of cells ( $\geq 20\%$  shown), for each cell type, where  $\geq 1$  transcript of a specific gene was detected. Color: Z-scores of mean gene expression values per cell type.
- (C) As in Fig2B, H&E images of tissue slices (top) and unbiased clustering results of the genome-wide, high-resolution spatial transcriptomics data (Open-ST) for biological replicates not shown in Fig2B.
- (D) UMAPs of the individual snRNA-seq samples can contain lymph node-associated immune and stromal populations (blue). snRNAseq and Open-ST replicates collected from independent mammary glands.
- (E) RCTD deconvolution scores (as in Fig2C, Methods) from lymph-node associated snRNAseq populations indeed map to the intramammary lymph node captured in Open-ST sample T2-#3.

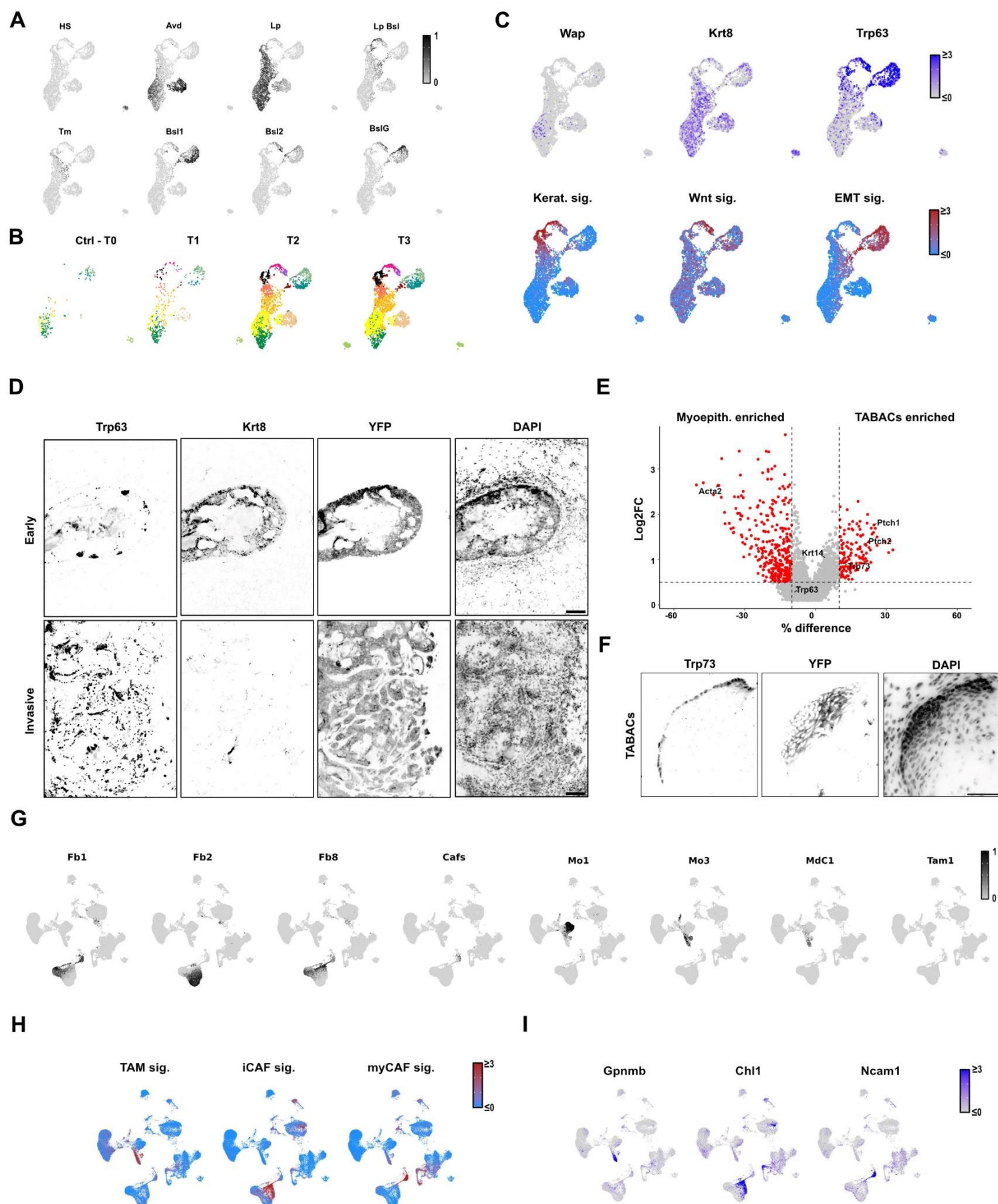
**Figure 3**



### Figure 3 Epithelial and stromal cells undergo extensive remodeling in the TME

- (A) Epithelial cells are remodeled after oncogene activation. UMAP (single nuclei RNA-seq data from epithelial cells, Methods) colored by enrichment in healthy (blue) vs tumor (red) samples.  $N = 5,428$  nuclei integrated from 12 samples (3 biological replicates/timepoint). Color scale clipped between 0.005 and -0.005 for visualization purposes.
- (B) Marker gene expression supports epithelial annotations. Dot size: percentage of cells ( $\geq 20\%$  shown), for each cell state, where  $\geq 1$  transcript of a specific gene was detected. Color: Z-scores of mean gene expression values per cell state. T: Tumor, Av: average.
- (C) As in (A, D, E) but colored by cluster annotations (see B).
- (D) Pseudotime captures tumor progression after oncogene activation. As in (A) but colored by pseudotime scores computed using “Palantir” (Methods), probabilistic modelling of tumor cell trajectories anchored to a Cell-of-origin. Grey: Luminal and myoepithelial cells. Hormone sensing not shown.
- (E) Tumor cells (YFP) down/up-regulate luminal/basal markers during tumor progression (Krt8 and Trp63, respectively). IF staining of early (left) and invasive (right) lesions. Random example ( $N = 3$ ). Scale bars: 50  $\mu\text{m}$ .
- (F) Markers of Tumor-associated basal cells (TABACs). As in (A, C) but colored by Z-scores for Trp73 expression (top) and AUCell enrichment for the ‘Hedgehog Signaling Pathway’ KEGG gene set (bottom; Methods).
- (G) TABACs (Trp73<sup>+</sup>) form a layer of non-recombined cells around transformed cells (YFP<sup>+</sup>). IF staining of an intermediate lesion. Random example ( $N = 3$ ). Scale bars: 50  $\mu\text{m}$ .
- (H) Stromal cells are remodeled in the TME. UMAP (single nuclei RNA-seq data from stromal cells, Methods) colored by enrichment in healthy (blue) vs tumor (red) samples as in (A).  $N = 15,424$  fibroblast and macrophage nuclei from 12 samples (3 biological replicates/timepoint). Color scale clipped between 0.05 and -0.05 for visualization purposes. M $\Phi$  - Macrophage
- (I) Dot plot as in (B) but of marker genes supporting stromal cell state annotations.
- (J) Stromal molecular phenotypes. As in H but colored by cell state annotations (see I).

