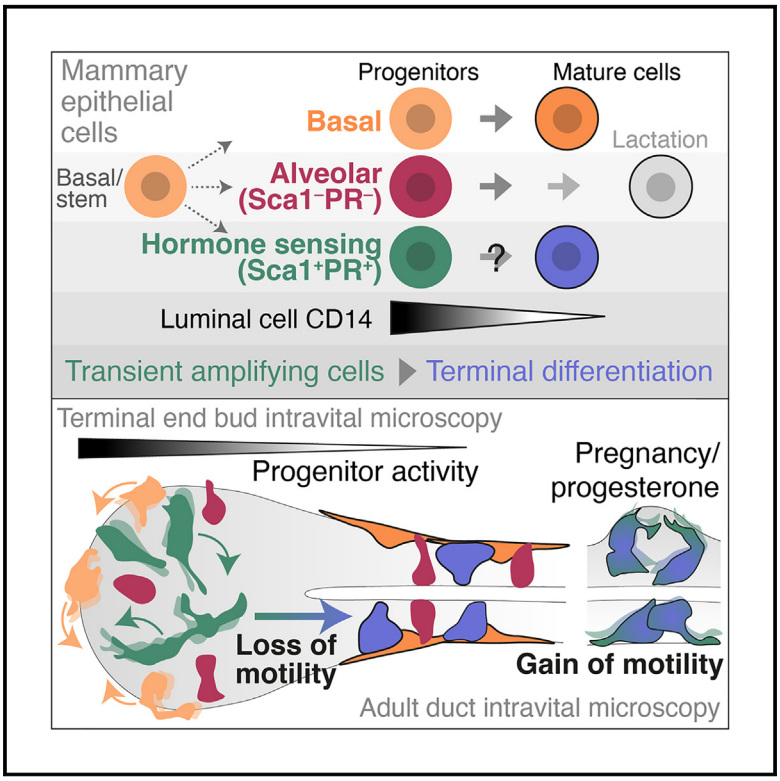
Cell Reports

Article

Hormone-responsive progenitors have a unique identity and exhibit high motility during mammary morphogenesis

Graphical abstract



Highlights

- CD14 and Sca1 definitively resolve hormone-receptor-positive luminal progenitors
- HR$^+$ progenitors are uniquely proliferative unlike HR$^+$ mature luminal cells
- HR$^+$ cells are highly elongated and motile within TEBs but not ducts
- HR$^-$ progenitors are columnar and static in TEBs and ducts

Authors

Caleb A. Dawson, Michael J.G. Milevskiy, Bianca D. Capaldo, ..., Yunshun Chen, Geoffrey J. Lindeman, Jane E. Visvader

Correspondence

visvader@wehi.edu.au

In brief

Dawson, Milevskiy, et al. investigate the phenotypic, cycling, and behavioral properties of the hormone receptor (HR)$^+$ and HR$^-$ mammary luminal lineages. They show that HR$^+$ progenitors have a distinct identity and are remarkably motile within terminal endbuds compared to their ductal HR$^+$ counterparts, suggesting that they help drive epithelial morphogenesis.

Dawson et al., 2024, Cell Reports *43*, 115073
December 24, 2024
https://doi.org/10.1016/j.celrep.2024.115073

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Hormone-responsive progenitors have a unique identity and exhibit high motility during mammary morphogenesis

Caleb A. Dawson,[1,2,3,10] Michael J.G. Milevskiy,[1,3,10] Bianca D. Capaldo,[1,3] Raymond K.H. Yip,[3,4] Xiaoyu Song,[1,2] François Vaillant,[1,3] Lexie Prokopuk,[1,3,9] Felicity C. Jackling,[1] Gordon K. Smyth,[5,6] Yunshun Chen,[1,3,5] Geoffrey J. Lindeman,[1,7,8] and Jane E. Visvader[1,2,11,*]

[1]ACRF Cancer Biology and Stem Cells Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[2]Immunology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[3]Department of Medical Biology, The University of Melbourne, Parkville, VIC 3010, Australia
[4]Advanced Technology and Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[5]Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[6]School of Mathematics and Statistics, The University of Melbourne, Parkville, VIC 3010, Australia
[7]Department of Medicine, Royal Melbourne Hospital, The University of Melbourne, Parkville, VIC 3010, Australia
[8]Parkville Familial Cancer Centre and Department of Medical Oncology, The Royal Melbourne Hospital and Peter MacCallum Cancer Centre, Parkville, VIC 3052, Australia
[9]Present address: The University of Queensland, Australian Women and Girls' Health Research Centre, QLD, Australia
[10]These authors contributed equally
[11]Lead contact
*Correspondence: visvader@wehi.edu.au
https://doi.org/10.1016/j.celrep.2024.115073

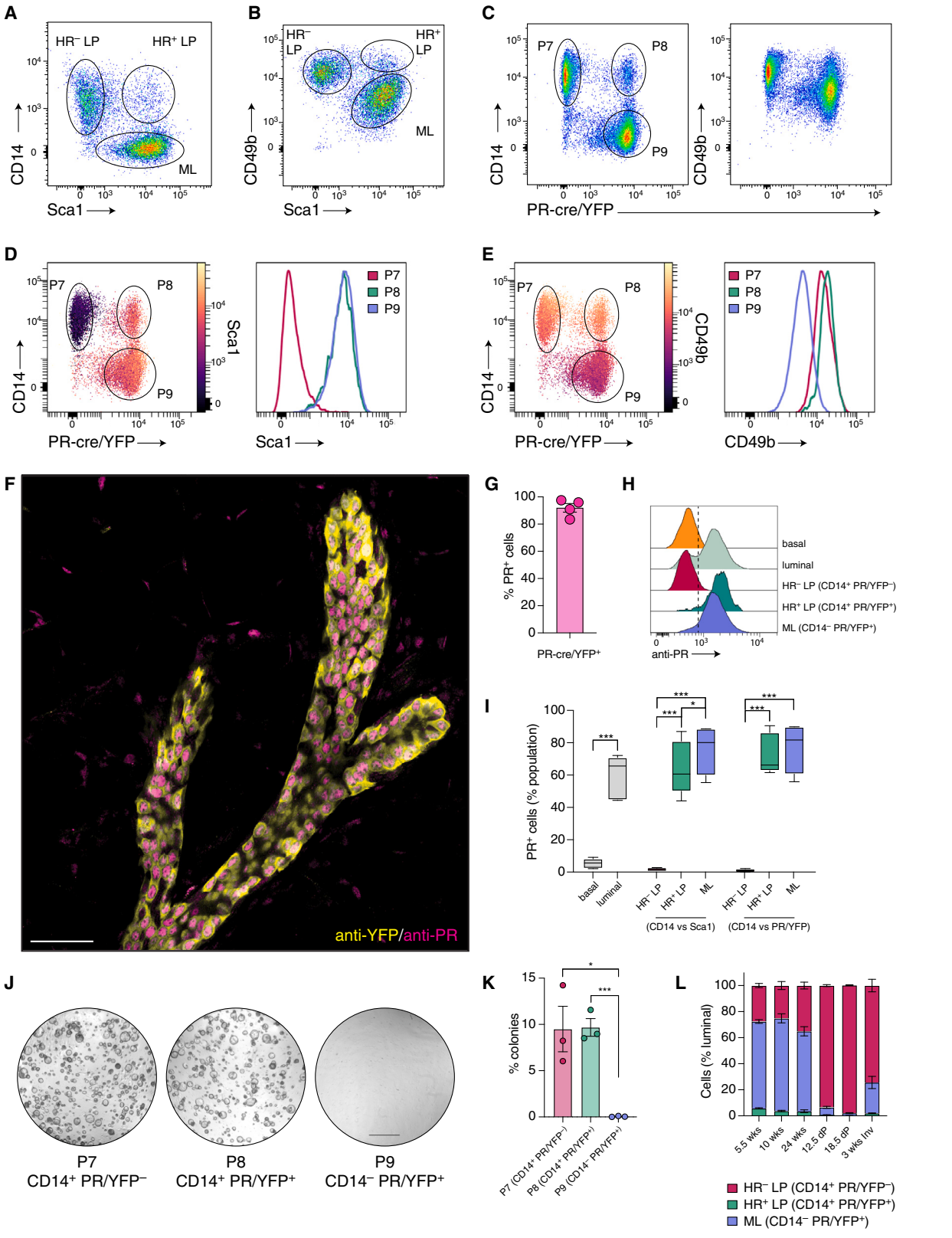
SUMMARY

Hormone-receptor-positive (HR$^+$) luminal cells largely mediate the response to estrogen and progesterone during mammary gland morphogenesis. However, there remains a lack of consensus on the precise nature of the precursor cells that maintain this essential HR$^+$ lineage. Here we refine the identification of HR$^+$ progenitors and demonstrate their unique regenerative capacity compared to mature HR$^+$ cells. HR$^+$ progenitors proliferate but do not expand, suggesting rapid differentiation. Subcellular resolution, 3D intravital microscopy was performed on terminal end buds (TEBs) during puberty to dissect the contribution of each luminal lineage. Surprisingly, HR$^+$ TEB progenitors were highly elongated and motile compared to columnar HR$^-$ progenitors and static, conoid HR$^+$ cells within ducts. This dynamic behavior was also observed in response to hormones. Development of an AI model for motility dynamics analysis highlighted stark behavioral changes in HR$^+$ progenitors as they transitioned to mature cells. This work provides valuable insights into how progenitor behavior contributes to mammary morphogenesis.