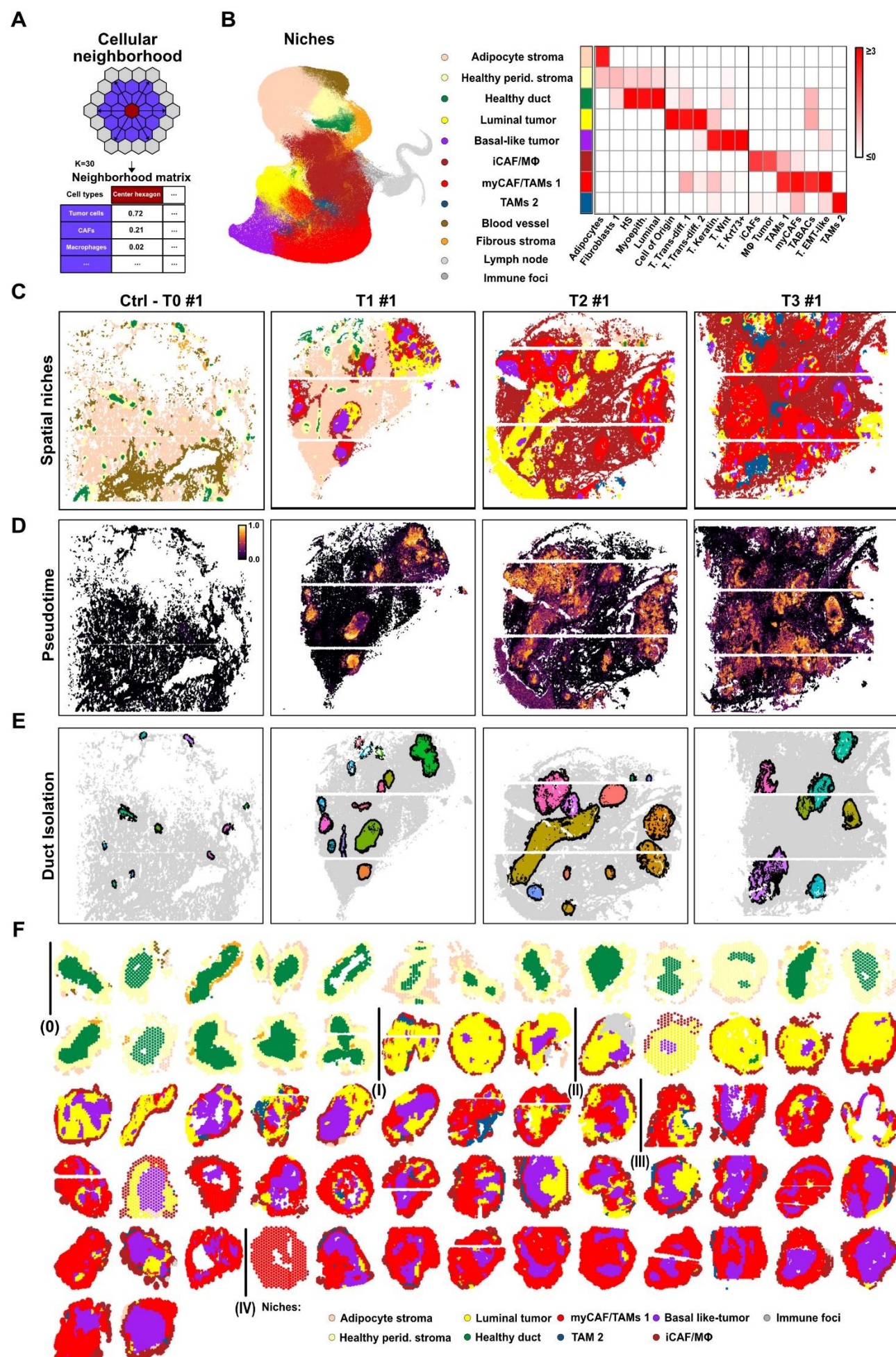
### SFigure 3



### SFigure 3

- (A) Epithelial cell states are similar to specific healthy and tumor reference populations. UMAP (single nuclei RNA-seq data from epithelial cells, as in Fig3A) colored by label transfer scores from a murine basal tumor single-cell atlas<sup>39</sup> (Methods).
- (B) As in (A) but colored by cluster annotations and split by experimental timepoint (T0–T3).
- (C) As in (A) but colored by Z-scores for luminal (Krt8) and basal (Trp63) marker gene expression (top) and AUCell enrichment for HALLMARK pathways upregulated by basal-like tumor cells (bottom; as in Fig3E, Methods).
- (D) Single-channel IF images corresponding to Fig3E. Scale bars: 50  $\mu$ m.
- (E) Volcano plot comparing gene expression between TABACs (right) and myoepithelial cells (left). Selected differentially expressed genes are highlighted. X: Difference in the percentage between TABACs and myoepithelial cells where  $\geq 1$  transcript of a specific gene was detected, Y: log2 fold change in average gene expression.
- (F) Single-channel IF images corresponding to Fig3G. Scale bars: 50  $\mu$ m.
- (G) Stromal cell states are similar to specific healthy and tumor reference populations. As in (G,H) but colored by label transfer scores from a murine mammary single-cell reference (see B).
- (H) Stromal cells upregulate cancer-associated signatures. UMAP (single nuclei RNA-seq data from stromal cells, as in Fig3H) colored by AUCell enrichment Z-scores for inflammatory cancer-associated fibroblasts (iCAFs), myofibroblasts (myCAFs) and tumor-associated macrophages (TAMs) signatures described in (Mayer et al., 2023)<sup>42</sup>.
- (I) Remodeled stromal cells upregulate specific markers. As in (G) but colored by Z-scores for mean gene expression.

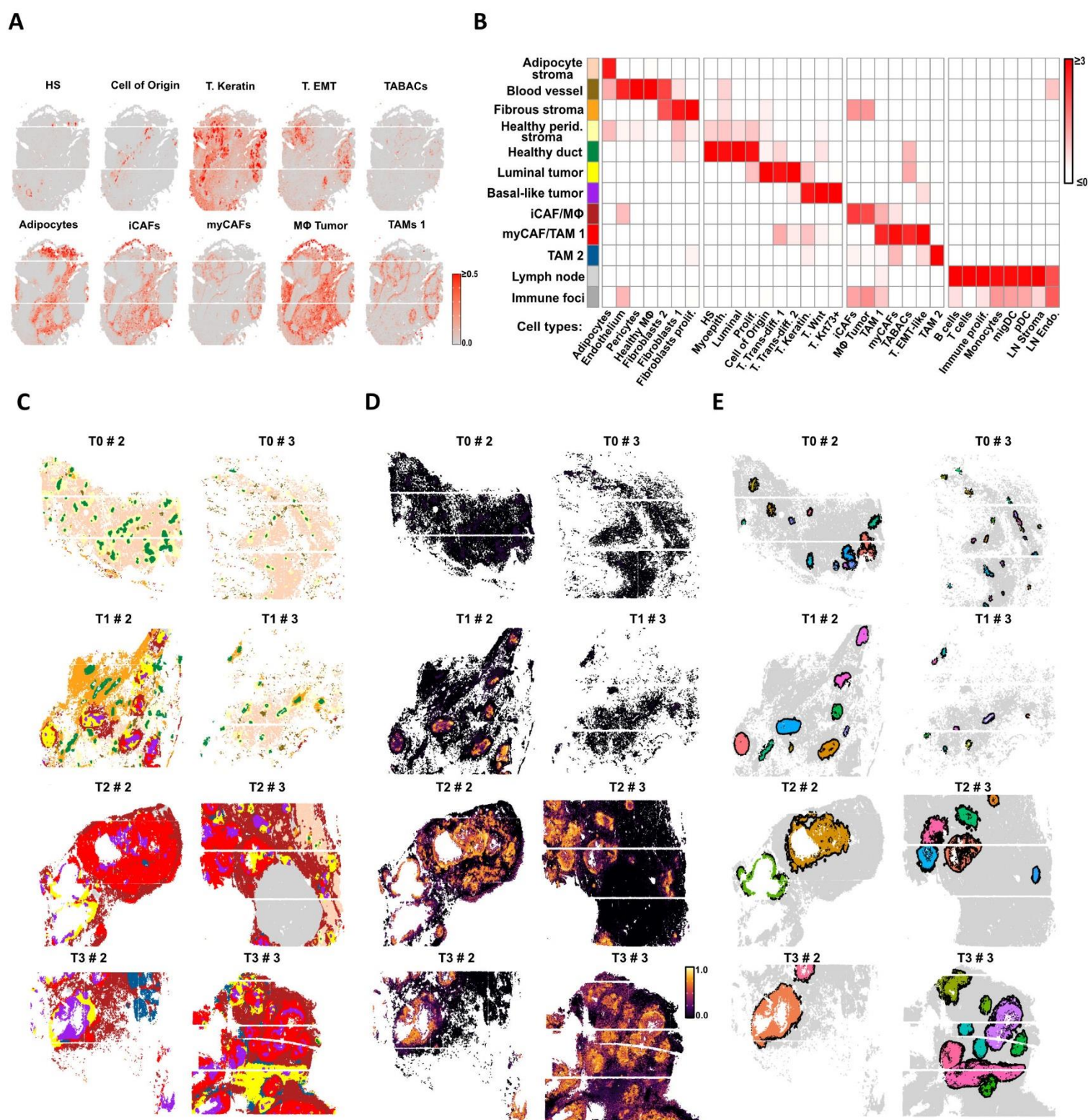
Figure 4



# **Figure 4. Tumor and stromal states are spatially organized in dynamic multicellular niches**

- (A) Definition of cellular neighborhoods. RCTD deconvolution scores are averaged per cellular neighborhood (*i.e.*, 30 closest pseudocells).
- (B) Epithelial and stromal populations are spatially organized. Cellular neighborhoods with similar cell state deconvolution scores are grouped into niches (Methods). Left: UMAP colored by niche identity. Right: Heatmap of Z-scores for mean RCTD deconvolution scores of epithelial and stromal populations. Additional niches and cell types in SFig4B.
- (C) Spatial mapping of epithelial and stromal niches during tumorigenesis. Open-ST samples (as in Fig2B) are colored by niche assignment (see B). Additional replicates in SFig. 4C.
- (D) Spatial mapping of tumor pseudotime scores (see Fig3C). As in (C) but colored by pseudotime label transfer scores (Methods). Additional replicates in SFig4D.
- (E) Duct segmentation (Methods). As in (C) and (D) but colored by duct assignment. Black: duct-stromal interface. Additional replicates in SFig4E.
- (F) Pseudotime ordering of 100+ ducts. Pseudocells in individual ducts colored by niche assignments. Ducts ordered by average pseudotime score. Stages 0-IV represent 5 pseudotime bins. Ducts with <250 pseudocells were removed (44 out of 111).