INTRODUCTION

The mammary gland largely develops in the postnatal phase, undergoing extensive expansion during puberty and pregnancy. The steroid hormones estrogen and progesterone, together with their cognate receptors (ER and PR, respectively), play pivotal roles in driving mammary gland morphogenesis.[1] Estrogen controls branching and elongation of the ductal tree through the mammary fat pad during puberty, while progesterone governs tertiary branching in the adult gland and alveologenesis during pregnancy.[2] Epithelial cells within the mammary ductal system are arranged in a bilayer, with an outer layer of myoepithelial (basal) cells and an inner layer of luminal cells. New duct formation in puberty is driven by highly proliferative terminal end buds (TEBs), which are large, club-like structures comprising an outer cap layer and a multi-layered body, that give rise to the basal and luminal lineages, respectively. Ductal growth during this phase occurs by proliferation and rearrangement of epithelial progenitor cells in TEBs, leaving behind a simple bilayer as they extend and branch through the fat pad.[3]

The luminal compartment can be separated into proliferative luminal progenitor (LP) cells and non-proliferative mature luminal (ML)[4,5] cells that are predominantly hormone-receptor (HR) positive (HR$^+$) (reviewed in Tharmapalan et al. and Fu et al.[6,7]). The LP pool is composed of a large HR$^-$ subset and a small HR$^+$ population,[5,8] with HR$^+$ progenitors presumed to be the precursors for the HR$^+$ ML population. Interestingly, lineage tracing studies have indicated that the ER$^+$ and ER$^-$ lineages are independently maintained during homeostasis.[9–11] Like mice, human mammary tissue contains both HR$^+$ and HR$^-$ LP cells, and HR$^+$ ML cells that lack proliferative activity.[12,13] Accordingly, in both human and mouse mammary tissue, most proliferative epithelial cells are ER$^-$/PR$^-$, consistent with steroid hormones exerting their mitogenic effects through paracrine signaling.[14–18] More recent



(legend on next page)

studies, however, have presented paradoxical findings on the murine luminal compartment. Most cell division *in vivo* was reported to occur within the HR$^+$ ML population rather than in LP cells, despite minimal clonogenic activity *in vitro*.[19,20] Hence, there remain unresolved discrepancies on the proliferative activity and progenitor function of mammary luminal populations. Accurate delineation of the HR$^+$ and HR$^-$ progenitor populations is pivotal to understanding the origins of the distinct subtypes of breast cancer.

Mammary epithelial progenitors not only maintain the mammary epithelial lineages but drive ductal elongation in TEBs by a process of collective migration. TEB cells do not extend into the tissue; rather, the TEB surface is smooth, and forward forces are generated by internal cell rearrangement into a bilayer while the TEB mass is sustained by proliferation.[3,21] As ducts mature, there is a transition from a disordered, multi-layered, and loosely adhered structure in the TEB to a lengthened and tightly adherent bilayer through radial cell intercalation.[22,23] This involves cell migration controlled through phosphatidylinositol 3-kinase (PI3K)/MEK/ERK signaling downstream of receptor tyrosine kinase activation[24] and physical constraint by the contractile myoepithelial layer and the basement membrane.[23] TEB branching occurs when a local reduction in cell motility and proliferation at the tip of the TEB creates a cleft,[25,26] while global branching patterns are governed by complex stochastic, mechanic, and cellular influences.[27–30]

Intravital microscopy (IVM) is an essential but challenging tool for revealing *in vivo* cell behavior. The light-scattering adipose tissue surrounding the mammary epithelium is particularly problematic and few studies have overcome this obstacle. IVM of normal development indicated that TEB cells are more migratory than duct cells[31] and that luminal TEB cells exhibit similar migratory behavior as seen in *ex vivo* studies.[32] We previously established lineage-specific *in vivo* imaging of cap cells over several hours,[33] where their ameboid migration and frequent entry into the TEB body confirmed a long-held prediction.[34,35] Modeling of cap cell fate showed that most cap cells that entered the body died within 4 h, after which 15% survived long term as the observable cap-in-body population.[33] However, a substantial gap remains between *ex vivo* and *in vivo* observations of TEB cell rearrangement during elongation. Moreover, there is no knowledge of whether HR$^+$ and HR$^-$ luminal cells contribute differently to ductal elongation.

To further explore the biological features of the HR$^+$ and HR$^-$ luminal lineages, we improved the identification of HR$^+$ LPs to enable a more robust study of their function and performed IVM on highly specific reporter mouse models. We demonstrate that proliferation *in vivo* is confined to LPs and not the HR$^+$ ML population. Of note, HR$^+$ LPs are highly proliferative but do not increase in number, suggesting asymmetric cell division, while HR$^-$ LPs are maintained as a larger pool, presumably poised for alveolar differentiation during pregnancy. At a molecular level, hormone-responsive progenitors share a proliferative signature with HR$^-$ alveolar-like precursors but display enhanced Wnt and Notch pathway signaling. IVM and AI analysis of cell-shape dynamics revealed that HR$^+$ progenitor cells migrate within TEBs and are vastly more motile in TEBs and alveolar buds than HR$^-$ progenitors and mature HR$^+$ luminal cells, suggesting that intercalation of HR$^+$ cells is the most significant driver of duct elongation. Collectively, this study reveals unexpected complexity and contrasting functions of the LP subsets during morphogenesis and maintenance of the mammary gland.

RESULTS

Refined resolution of mammary luminal cells based on CD14, Sca1, and PR expression

Luminal cells have been previously fractionated into CD49b$^+$Sca1$^+$ (HR$^+$ LP), CD49b$^+$Sca1$^-$ (HR$^-$ LP), and CD49b$^-$Sca1$^+$ (HR$^+$ ML) populations in the mouse mammary gland,[5] yielding variable separation of the LP subsets. To improve LP stratification, we performed flow cytometry using the alternative LP marker CD14[36,37] together with Sca1[38] on the luminal compartment (Lin$^-$CD29^{lo}CD24$^+$) and compared this with CD49b and Sca1 expression. The three populations were clearly discriminated in both C57BL/6 and FVB/NJ mice using CD14 and Sca1 (Figures 1A and S1A), whereas gating with CD49b and Sca1 resulted in ambiguous populations (Figures 1B and S1B). In C57BL/6 mice, ~4.4% HR$^+$ LPs, ~27% HR$^-$ LPs, and ~68.6% HR$^+$ ML cells were resolved (Figure S1C) while FVB/NJ mice were enriched in HR$^-$ LPs (CD14$^+$, 49%) compared to C57BL/6 mice (25%; t test = 0.001) but maintained a similar basal:luminal cell ratio (Figures S1C–S1E).

To directly assess the HR status of luminal cells, we crossed PR-cre[39] with Rosa26-LSL-eYFP mice (PR-cre/YFP) to mark PR$^+$ cells and their progeny. As expected, YFP was highly enriched in the luminal but not the basal (0.29%) compartment (Figures S1F–S1H). Stratification based on PR-cre/YFP and CD14 resulted in three robust cell populations where the majority of YFP$^+$ cells expressed Sca1 and the luminal marker Prom1[10,40] (Figures 1C, 1D, S1I, and S1J). As CD49b yielded poor separation, we focused on CD14 for subsequent studies (Figures 1E and S1K). High-resolution 3D confocal imaging of PR-cre/YFP

Figure 1. Co-expression of the progesterone receptor with CD14 and Sca-1 refines the mammary luminal compartment

(A and B) Luminal cell gating on mammary epithelial cells from adult C57BL/6 mice based on Sca1 and CD14 (A) or CD49b (B) expression.

(C) FACS gating strategy for luminal cells isolated from PR-cre/R26R-EYFP (9 weeks) mice.

(D and E) Histograms showing expression of Sca1 (D) and CD49b (E) across the CD14 and PR-cre/YFP-defined luminal populations.

(F) 3D confocal image of an adult gland showing co-expression of PR-cre/YFP and PR protein (whole mount, Figure S1L). Scale bar, 50 μm.

(G) Quantification of PR protein and PR-cre/YFP double-positive cells in 3D confocal images (n = 4 mice).

(H) Overlay histogram of PR protein expression in different subsets as determined by intracellular FACS analysis.

(I) Box and whisker plot (median, Tukey) showing the percentage of PR$^+$ cells within each population (n = 5).

(J) 3D colony-forming assays of luminal cells isolated from 9-week-old PR-cre/YFP mice (n = 3). Scale bar, 1 mm

(K) Quantification of 3D colony assays in (J) (n = 3).