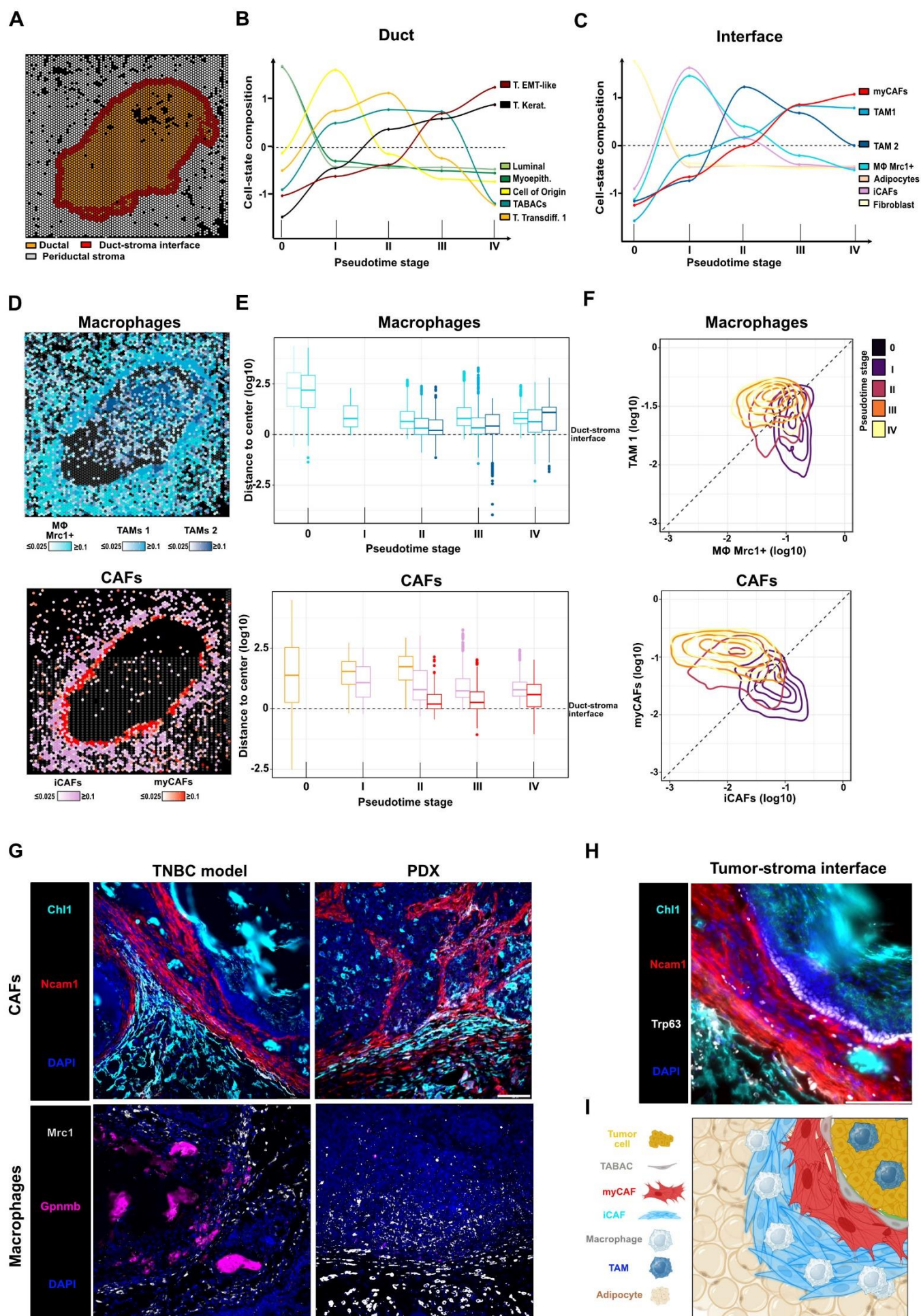
**SFigure 4**



## SFigure 4

- (A) Single-nucleus RNA-seq epithelial and stromal states mapped to spatial transcriptomics. RCTD deconvolution scores (see Fig2C, Methods) for selected epithelial and stromal populations are shown in Open-ST sample T2 #1.
- (B) As in Fig4B but showing Z-scores for mean RCTD deconvolution scores of all epithelial, stromal, and immune populations across all identified multicellular niches.
- (C) Spatial mapping of multicellular niches across additional biological replicates not shown in Fig4C. Color legend in B.
- (D) Spatial mapping of tumor pseudotime scores across additional biological replicates not shown in Fig4D.
- (E) Spatial segmentation of individual ducts across additional biological replicates not shown in Fig4E.

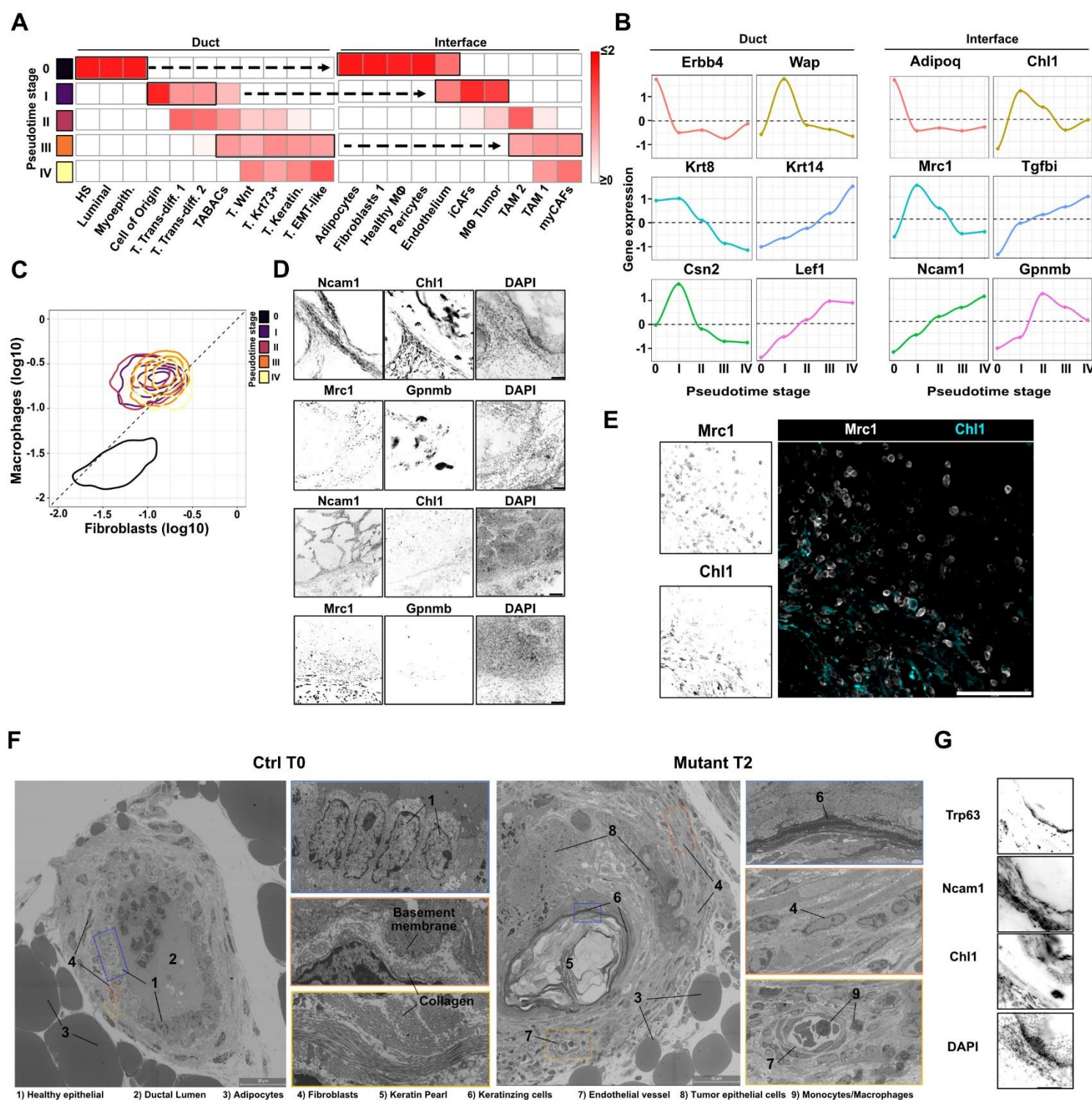
Figure 5



## Figure 5 Spatiotemporal dynamics of tumor-stromal interface remodeling

- (A) Duct isolation enables the study of the duct-stromal interface (red) (Methods).
- (B) Composition dynamics of epithelial phenotypes within ducts. Cell-state compositions are defined as Z-scores of mean, stage-specific RCTD deconvolution scores (as in SFig4A top row). Additional epithelial states in SFig4A, left.
- (C) Composition dynamics of stromal phenotypes at the duct-stromal interface. As in B but related to stromal populations mapped to the duct-stromal interface. Additional stromal states in SFig4A, right.
- (D) Stromal phenotypes are spatially organized in the periductal stroma. Spatial transcriptomics of the same duct shown in (A), displaying macrophage (top) and CAF (bottom) cell-state composition.
- (E) Spatial organization of stromal phenotypes across duct stages (time). Box plots show the distance of individual macrophage (top) and fibroblast (bottom) populations to the tumor–stroma interface (Methods). Boxes omitted for populations detected in <150 stage-specific pseudocells. Color legend in C.
- (F) Progressive remodeling of stromal states at the tumor-stromal interface. Phase diagrams of stromal dynamics during tumor progression. Cell-state composition of neighbourhoods at the duct-stromal interface (here shown in log10 scale) are used to compute 2D kernel densities for each duct stage. Stage 0 not shown as remodeled stromal populations only appear later (see C).
- (G) Cell type distribution at the duct-stromal interface. IF staining of iCAFs (Chl1<sup>+</sup>), myCAFs (Ncam1<sup>+</sup>), macrophages (Mrc1<sup>+</sup>), and TAMs (Gpnmb<sup>+</sup>) in an intermediate-stage tumor (left) and a PDX(#3) model (right). Scale bars: 50  $\mu$ m.
- (H) TABACs (Trp63<sup>+</sup>) form a layer between tumor cells (YFP<sup>+</sup>) and myCAFs (Ncam1<sup>+</sup>) interface. IF staining showing the tumor–stroma interface in a primary tumor. Scale bars: 100  $\mu$ m.
- (I) Graphical representation of cell types organization at the tumor-stromal interface.

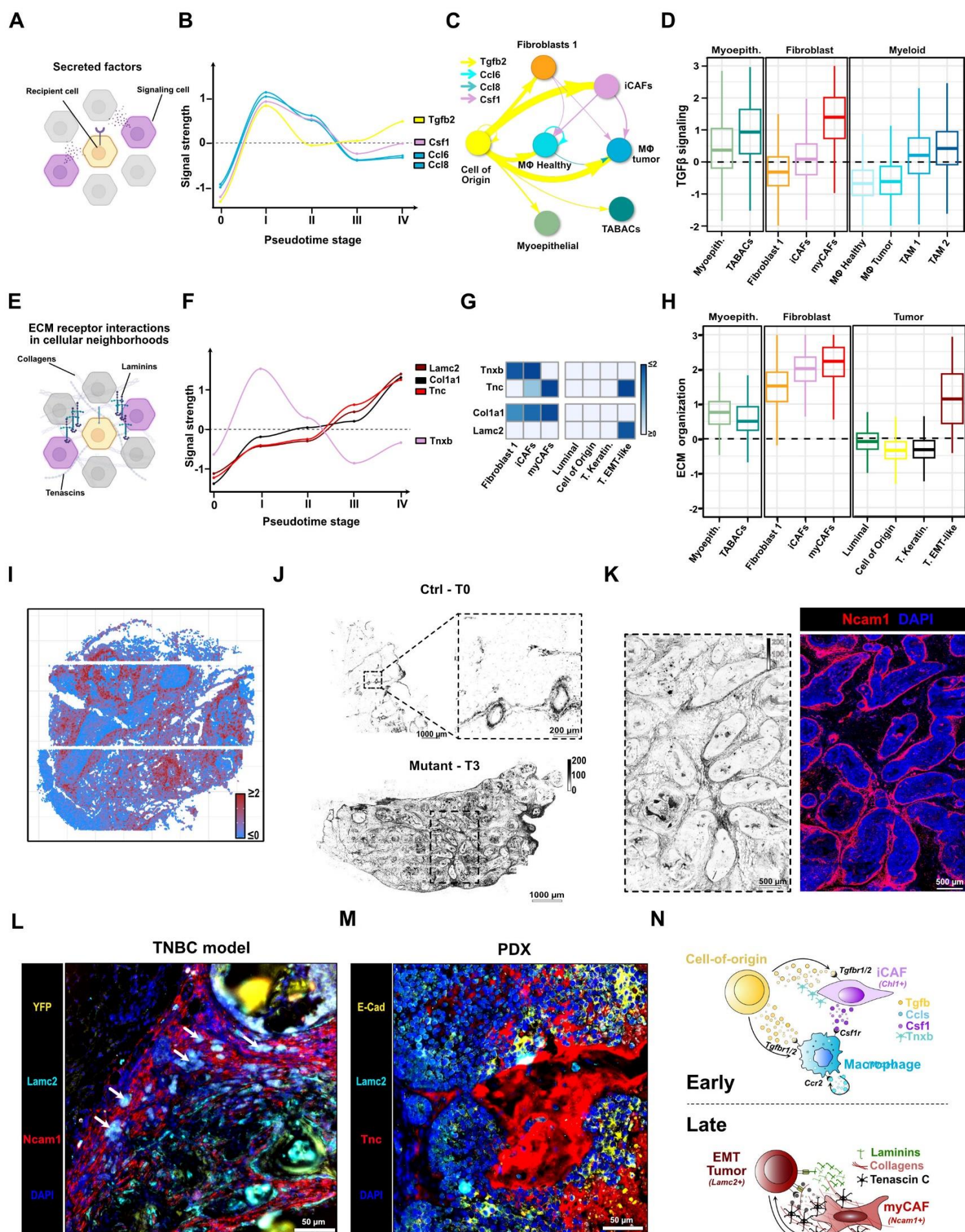
## SFigure 5



## SFigure 5

- (A) Specific stromal compositions at the duct-stromal interface follow specific epithelial phenotypes within ducts. Cell-state composition (Z-scores of mean RCTD deconvolution scores) of epithelial populations in the duct (as in Fig5B) and stromal populations in the duct-stromal interface (as in Fig5C).
- (B) Marker gene expression follows population dynamics. Z-scores of mean expression of epithelial marker genes in the duct compartment and stromal markers genes in the duct-stromal interface.
- (C) A macrophage-rich/fibroblast-rich 'hot fibrosis' state is established at the tumor-stromal interface. As in Fig5F but showing the cumulative abundance of fibroblast and macrophage states and including healthy ducts.
- (D) Single-channel IF images corresponding to Figure 5G. Individual channels shown for iCAFs (Chl1<sup>+</sup>) and macrophages (Mrc1<sup>+</sup>). Scale bars: 50  $\mu$ m.
- (E) iCAFs are closely associated with macrophages in the periductal stroma, from T2. IF staining of iCAFs (Chl1<sup>+</sup>) and macrophages (Mrc1<sup>+</sup>) in T2 tumor. Random example (N=3)
- (F) Transmission electron microscopy of mammary gland ducts derived from healthy and tumor-bearing animals. T0 and T2 Tumor lesion with disrupted epithelial architecture, dense fibroblasts layer around the tumors and multiple macrophages infiltrating the stroma.
- (G) Single-channel IF images corresponding to Figure 5H. Individual channels of IF images shown for Ncam1<sup>+</sup> myCAFs, Trp63<sup>+</sup> myoepithelial basal cells, YFP<sup>+</sup> tumor cells, and DAPI. Scale bars: 50  $\mu$ m.