(L) Population dynamics of the luminal compartment during mammary development ($n \geq 3$). dP, days pregnancy; Inv, involution. Significance was tested using two-tailed t tests: $^*p < 0.05$, $^{***}p < 0.001$. Data are shown as mean ± SEM (G and K) and mean percentage ±SEM (L).

glands confirmed that approximately 92% of YFP$^+$ cells expressed PR protein (Figures 1F, 1G, and S1L). Moreover, intracellular fluorescence-activated cell sorting (FACS) analysis demonstrated that PR protein expression was highly enriched in YFP$^+$ cells (Figure 1H), with no expression in HR$^-$ LPs (Figure 1I). The presence of some YFP$^+$ cells without PR protein may be due to the turnover of PR protein as cells enter the S phase of the cell cycle.[41]

Colony-forming cell assays in Matrigel indicated that progenitor activity was restricted to the CD14$^+$ populations (Figures 1J and 1K), supporting the designations of the CD14$^+$Sca1$^-$ subset comprising HR$^-$ LP cells, the CD14$^+$Sca1$^+$ subset comprising HR$^+$ LPs and the CD14$^-$ Sca1$^{+/-}$ subset representing HR$^+$ ML cells. We observed a large expansion of the HR$^-$ LP (CD14$^+$YFP$^-$ or CD14$^+$Sca1$^-$) population during pregnancy and 3 weeks post involution (Figure 1L). Conversely, the HR$^+$ LP (CD14$^+$YFP$^+$ or CD14$^+$Sca1$^+$) population was highest during puberty and declined in pregnancy before increasing post involution in a manner similar to HR$^+$ ML cells (CD14$^-$YFP$^+$ or CD14$^-$Sca1$^{+/-}$)(Figure 1L).

RNA-seq analysis implicates HR$^+$ LP cells as a luminal intermediate population

To molecularly characterize the luminal populations, we performed bulk RNA sequencing (RNA-seq) analysis of HR$^-$ LPs, HR$^+$ LPs, and HR$^+$ MLs derived from PR-cre/YFP-expressing mice. These data revealed that HR$^+$ LPs shared a striking similarity to both the HR$^+$ ML and HR$^-$ LP populations (Figure S2A). Analysis of luminal lineage markers in HR$^+$ LPs confirmed high expression of key HR$^+$ ML-associated genes, including the transcription factor *Tbx3*.[42] We also observed high levels of LP-associated genes (*Kit*, *Hey1*, and *Tspan8*) and intermediate expression of others (*Elf5*, *Itgb3*/*Cd61*, and *Cd55*)[5,37] within HR$^+$ LPs (Figure 2A), aligning with previous qPCR findings.[5] Despite similar transcriptomes, differentially expressed gene (DEG) analysis detected 1,023 DEGs between the HR$^+$ LP and ML populations, with 993 of these genes upregulated in HR$^+$ LPs. Of these, 984 (99%) were also higher in HR$^-$ LPs than in the HR$^+$ ML population, indicating that the differences between HR$^+$ LP and ML cells reflect a common progenitor cell program. DEG analysis of the two LP populations found 3,269 genes with 1,606 and 1,663 upregulated genes in the HR$^+$ and HR$^-$ populations, respectively. Definitive HR$^-$ lineage genes downregulated in HR$^+$ LP cells included *Lbp*, *Wfdc18*, *Cebpd*, *Cebpb*, *Nrg2*, *Lalba*, *Runx2*, and *Wnt5b*. All 1,606 DEGs for HR$^+$ LPs were also significantly enriched in HR$^+$ MLs compared to HR$^-$ LPs. These data are reflected in signature analyses, which indicate that HR$^+$ LP cells share a hormone-related signature with HR$^+$ ML cells as well as a progenitor signature with HR$^-$ LP cells (Figure 2B). Analysis of our single-cell RNA-seq (scRNA-seq) datasets[44] confirmed clear demarcation of luminal populations by *Cd14* and *Sca1* but promiscuous *Itga2*/*Cd49b* expression independent of flow cytometric gating (Figures S2B and S2C). Together, these findings indicate that HR$^+$ LP cells correspond to the luminal intermediate cluster identified in scRNA-seq studies.[44,45]

Upon examining differences in molecular pathways between the luminal populations using Gene Ontology (GO) analysis, HR$^+$ LPs were found to be enriched for Wnt, Notch, estrogen signaling, nucleosome, mammary development, and proliferation pathways (Figure 2C). Both the Wnt[46,47] and Notch[11,48,49] pathways play important roles in mammary gland development. HR$^+$ LP cells were enriched for *Wnt4*, *Wnt5a*, *Wnt11*, *Wnt7b*, and *Jag1* expression compared to HR$^-$ LPs, and for *Wnt5b*, *Wnt6*, *Wnt9a*, *Numb*, *Notch1*, *Hey1*, and *Hey2* expression compared to the HR$^+$ ML population (Figures 2D and 2E). The increased expression of cell-cycle genes including *Cdt1*, *Wee1*, *Pcna*, and *Rb1* in the CD14$^+$ populations is in keeping with their progenitor function (Figure 2F). Interestingly, cell-cycle genes up in both HR$^+$ and HR$^-$ LP cells were associated with positive regulation of mitosis and DNA replication, while those up in HR$^+$ cells were associated with microtubules, nuclear condensation, spindle organization, and chromosome segregation, consistent with differences observed in structural proteins and chromatin compactions.[37,43,50]

Both progenitor populations within the luminal compartment are actively cycling

To address the reported discrepancy between progenitor activity and *in vivo* cell proliferation, adult PR-cre/YFP mice (C57BL/6 background) were injected with 5-ethynyl 2′-deoxyuridine (EdU) and glands collected 2 h later to exclusively capture cells that had entered S phase (Figures 3A, 3B, and S2D). HR$^+$ LP cells were found to have the highest level of EdU staining in this strain, representing ~40% of EdU$^+$ epithelial cells. The presence of proliferating (Ki67$^+$) PR$^+$ and PR$^-$ cells in the luminal layer of adult mice was confirmed by 3D confocal microscopy (Figure 3C).

To label S-phase cells as well as those that had recently completed S phase,[20] adult FVB/NJ mice (a widely used strain in the field) were injected with EdU and collected 6 h later (Figure 3D). Although the majority of EdU$^+$ cells in this strain were HR$^-$ LPs (~55.5% EdU$^+$ cells compared to ~10.9% HR$^+$ LPs), HR$^+$ LP cells were found to be the most proliferative luminal subset and comprised 6.5% EdU$^+$ cells compared to 5.1% for HR$^-$ LPs) (Figures 3D, S2E, and S2F). To test the hypothesis that HR$^+$ LPs were undergoing differentiation and thus disappearing from the proliferative pool, we injected mice with EdU three times over a 24-h period to label cells in S phase and those that had moved through the S/G_2/M phases. It has been estimated that mammary epithelial cells take 20 h to complete each round of cell division.[20] A large increase in EdU labeling of HR$^+$ ML cells occurred only at the 26-h time point (24.3% of EdU$^+$ epithelial cells), revealing that HR$^+$ ML cells are not actively cycling but rather are the likely progeny of CD14$^+$ progenitors (Figure 3D). To further probe the cell-cycle status of the different cellular subsets, we used the fluorescent, ubiquitination-based cell-cycle indicator (Fucci2) mouse model[51] crossed with MMTV-cre(A)[52] (Figures 3E and S2G). In adult mice, the only populations that transitioned from G_1 to S phase (hCtd1$^+$ hGeminin$^+$ cells) were the basal Tspan8$^-$ and LP populations (Figures 3F, S2H, and S2I). Importantly, HR$^+$ ML cells were rarely positive for hCtd1 (DNA licensing and initiation of S phase) or hGeminin (G_2/M entry) (Figures 3F and S2I). Differentiation is thought to occur in late G_1 as cells enter S phase.[53,54] Most hGeminin$^+$ HR$^+$ LP cells occurred in late G_1/S phase while the majority of hGeminin$^+$ HR$^+$ ML cells were in late S/G_2/M, consistent with their transition from progenitor to mature cells. Together these data indicate that HR$^+$ ML cells are terminally differentiated and incapable of entering the cell cycle, supporting the hypothesis that HR$^+$ LP cells serve as precursors for the HR$^+$ ML population.