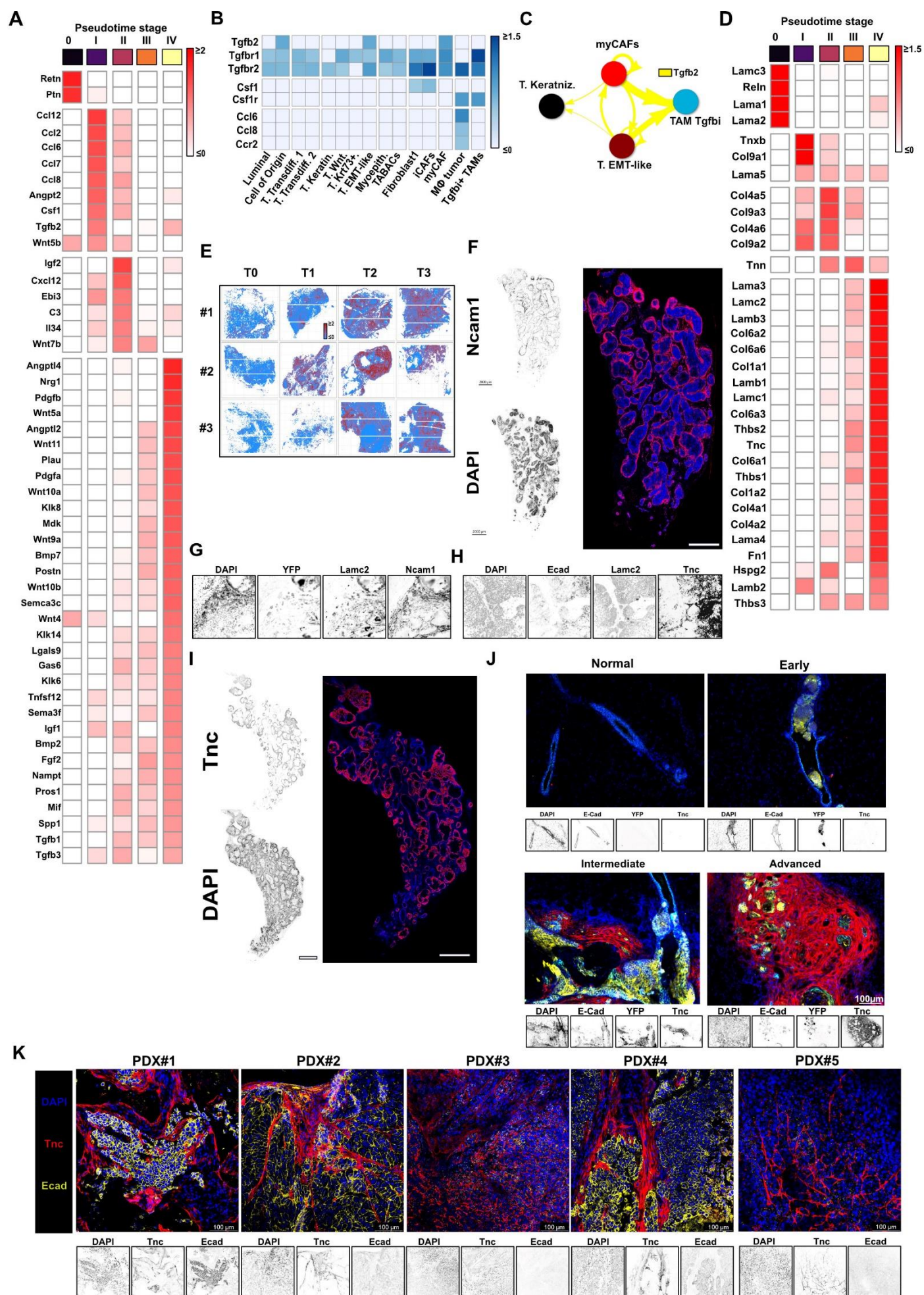
**Figure 6**



## Figure 6. Coordinated and transient signaling axes orchestrate early interface remodeling

- (A) Spatial analysis of secreted ligand-receptor interactions in cellular neighborhoods. Ligand spatial activity is computed based on co-expression of ligand and its receptors in neighboring cells (Methods). Pairs of ligand and receptor pairs from CellChat<sup>44</sup>.
- (B) Transient induction of signaling axes at the early (Stage I) tumor-stromal interface. As in Fig5B but showing Z-scores of mean spatial ligand activity, computed for pseudocells forming the tumor-stromal interface. Additional Stage I and other stage-enriched ligands shown in SFig6A.
- (C) Tgfb2 signaling from the cell-of-origin is received by early-interface populations. snRNA-seq data is used to identify the directionality of stage-specific interface interactions (see B). Arrow width is proportional to geometric of ligand/receptor expression in the sender/receiver populations mapping to the early duct-stromal interface (Stage I). Arrows shown for receptor and ligand expression detected in >30% of cells.
- (D) Intracellular targets of TGFb signaling are upregulated in early-interface populations. Box plots showing Z-scores of AUCell enrichment for HALLMARK 'TGFβ signaling' gene set in snRNA-seq populations mapping to the duct-stromal interface.
- (E) Spatial analysis of ECM-receptor interactions in cellular neighborhoods. As in A but focusing on ECM ligand–receptor pairs.
- (F) Progressive ECM remodeling at the tumor-stromal interface. As in B but showing the dynamic of selected ECM ligands. All stage-enriched ligands are shown in SFig6D.
- (G) Heatmap showing mean ligand expression across selected populations from snRNA-seq data.
- (H) myCAFs orchestrate ECM organization. Z-score for AUCell enrichment of 'ECM organization' REACTOME gene set in selected snRNAseq populations.
- (I) ECM remodeling at the tumor-stromal interface. As in H but scoring a representative spatial transcriptomic sample (T2 #1). All replicates shown in SFig6F.
- (J) Collagen is deposited at the tumor-stromal interface. Second harmonic generation (SHG) imaging of collagen deposition in healthy control T0 (left) and T3 tumor (right) tissue sections. Zoom-out and zoom-in views shown. Scale bars: 1000 μm (overview), 200 μm (inset).
- (K) Collagen deposition colocalizes with myCAFs at the tumor-stromal interface. IF and SHG imaging of the same tumor region from consecutive sections. Top: Ncam1<sup>+</sup> myCAFs shown by IF. Bottom: corresponding SHG image showing collagen deposition in the same field. Scale bars: 50 μm.
- (L) EMT-like tumor cells invade in a remodeled ECM at the tumor-stromal interface. IF staining in a T3 tumor showing Lamc2<sup>+</sup>/YFP<sup>+</sup> tumor cells in proximity to Ncam1<sup>+</sup> myCAFs. Random example (N=3), Scale bar: 50 μm.
- (M) Lamc2<sup>+</sup> tumor cells infiltrate in a remodeled ECM in PDX samples. IF staining of a PDX sample showing Lamc2<sup>+</sup>/Ecad<sup>+</sup> tumor cells in a Tnc<sup>+</sup> stroma. Scale bar: 50 μm. Random example (N=3)
- (N) Graphical representation of signaling events in early and late tumorigenesis.

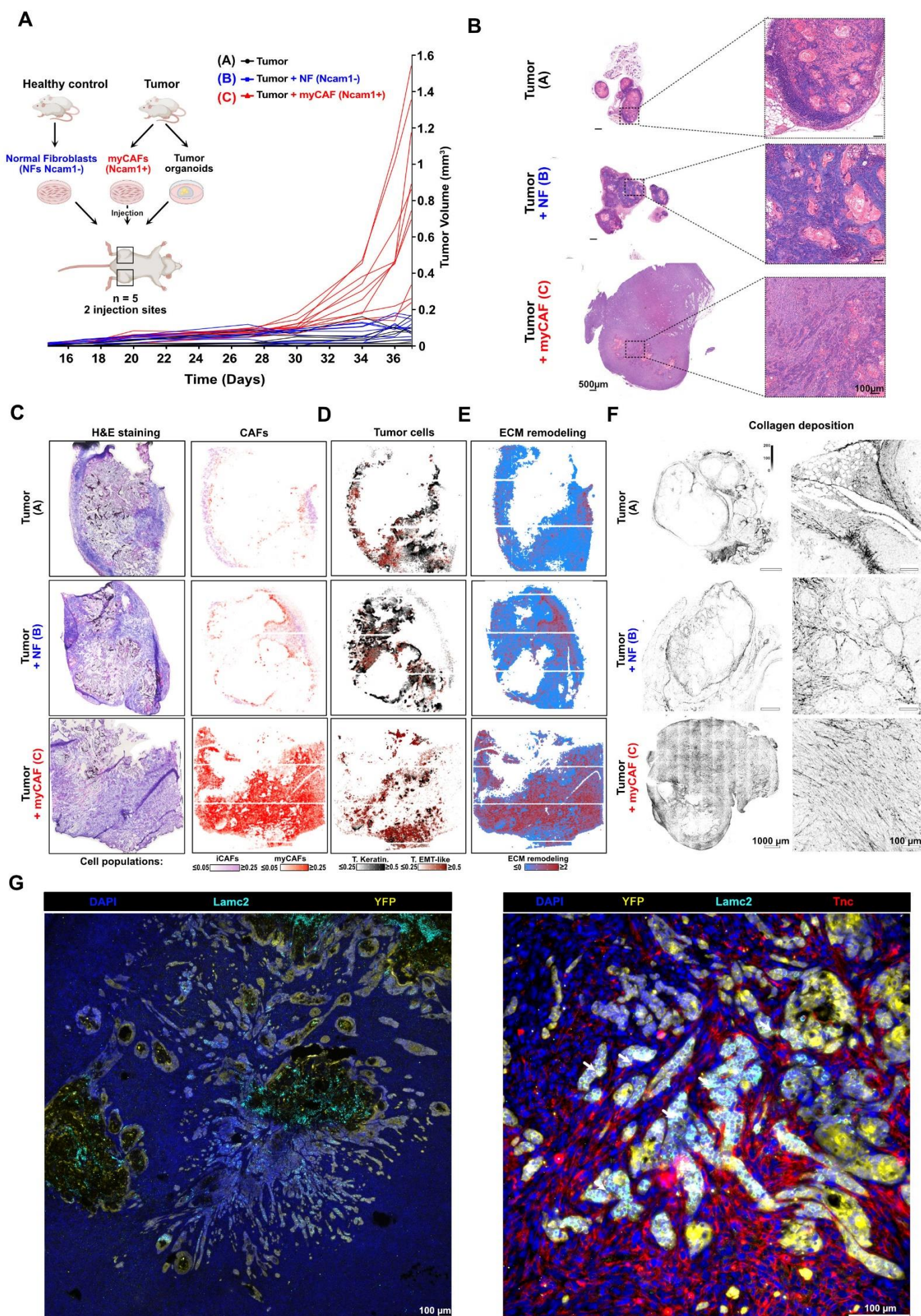
## SFigure 6



## SFigure 6

- (A) Stage-specific secreted interactions at the tumor-stromal interface. Heatmap showing Z-scores of mean spatial ligand activity (computed for spatial pseudocells, as in Fig6B). Stage-specificity is defined by both  $\log_2FC > 0$  in average ligand activity and non-zero score in at least 5% of stage-specific interface pseudocells.
- (B) Early interface ligand and receptor expression in single-nuclei populations. Heatmap showing average expression of selected early-stage ligands and their receptors across snRNA-seq populations.
- (C) myCAFs and EMT tumor cells secrete Tgfb2 towards TAMs at the late tumor-stromal interface. As in Fig6C but showing Tgfb2 interactions snRNAseq population mapping to the advanced (stage III/IV) tumor-stromal interface.
- (D) Stage-specific ECM ligand-receptor interactions at the tumor-stromal interface. As in A but related to ECM ligands.
- (E) Spatial activity of ECM organization pathway localizes to the tumor-stromal interface. As in Fig6I but across all Open-ST samples.
- (F) Whole-gland immunofluorescence image corresponding to Figure 6J. Merged and single-channel images for Ncam1 and DAPI are shown in a T3 tumor section. Random example (N=3)
- (G) Individual channels of IF images shown for Ncam1, Lamc2, YFP, and DAPI corresponding to Fig 6K. Scale bars: 50  $\mu\text{m}$ .
- (H) Individual channels of IF images shown for Tnc, Lamc2, Ecad, and DAPI corresponding to Fig 6I. Scale bars: 50  $\mu\text{m}$ .
- (I) Whole-gland immunofluorescence merged image and single-channel images for Tnc and DAPI are shown in a T3 tumor section. Random example (N=3)
- (J) Immunofluorescence merged image and single-channel images for Ecad, Tnc and DAPI are shown in all stages T0-T3 section. Random example (N=3)
- (K) Immunofluorescence images and single-channel images for Ecad, Tnc and DAPI are shown in a panel of human TNBC PDXs sections, Scale bars: 100  $\mu\text{m}$ .