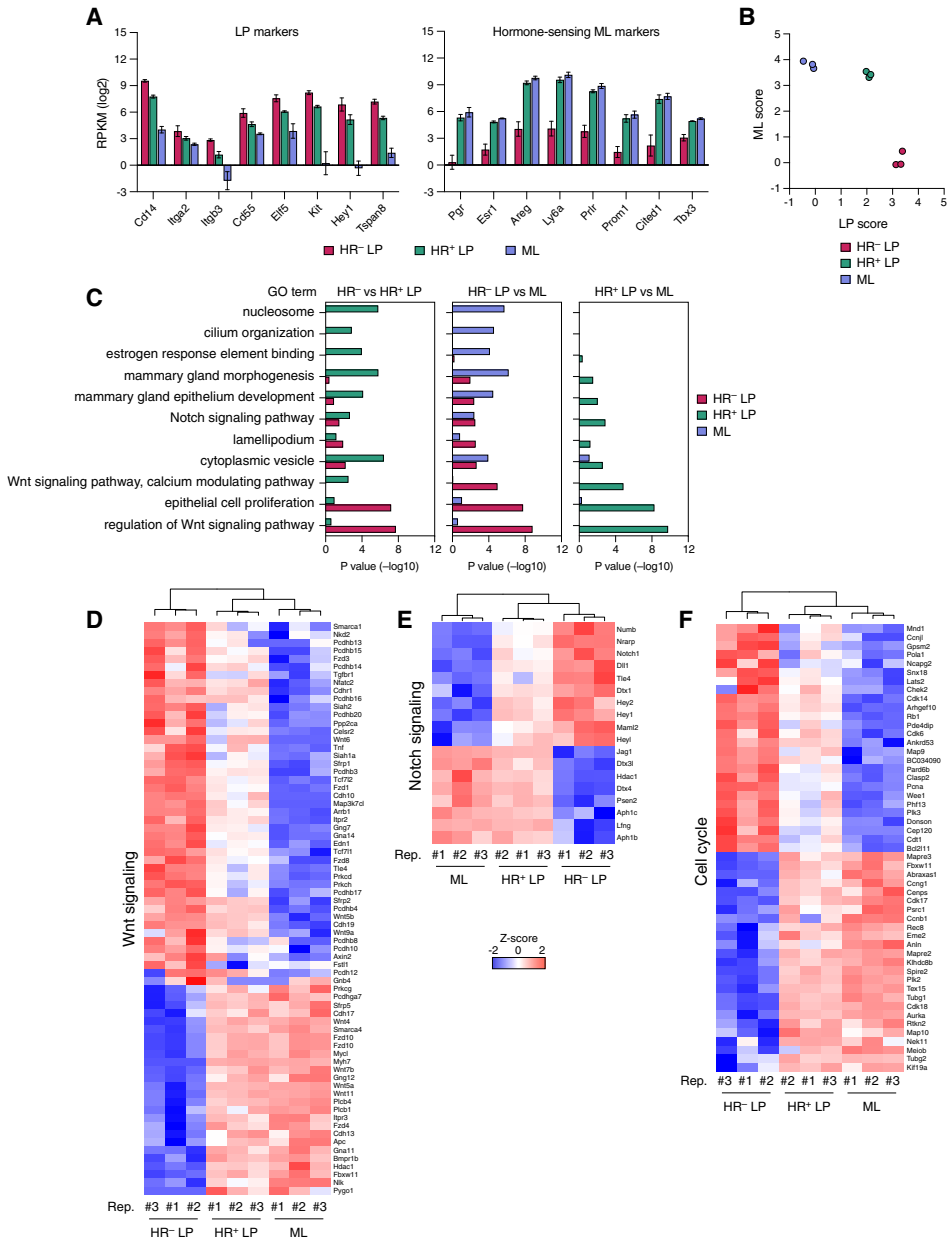


Figure 2. HR⁺ progenitors display bi-lineage luminal gene expression

(A) Expression of luminal progenitor (LP) and mature luminal (ML) markers across the populations defined by PR-cre/YFP and CD14 expression.

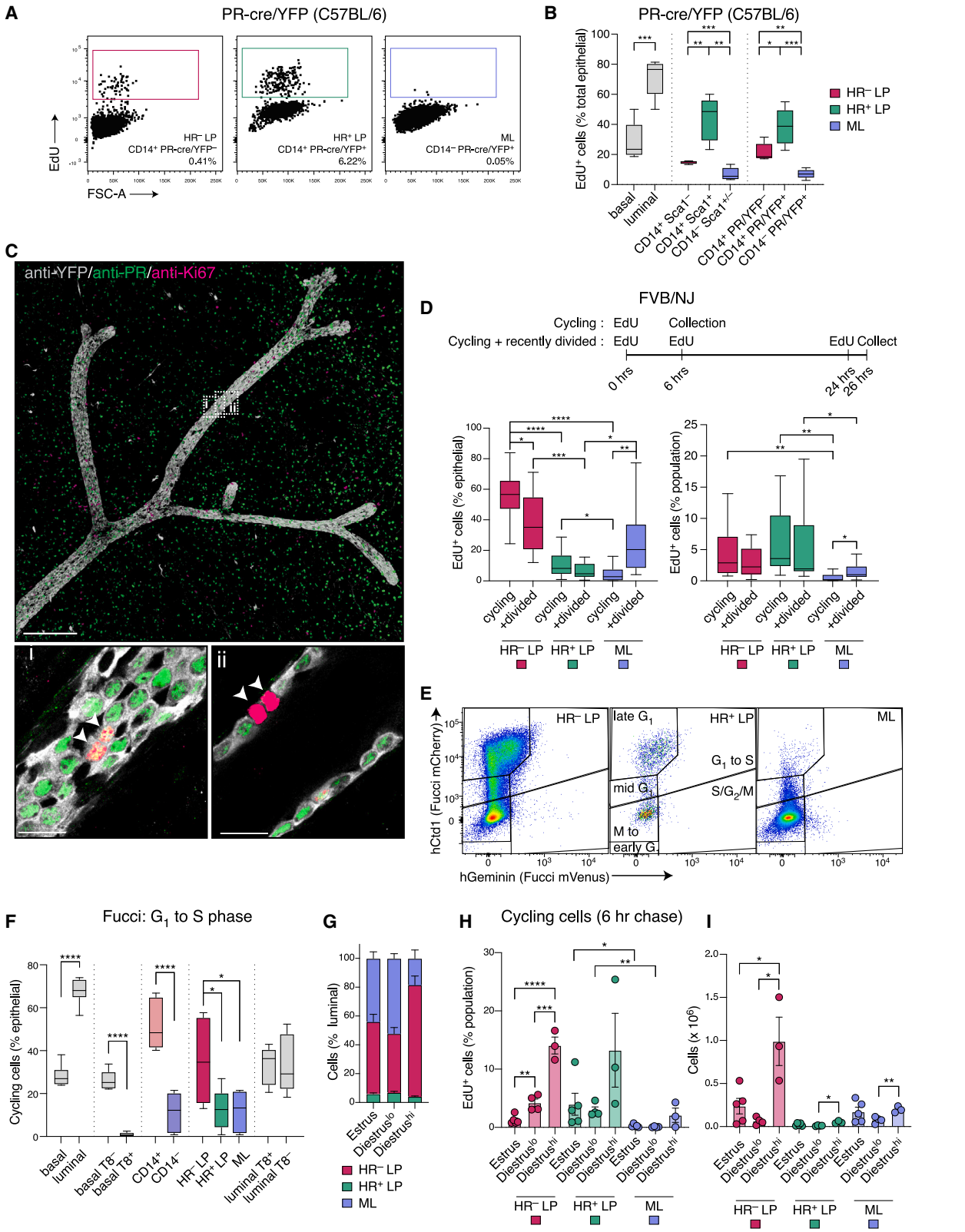
(B) Signature analysis of luminal cell transcriptome scores against the LP and ML gene signatures.[43]

(C) GO analysis highlighting pathways significantly enriched in pairwise comparisons of the different populations defined by CD14 and PR-cre/YFP.

(D–F) Expression of differentially expressed Wnt (D) and Notch (E) signaling genes and those associated with the cell cycle (F) in the CD14 PR-cre/YFP populations. Data are shown as mean ± SD (A).

Epithelial cells display differential population dynamics during the estrus cycle

To explore luminal proliferation over the estrus cycle, we euthanized mice and immediately staged their estrus phase. We used this approach since staging live mice through vaginal stimulation can induce a pseudopregnancy,[55] and this could influence luminal cell behavior. A previous report suggested that two patterns of proliferation exist in diestrus, with a high and low proliferative phase (diestrus^{hi} and diestrus^{lo}).[20] Consistent with this observation, of the seven mice analyzed in diestrus, three mice exhibited >5% EdU^{+} epithelial cells, while four had EdU levels comparable to those in estrus (Figure S3A). The previously reported increase in basal cells during diestrus[56] appears to associate with the diestrus^{hi} group, coinciding with a marked decrease in HR^{+} LP and ML cells (Figures 3G, S3B, and S3C). Interestingly, the majority of HR^{+} ML cells lost Sca1 expression



(legend on next page)

during high proliferative diestrus (Figures S3B and S3D). Most epithelial cells increased in proliferation from diestrus^{lo} to diestrus^{hi}, with HR$^+$ ML cells showing a trend ($p = 0.1$) (Figures 3H and S3E). Despite increases in proliferating HR$^+$ LP in diestrus^{hi}, HR$^-$ LPs made the largest contribution to the total proliferative pool (54% of EdU$^+$ epithelial cells) (Figures S3F and S3G).

To estimate cell proliferation based on total cell number, we determined the absolute number of proliferative cells within the luminal populations by FACS analysis (Figures 3I and S3H). On average there were 238 × 10^3 HR$^-$ LPs cells in estrus, increasing to 990 × 10^3 cells during diestrus^{hi}. There were 23 × 10^3 HR$^+$ LPs in estrus decreasing to 9.6 × 10^3 in diestrus^{lo} and 15 × 10^3 in diestrus^{hi}. Similarly, ML cells were lowest in diestrus^{lo} at 76 × 10^3 cells and increased to 172 × 10^3 in estrus and 198 × 10^3 cells in diestrus^{hi}. Overall, despite a similar proportion of proliferating cells in both LP populations, HR$^+$ LP cell numbers did not increase during diestrus, suggesting that many cells in this pool undergo rapid differentiation.