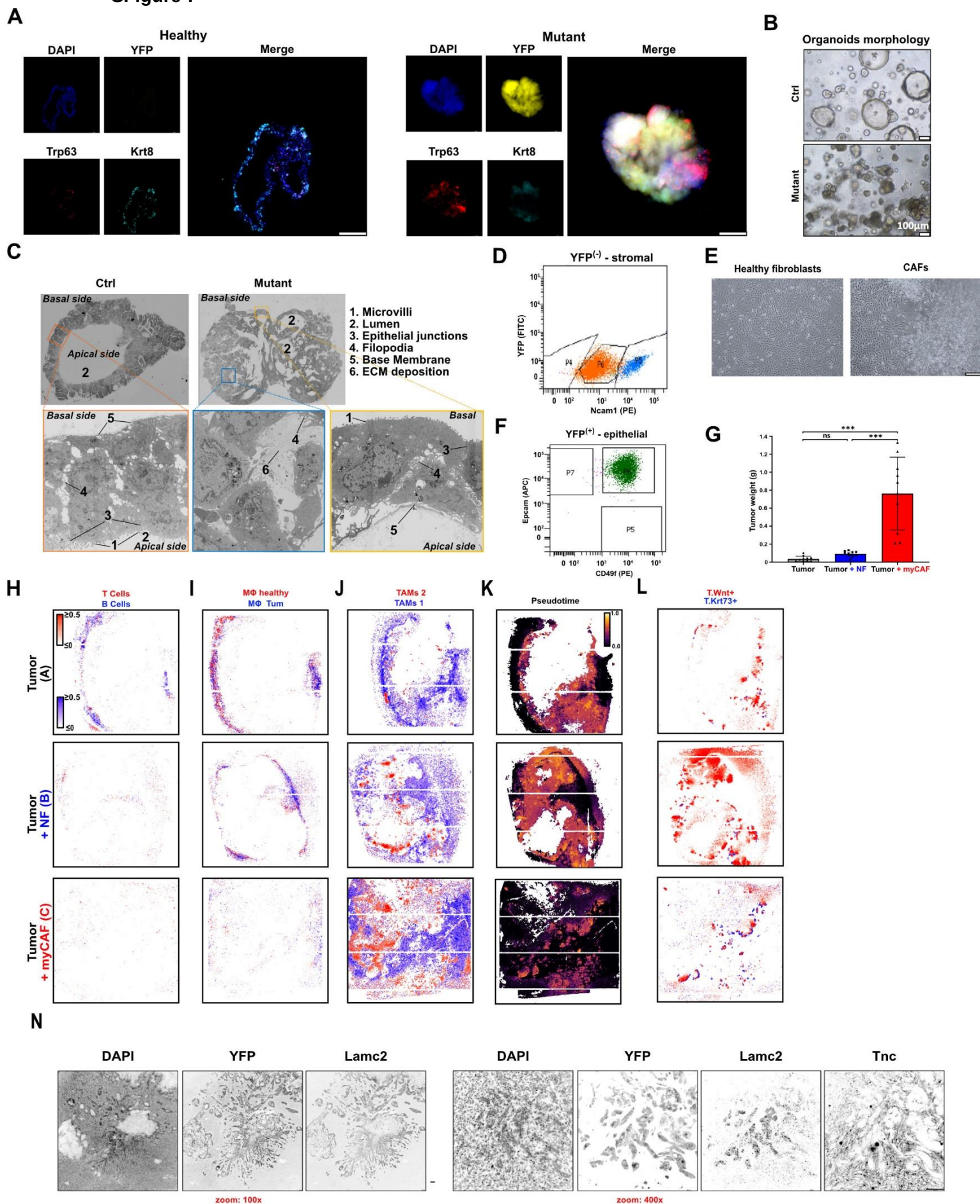
**Figure 7.**



## Figure 7. Myfibroblasts promote tumor growth and invasion

- (A) myCAFs promote tumor growth when co-transplanted with tumor cells in healthy immunocompetent mice. Top: experimental design for orthotopic co-transplantation. Tumor cells were injected alone (Group A), with healthy Ncam1<sup>-</sup> fibroblasts (Group B), or with tumor-derived Ncam1<sup>+</sup> myCAFs (Group C) into the mammary fat pad of immunocompetent FVB/N mice. Bottom: Tumor volume was measured daily from day 15 to day 37, the experimental endpoint defined by maximum allowable tumor size. Tumor growth at individual injection sites is shown over time in mm<sup>3</sup> (N=10 per group).
- (B) myCAF co-transplant promotes disorganized tumor histology. Hematoxylin and eosin (H&E) staining of representative tumors from Groups A–C. Random example (N=3) Scale bars: 500  $\mu$ m.
- (C) Spatial transcriptomic (Open-ST) profiling of transplant tumors. Spatial mapping (RCTD deconvolution as in Fig2C and 5D) of iCAF and myCAFs snRNAseq populations onto Open-ST data from one tumor/group. Immune populations shown in SFig7H–J.
- (D) myCAFs promote invasive tumor phenotypes. As in C but showing RCTD deconvolution scores for EMT-like and keratinizing snRNAseq tumor populations. Tumor pseudotime and additional tumor populations shown in SFig7K–L.
- (E) myCAFs promote ECM remodeling. As in Fig6I but related to transplant tumors in C.
- (F) myCAFs promote the deposition of a pro-invasive ECM. Second harmonic generation imaging (as in Fig6J) of collagen deposition at tumor injection sites for Groups A–C. For each group, a zoomed-out view (scale bar: 1000  $\mu$ m) and a corresponding zoom-in (scale bar: 100  $\mu$ m) are shown. Collagen fiber organization and density are visualized across the tumor and surrounding stroma.
- (G) myCAFs promote tumor invasion in a remodeled ECM. IF staining of a Group C tumor. Left: overview of the tumor area showing Lamc2 and YFP. Right: magnified region displaying colocalization of Lamc2, YFP, and Tnc. Arrows indicate sites of colocalization. Scale bars: 500  $\mu$ m (left), 100  $\mu$ m (right).

## SFigure 7



## SFigure 7

- (A) Tumor organoids display a basal phenotype. IF staining of organoid lines from healthy (left) and mutant (right) mammary glands. Shown channels: DAPI, YFP, Trp63, and Krt8. Scale bars: 100  $\mu$ m.
- (B) Tumor organoids display a disorganized architecture with loss of the central lumen. Brightfield images of organoid cultures showing the structural organization of epithelial cells. Healthy organoids form acinar-like structures with central lumens; tumor organoids display disorganized architecture. Scale bars: 100  $\mu$ m.
- (C) Tumor organoids display protrusions invading the surrounding ECM. Transmission electron microscopy of organoids derived from healthy and tumor-bearing mammary glands. Tumor organoids show disrupted epithelial architecture, loss of apical–basal polarity, and protrusions extending into the extracellular matrix. Scale bars: as indicated.
- (D) Tumor fibroblast cultures Ncam1<sup>+</sup> and Ncam1<sup>-</sup> negative populations. Flow cytometry analysis of tumor-derived fibroblasts stained with PE-conjugated anti-Ncam1 antibody. Cells were gated negative for Ptprc and EpCAM. Distinct Ncam1<sup>+</sup> and Ncam1<sup>-</sup> populations are shown.
- (E) Tumor fibroblast cultures display heterogenous morphologies. Brightfield images of fibroblast cultures derived from healthy and tumor-bearing mammary glands showing differences in morphology. Scale bars: 100  $\mu$ m.
- (F) Epithelial cells used for injection experiment were enriched for LP state. Flow cytometry plot showing the sorting strategy for YFP<sup>+</sup> epithelial cells using PE-CD49f and APC-EpCAM staining, based on the luminal progenitor gating approach described in Rosenbluth et al., 2020. Random example (N=3).
- (G) myCAFs co-injection leads to increased tumor weight at study termination. Fatpad weights (in gram) at the day of study termination for Groups A, B, and C. Bars represent endpoint tumor mass in grams. *N* =10.
- (H-J) Spatial mapping of lymphocytes (H), macrophages (I) and TAM (J) snRNAseq populations in Open-ST transplant samples (as in Fig7C).
- (L) Spatial mapping of tumor pseudotime (as in Fig 4D) but related to transplant samples in H-J.
- (M) As in (H-J) but related to selected tumor populations not shown in Fig7D.
- (N) Immunofluorescence staining of tumors from group A-C Shown channels: DAPI, YFP, Ncam1. Scale bars: 100  $\mu$ m.
- (O) Single immunofluorescence channels related to Figure 7E. Single-channel immunofluorescence images corresponding to Figure 7E. Channels shown: Lamc2, YFP, Tnc, and DAPI. Scale bars: 100  $\mu$ m.