TEB HR$^+$ cells are proliferative and enriched for genes associated with cell migration

We further explored the localization of HR$^+$ cells in puberty when they are highest in proportion (HR$^+$ LP 5.9%, ML 66.6%) (Figure 1L). Comparison of the frequency of HR$^+$ cells in TEB units and subtending ducts microdissected from PR-cre/YFP mice revealed a doubling in the percentage of YFP$^+$ cells in TEBs (Figure 4A). To study HR$^-$ cells within these structures, we utilized the Elf5-rtTA/TetO-Cre/tdTomato model,[57] which expresses GFP in Elf5$^+$ LP cells that are primarily CD14$^+$ (Figure 4B). Lineage tracing based on *Elf5* almost exclusively gave rise to ER$^-$ cells after a 4-week chase (Figure S3I), with only one cell found to express ER by 3D imaging of several millimeters of tissue, thus validating the use of this model to mark HR$^-$ LP cells. Both HR$^+$ (PR-cre/YFP$^+$) and HR$^-$ (Elf5-GFP$^+$) cells exhibited high proliferation within TEBs versus ducts (Figures 4C, 4D, and S4A), consistent with TEBs being highly enriched for progenitor cells.[3] Interestingly, CD14 did not correlate with progenitor activity in the TEB (Figure S4B), contrasting with the adult gland.

To further characterize differences between HR$^+$ and HR$^-$ cells in TEBs versus ducts, we performed RNA-seq on cells isolated from microdissected structures from the Elf5-GFP and PR-cre/YFP models (Figures 4E and S4C). All TEB populations were found to be enriched for cell cycle genes (Figures 4F and S4D). Differential Wnt signaling was particularly evident in HR$^+$ cells, whereas enhanced Notch signaling was seen in HR$^-$ duct cells, in line with previous data.[11] Differential analysis identified 140 DEGs between HR$^+$ TEB cells versus HR$^-$ TEB and HR$^+$ duct populations, encompassing 77 upregulated and 63 downregulated genes (Figure 4G). Interestingly, CD61 (*Itgb3*), which enriches for LP cells in the adult gland,[4] poorly correlated with *Elf5* expression in TEBs (Figure 4B) but was highly expressed in TEB HR$^+$ cells. GO analysis revealed HR$^+$ TEB cell enrichment of migration, locomotion, chemotaxis, lamellipodia, and cell adhesion pathways including genes involved in focal adhesion kinase (FAK) and PI3K signaling (*Pik3cd*, *Ptk2*, *Itgb3*, *Cxcl13*, *Slit2*, *Sdc1*, and *Coro1c*) (Figure 4H). Collectively, these data indicate an enrichment of highly proliferative HR$^+$ cells within the TEB that are distinct from TEB HR$^-$ progenitor cells and are predicted to execute discrete functions associated with cell motility and adhesion.

Intravital imaging of single progenitor cell behavior

To explore differences in behavior between HR$^+$ and HR$^-$ progenitors in TEBs and predominantly mature cells in ducts, we performed subcellular resolution 3D intravital microscopy (IVM)[33] of these structures. To reveal single-cell behavior, we used lineage-specific Cre expression to recombine the Confetti allele[58] in basal cells (K5-rtTA/TetO-Cre/Confetti),[57] HR$^+$ cells (PR-cre/Confetti), and HR$^-$ LPs (Elf5-rtTA/TetO-Cre/Confetti)[57] (Figure 5A). Recombination was activated in the inducible K5- and Elf5-rtTA models with doxycycline at 4 weeks. IVM was performed at 5 weeks (for TEBs) and at 6 or 9 weeks (for ducts), imaging every 10 min for several hours using our modified skin-flap method with fluorescence-guided microdissection as previously described[33] (Figures 5B–5L and S4E–S5I). Only large TEBs with a multi-layered body were evaluated, and only the first 5 h of imaging were analyzed as increased cell death occurs beyond this point.[33]

Cap and HR$^+$ cells are migratory and motile within the TEB

As we previously documented, cap cells take on complex shapes with multidirectional protrusions as they migrate around

Figure 3. Luminal cell proliferation is restricted to progenitor cells *in vivo*

(A) FACS gating for EdU labeling of luminal cells in PR-cre/YFP mice.

(B) Percentage of EdU$^+$ mammary epithelial cells in PR-cre/YFP mice (9 weeks); mice were collected 2 h following EdU injection ($n = 5$). Cells were stratified by Sca1 and CD14 (HR$^-$ LP, HR$^+$ LP, and ML) or PR-cre/YFP and CD14 (P7, P8, P9, P10) expression.

(C) 3D confocal image of an adult (9 weeks) PR-cre/YFP mouse mammary gland. Whole mount (top) and optical sections (bottom) indicate (i) PR-cre/YFP$^+$PR$^+$Ki67$^+$ cells and (ii) PR-cre/YFP$^-$PR$^-$Ki67$^+$ cells, marked by arrows. Scale bars: whole mount, 200 μm; optical sections, 20 μm ($n = 3$ mice).

(D) EdU$^+$ cells, as a percentage of EdU$^+$ total epithelial cells (left) or of population (right). FVB/NJ mice (10–11 weeks old) were injected with EdU and collected 6 h later (6 h, $n = 14$) or injected twice with EdU on day 1 (6 h apart) and again on day 2 and collected 2 h later (26 h, $n = 13$). Luminal cells were stratified by Sca1 and CD14 as above.

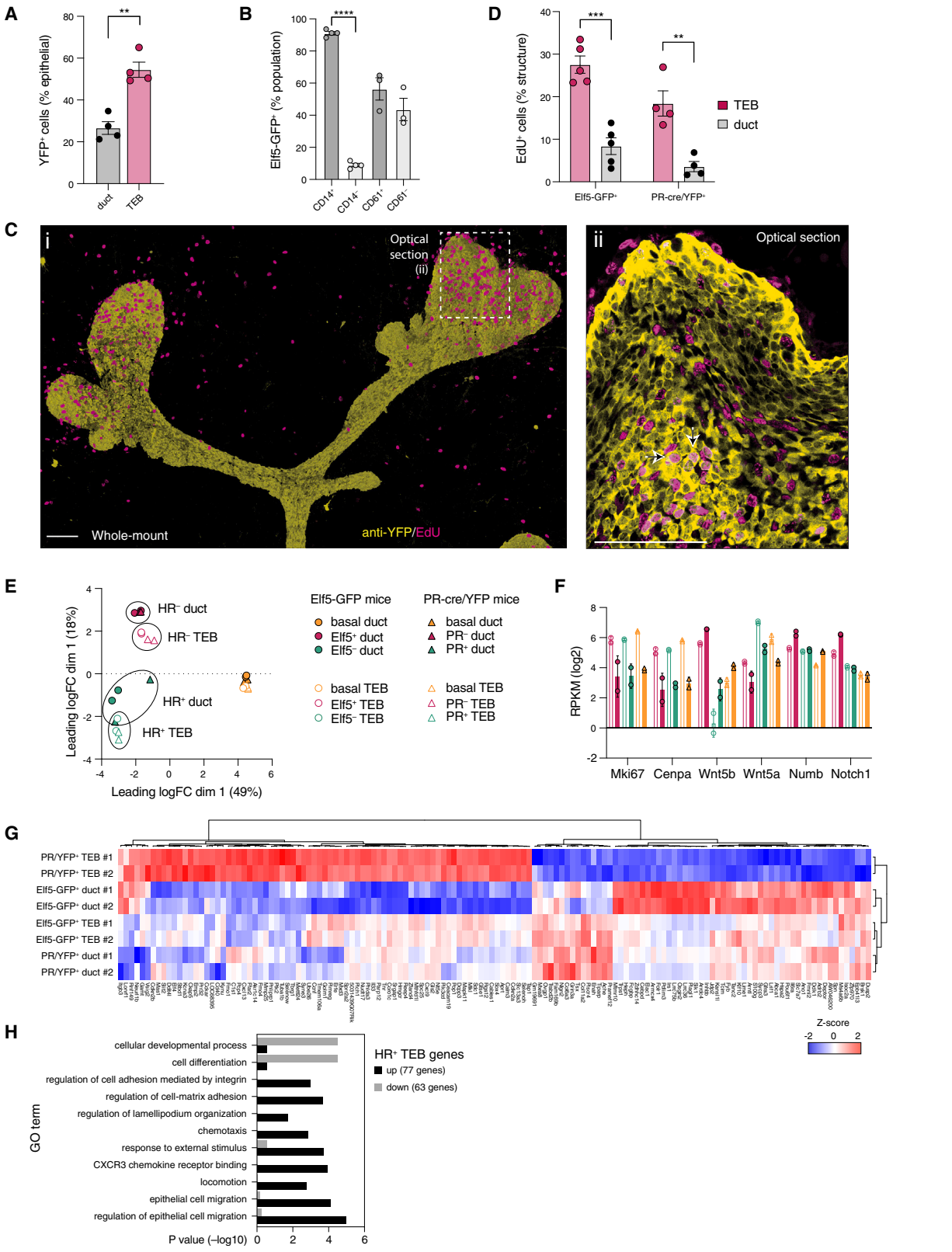
(E) FACS gating strategy for MMTV-cre/Fucci2 adult mice. The mid-to-late G_1 phase gates were merged for subsequent analyses.

(F) Cells marked as transitioning from G_1 to S phase in MMTV-cre/Fucci2 mice ($n = 6$). Shown is the percentage of cells within the epithelial population gated as G_1 to S.

(G) Percentage of luminal cell populations in estrus ($n = 5$), diestrus low (lo, $n = 4$), and diestrus high (hi, $n = 3$).

(H) EdU$^+$ cells as a percentage of population.

(I) Number of luminal cells in estrus, diestrus^{lo}, and diestrus^{hi} phases. Cell numbers were determined using FACS population percentages and counts for single-cell suspensions prior to sorting. Estrus ($n = 5$), diestrus low (lo, $n = 4$), and diestrus high (hi, $n = 3$). p values were determined using two-tailed t tests: $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$. Data are shown as median with Tukey whiskers (B, D, and F), mean percentage ±SEM (G), and mean ± SEM (H and I).



(legend on next page)

the cap layer in an ameboid fashion (Figures S4E and S4F).[33] In contrast, IVM of HR$^-$ LPs in TEBs revealed rounded and columnar shapes and low migration (Figures 5B, 5C, and S5A; Video S1). These cells sent short, dynamic extensions within the TEB and to the TEB edge (Figure S5B). Subsequent 3D super-resolution imaging of these same cells showed that the extensions protrude into and through the cap cell layer, presumably contacting the basement membrane (Figure S5C), as observed for luminal cells of unknown lineage by IVM,[32] in fixed tissue[59] and in organoids.[22]

Unexpectedly, IVM of HR$^+$ progenitors in TEBs revealed a striking motility and radial elongation, highlighting a contrast between HR$^-$ and HR$^+$ TEB cells (Figures 5D–5H, S5D, and S5E; Video S2). HR$^+$ cells migrated extensively throughout the TEB body, often extending and retracting long, thin projections as they moved both toward and away from the TEB surface (Figure 5G). They frequently sent dynamic protrusions toward the leading TEB edge while often having a distal point anchored in the inner TEB (Figure 5H). These findings suggest that only HR$^+$ cells correspond to the highly motile inner cells previously found to aid branching in *ex vivo* epithelial structures.[22,26] Progenitor activity of TEB HR$^+$ cells was confirmed by the observation of many cell divisions (approximately 0.6 times per hour) in stark contrast to ducts (none) (Figures S5J and S5K).

Motile HR$^+$ cells occur in TEBs and not ducts

To assess how the behavior of luminal cells changes throughout differentiation, we performed IVM on mature ducts in 6- to 9-week-old mice. HR$^-$ cells displayed a similar columnar phenotype to those in the TEB with no migration and short, dynamic protrusions at the basal surface (Figures 5I–5J, S5F, and S5G; Video S3). HR$^+$ cells showed distinct morphology from their TEB counterparts with a static conoid shape created by a small apical surface and a large basal surface that changed little over several hours (Figures 5K–5L, S5H, and S5I; Video S4).

Cap and HR$^+$ progenitors are more migratory than HR$^-$ progenitor cells in the TEB

To compare the migration of each population, cells were manually tracked in 3D over several hours, generating cell migration tracks and vectors of cell displacement (Figures S6A–S6C). Most cap cells migrated throughout the TEB and many to a large extent, especially near the TEB tip. Most HR$^+$ cells migrated to a smaller degree with some making large transitions, while most HR$^-$ LPs were stationary and only a few made significant movements. Migratory cap and HR$^+$ cells were displaced more than 50 μm over 10 h, over a third of the width of a typical TEB. To quantify these observations, we compared the average cell speed (distance traveled during 1 h) (Figure 5M) and average displacement (distance between start and end positions over 1-h periods) (Figure 5N) for each cell. Cap cells demonstrated the highest speed (median 8.1 μm/h) and displacement (5.1 μm/h). Speed was similar for HR$^+$ (6.3 μm/h) and HR$^-$ (5.6 μm/h) luminal cells, but HR$^+$ cells had greater displacement (4.6 vs. 3.1 μm/h), indicating more directional migration. Furthermore, there was a dramatic decrease in speed and displacement from TEB to duct for HR$^+$ cells (1.4- and 3.2-fold less, respectively) but not for HR$^-$ cells (speed, no change; displacement, 1.7-fold). The intercalation of motile TEB luminal cells into a stable bilayer is predicted to generate the force that elongates ducts during mammary gland morphogenesis.[23] HR$^+$ cells may provide more of this force since they had a much larger shift in dynamics than HR$^-$ cells following intercalation.

Creating a high-throughput pipeline for *in vivo* cell-shape analysis

To further exploit the 3D IVM dataset of mammary cell dynamics during morphogenesis, we developed an advanced AI analysis pipeline to measure the morphology of thousands of mammary cells over time, termed mammary cell-shape computation over time (MaSCOT-AI) (Figure 6A). The 27 3D videos were split into 350 2D videos spanning one cell layer to capture single-cell behavior. We then manually trained our MaSCOT-AI Cellpose 2.0 cell segmentation model[60] until it consistently outlined cell shapes accurately. This was integrated into a Trackmate-Cellpose script[61] that analyzed the shape of every cell over time in all 350 2D videos. A set of 5,619 cell tracks was finalized after quality control (tracks length >30 min, ensuring constant cell size and manual verification). The most independent cell morphology descriptors were aspect ratio (cell length/width ratio), for which increasing values reflect elongation, and solidity (cell area/convex hull area), for which decreasing values indicate growing complexity (Figures 6B, S6D, and S6E). These descriptors were partially redundant with circularity (Figures S6F–S6H).

Cap and HR$^+$ TEB cells have more dynamic shapes than mature HR$^+$ cells

To visualize the spread of cell behaviors and summarize the large datasets, we averaged the measurements over time for each cell and selected representative cell tracks at several percentiles. The 95th percentile of aspect ratio and fifth percentile for solidity demonstrated accurate segmentation and representative cell

Figure 4. HR$^+$ cells in TEBs are proliferative and express genes involved in cell motility

(A) Percentage of YFP$^+$ cells in the ducts and TEBs of PR-cre/YFP mice at 5 weeks ($n = 4$).

(B) FACS-based stratification of adult luminal cells from Elf5-GFP$^+$ mice ($n \geq 3$) based on CD14 or CD61 expression.

(C) 3D confocal image of a PR-cre/YFP mammary gland during puberty (5 weeks): whole mount (i) and optical section (ii), stained for YFP and EdU ($n = 4$ mice). Arrows indicate double-positive cells within the TEB. Scale bars, 100 μm.

(D) Percentage of EdU$^+$ cells, 2 h after EdU injection, in TEBs and ducts of Elf5-GFP and PR-cre/YFP mice by 3D confocal imaging and image quantification ($n \geq 4$).

(E) Multidimensional scaling (MDS) plot of RNA-seq data from FACS-isolated mammary epithelial cells derived from Elf5-GFP and PR-cre/YFP pubertal mice ($n = 2$).

(F) Expression of cell-cycle, Wnt, and Notch pathway genes in TEB and duct epithelial cells.

(G) Expression of 140 genes, either up- (77) or down-regulated (63) between TEB PR-cre/YFP$^+$ cells compared to TEB Elf5-GFP$^+$ cells and duct PR-cre/YFP$^+$ cells.

(H) GOs for the 140 DEGs shown in (H). p values were determined using two-tailed t tests: $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$. Data are shown as mean percentage ±SEM (A, B, and D) and mean ± range (G).

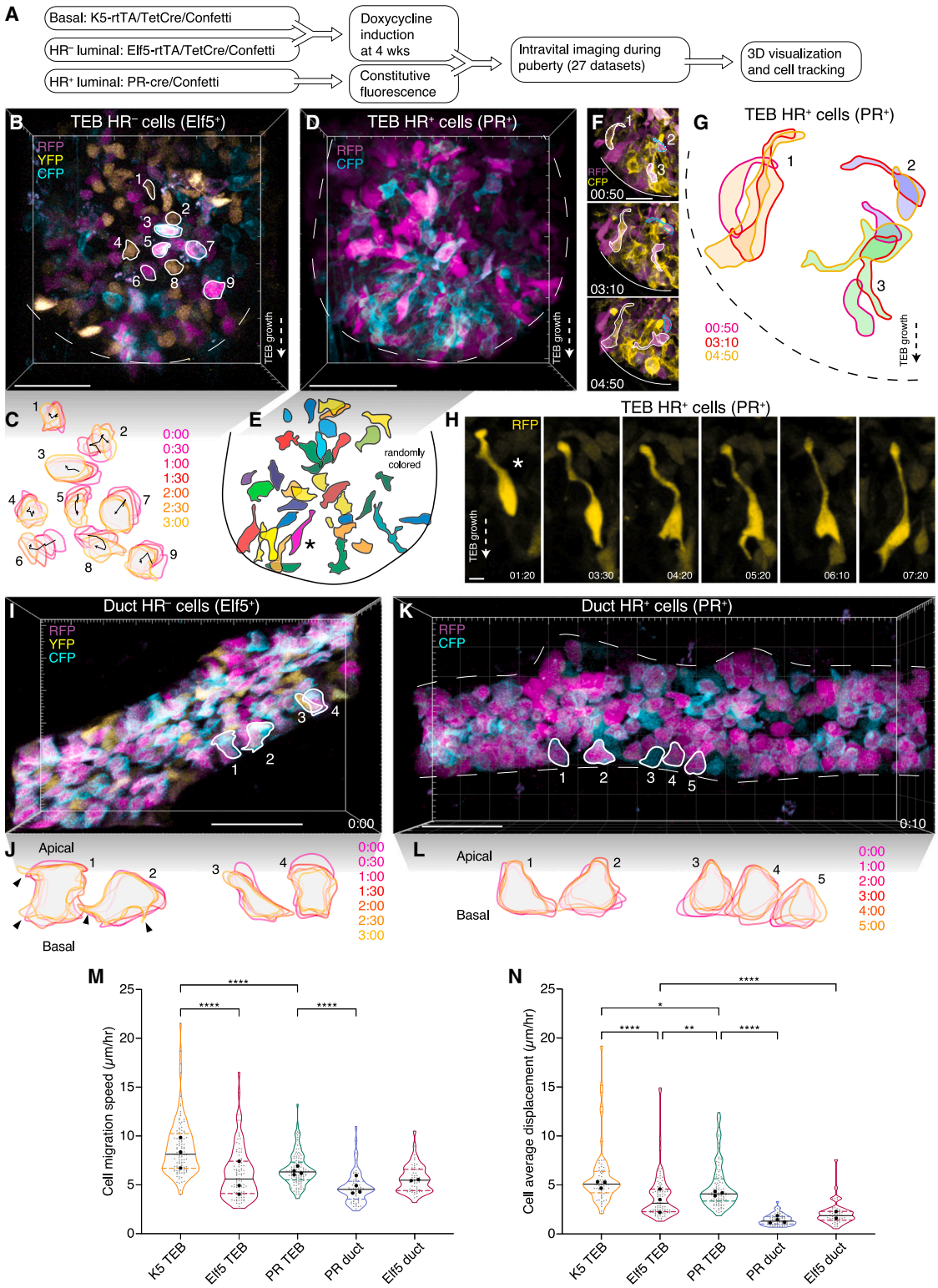


Figure 5. 3D single-cell intravital microscopy of morphogenesis reveals unique migratory features of HR⁺ LPs within the TEB

(A) Flow diagram of multicolor fluorescent labeling of mammary epithelial lineages and intravital microscopy.

(B) Snapshot from a 3D intravital video of an Elf5-rtTA/Tet-Cre/Confetti TEB (HR⁻ LP) at 5 weeks after doxycycline induction at 4 weeks ($n = 3$). Scale bar, 50 μm.

(C) Manually outlined cells over time as numbered from (B). Time indicated in hours:minutes and indicated by color, corresponding with cell outline color.

(legend continued on next page)

shapes over time for each population (Figures 6C and 6D). Reduction of the data to 2D resulted in a slight loss of complexity compared to actual cell shapes in 3D (Figures 5B–5L); however, this trade-off was necessary to analyze many cells. We plotted the aspect ratio and solidity over time for two representative cells for each percentile with the tracks smoothed by a running average (Figures 6E and 6F). Cap and HR$^+$ TEB cells displayed large peaks in elongation and dips in solidity that were absent in duct HR$^+$ cells. The maximum measurements per cell revealed the extent of elongation and complexity over whole populations (Figures 6G and 6H) and variation of these measurements over time reflected the degree of cell-shape dynamics (Figures S6I and S6J). HR$^+$ TEB cells were similar to cap cells in the extent and variation of both elongation and complexity. Both populations were significantly higher than HR$^-$ TEB cells for these parameters. HR$^-$ cells showed a minor increase in elongation from the TEB to duct but no change in solidity or variation of either measurement. By contrast, HR$^+$ TEB and duct cells were completely opposed, representing the highest and lowest populations for each measurement including variation. This reinforces the dramatic transition that these cells make from motile to sessile during duct formation. Collectively, MaSCOT-AI is a powerful tool for interrogating *in vivo* cell shapes during mammary gland development, with further opportunities for interrogating cell behavior during other developmental stages and oncogenesis. We established that HR$^+$ cells are unique in their transition from highly elongated and dynamic cells to sessile conoid cells during differentiation.

Cap and HR$^+$ TEB cells have more frequent peaks in complexity and elongation

The high quality of our single-cell morphology data allowed us to interrogate the frequency of motile behavior over time. We counted the total number of small, short variations (a single increase then decrease in complexity or elongation >1%) and large, long variations (two increases then two decreases in complexity or elongation, each >1%). These counts were divided by the cumulative time that all cells were tracked (minus start and end periods) to generate a frequency for each population. Plotting short vs. long variations for both solidity and aspect ratio revealed that all TEB measurements except Elf5$^+$ cell solidity had more frequent large, long variations than small, short variations (Figure S6K). HR$^-$ ductal cells had frequent short oscillations in both complexity and elongation compared to static HR$^+$ ductal cells. Comparing the frequency of large, long complexity and elongation peaks and overlaying the mean cell maximums for these measurements created a "motility plot" that coalesced much of the information generated by MaSCOT-AI. This comparison demonstrated that cap and HR$^+$ TEB cells are the most frequently dynamic for elongation and complexity, respectively, and that HR$^+$ cells are the highest on average for both measurements (Figure 6I).

Progesterone and pregnancy induce HR$^+$ cell motility in alveolar buds

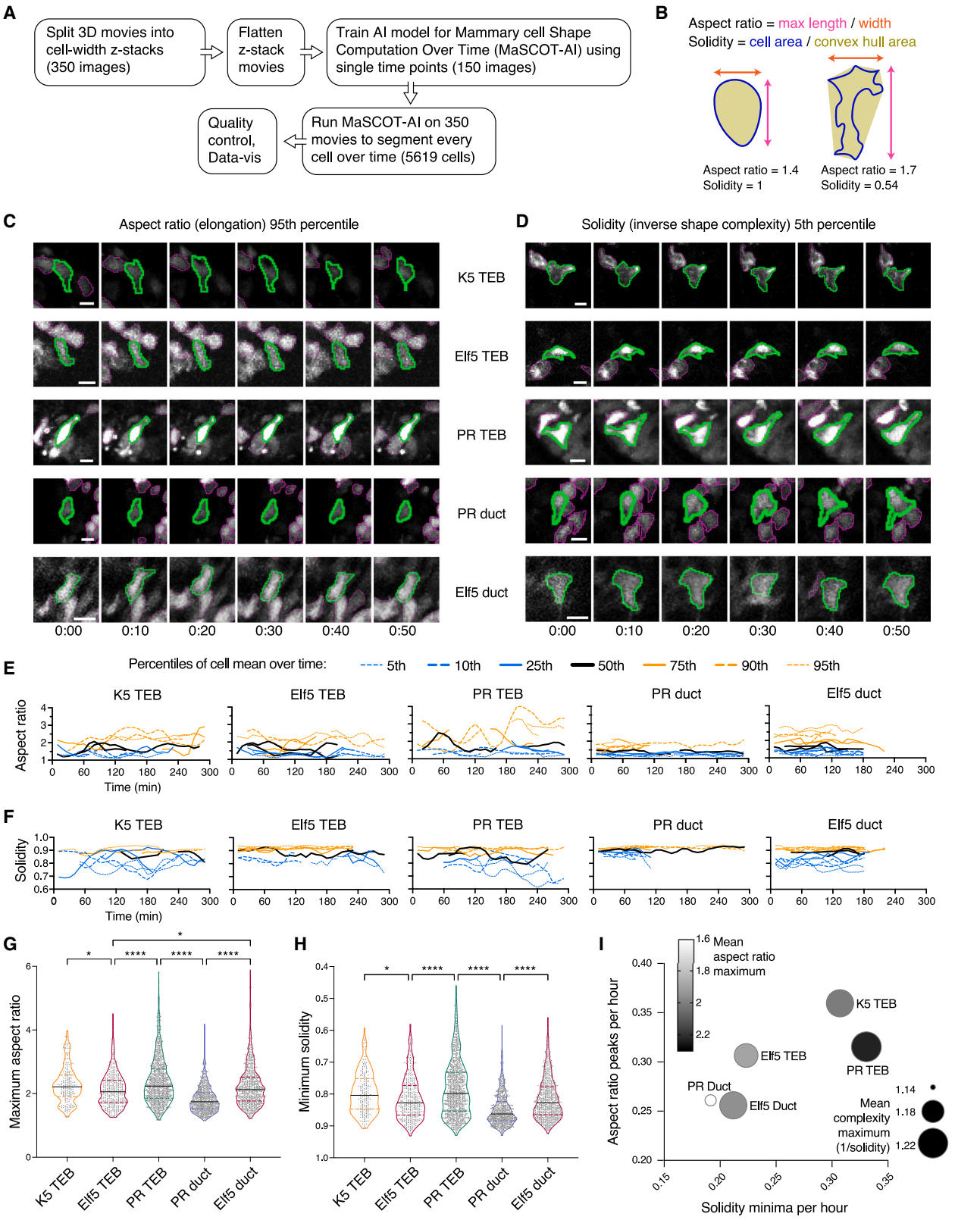
As HR$^+$ cell motility and migration were associated with morphogenesis in puberty, we hypothesized that this also occurs in pregnancy. Indeed, IVM of PR-cre/Confetti mice at 6.5 days of pregnancy (dP) followed by 3D visualization demonstrated the induction of complex behavior (Figures 7A–7E; Video S5). The most highly elongated and complex shapes were seen in the rapidly expanding alveolar buds. Cells did not migrate in the buds but rather extended dynamic filopodia in the direction of growth. Short, dynamic projections on the basal side were also seen in the ducts in pregnancy. To study the effects of a potent hormonal stimulus in isolation, we performed IVM on virgin adult mice treated with the synthetic progestin, medroxyprogesterone acetate (MPA). MPA treatment for 6 days induced dynamic behavior of HR$^+$ cells within ducts compared to the static cells observed in vehicle-treated mice (Figures 7F–7K, S7A, and S7B; Videos S6 and S7).

We next performed high-throughput MaSCOT-AI analysis on HR$^+$ cell IVM during pregnancy or upon MPA treatment, from which representative cells are shown (Figures S7C and S7D). Plotting smoothed measurements of solidity (Figure 7L) and aspect ratio (Figure S7E) clearly demonstrated an increase in complexity and elongation in pregnancy and with MPA treatment. Comparing the average cell maxima/minima (Figures S7F and S7G) revealed global increases in HR$^+$ cell elongation and complexity. The variation in these measurements over time was also increased, demonstrating enhanced cell dynamics (Figures S7H and S7I). We then investigated where these conditions fell in our motility frequency analysis (Figure 6I) and found that they displayed motile behavior at a rate intermediate between HR$^-$ and HR$^+$ TEB cells and at a similar magnitude to HR$^-$ TEB cells (Figure 7M). Finally, we separated HR$^+$ cells in ducts versus alveolar buds and confirmed that, although elongation was unchanged (Figure S7J), cells in the developing buds displayed greater complexity (Figure 7N).

DISCUSSION

Around 99% of breast cancers have their cellular origin in the luminal lineage.[62] Thus, a comprehensive understanding of the

(D) 3D intravital snapshot of a PR-cre/Confetti (HR$^+$ LP, $n = 4$) TEB at 5 weeks. Scale bar, 50 μm.
(E) Manually outlined cells from (D), randomly colored to facilitate differentiation. Star indicates cell highlighted in (H).
(F) Zoom of cells in (D).
(G) Manually outlined cells from (F). Cell outline indicates the time point, cell fill color indicates individual cells.
(H) Zoom of the radially oriented HR$^+$ LP in the PR-cre/Confetti TEB IVM video starred in (E). Scale bar, 7 μm.
(I) 3D intravital snapshot of an Elf5-rtTA/Tet-Cre/Confetti duct at 6 weeks after doxycycline induction at 4 weeks ($n = 2$). Scale bar, 50 μm.
(J) Manually outlined cells over time as numbered from (I). Arrowheads indicate dynamic protrusions. Overlapping cells were separated.
(K) 3D intravital snapshot of a PR-cre/Confetti duct at 6 weeks ($n = 4$). Scale bar, 50 μm.
(L) Manually outlined cells over time as numbered from (K).
(M) The average distance migrated by cells per hour. Gray dots are individual cells and dataset medians are black dots.
(N) Average cell displacement per hour. Significance was determined by one-way ANOVA between selected pairs with Sidak correction for multiple comparisons: $^*p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$. Data are shown as snapshots measured in hours:minutes (B–L) and median with quartiles (M and N).



(legend on next page)

HR^+ and HR^- luminal subpopulations is paramount for understanding both normal development and the etiology of the different subtypes of breast cancer. Here we show that both HR^- and HR^+ LP cells are highly proliferative *in vivo*, whereas HR^+ ML cells are non-proliferative, consistent with the low clonogenic activity of ML cells in mouse and human tissue.[4,12] HR^+ LPs constitute a small and variable progenitor fraction dependent on mouse strain, express lower levels of *Esr1* and *Pgr* mRNA compared to the ML population, and appear to rapidly differentiate as their numbers do not increase during the cell cycle. This small population of HR^+ LP cells is likely to account for the rare proliferative ER/PR^+ cells observed *in vivo*. $CD14^-Sca1^{+/-}$ ML cells only demonstrated EdU positivity when cells had completed S phase, indicating that HR^+ ML cells are the progeny of recently divided cells, most likely the HR^+ LP population. There are three potential factors that account for our findings versus previous studies, which concluded ML cells are the most proliferative[19,20]: firstly, CD14 together with Sca1 provided more robust fractionation of the HR^+ LP population than afforded by CD49b and Sca1, which is important given the small size of this subset. Secondly, we euthanized mice prior to estrus staging since previous studies had shown that vaginal stimulation often results in a pseudopregnancy.[55,63,64] Thirdly, HR^+ LP cells were not analyzed in one study, which addressed only $CD49b^+Sca1^-$ and $CD49b^-Sca1^+$ cells.[20] These factors confound interpretation of data and have led to the incorrect conclusion that HR^+ ML cells are highly proliferative *in vivo*.

The mammary epithelium exhibits dynamic and cyclical proliferation during the estrus cycle. Progesterone levels peak during diestrus, yielding the highest numbers of epithelial cells in the murine mammary gland.[5,56] Diestrus, however, does not always lead to increased proliferation and can be associated with either a low or high proliferative state.[20] Using the Fucci2 cell-cycle reporter, Shehata et al. demonstrated that, in diestrus mice showing low-level proliferation, <1% of $CD49b^+Sca1^-$ and $CD49b^-Sca1^+$ cells were in the late G_1 to M phase, while, in high-proliferation diestrus mice, cells in these phases increased to ~20% and ~10%, respectively.[20] We also observed a dichotomy of epithelial proliferation in diestrus, where it ranged from ~2% to ~9% between mice. Analysis of the subsets defined by CD14 and Sca1 during the estrus cycle showed that cell numbers increased for basal, total luminal, and HR^- LP populations in diestrus compared to estrus. Despite increased proliferation, HR^+ LP cells did not increase in number between estrus- and diestrus-stage glands, indicative of cell loss from this pool via differentiation.

Unexpectedly, HR^+ and HR^- LPs were found to have distinct behavioral functions. HR^+ cells in TEBs and developing alveoli are elongated and motile, especially within TEBs where they are migratory (Figure 7O). Only the highly migratory cap cells of the TEB had a higher rate of displacement than HR^+ TEB cells. These two populations had a comparable frequency of motile behavior and HR^+ cells had the greatest extent of elongation and complexity. By comparison, HR^- progenitor cells were mostly columnar in TEBs and ducts and had low motility in both structures. Their behavior was typified by short, dynamic projections into the basal layer. This division of behavioral function in the luminal compartment was unanticipated. Mammary gland morphogenesis in puberty is an important model of tissue formation but much remains to be determined, including how *ex vivo* and modeling data relate to *in vivo* processes. Our *in vivo* data uncover a previously unappreciated aspect of the morphogenesis of mammary epithelial tubes in the compartmentalized roles of luminal cells.

IVM produces rich resources of cell behavior data but raises the challenge of drawing robust conclusions that take full advantage of the depth of information. Advances in AI image analysis, such as the Cellpose[60] program used here, are enabling the field to quantitatively assess cell dynamics over time. Training an AI model (MaSCOT-AI) to accurately segment mammary epithelial cells in our dataset made it possible to run analyses over thousands of cells and draw key conclusions on lineage-specific behavior. This yielded not only comparative measurements of cell populations but also insights into the frequency of cell-shape changes. These measurements reflect the rate and timing of cell elongation and adoption of complex forms that are a direct reflection of their structural and developmental roles, culminating in a motility plot that combines the frequency and magnitude of cell elongation and complexity (Figure 7M). This provides a map of cell behavior that, together with measurements of speed and displacement, provides an unprecedented description of *in vivo* mammary cell behavior. In the future, data from IVM of individual epithelial populations at different developmental stages and disease states could be overlayed onto our reference dataset to reveal the effect of perturbations on mammary cell behavior.

The most striking observation from our IVM studies was the stark difference in HR^+ cell behavior between TEBs and ducts. An earlier prediction was that luminal cell radial intercalation into the bilayer generates the force for ductal elongation.[23] The observation that HR^- cells have similar behavior in TEBs and ducts suggests that they play a more passive role in duct elongation. HR^+

Figure 6. AI analysis pipeline indicates that motile HR^+ progenitors differentiate into sessile mature cells

(A) Flow diagram of the MaSCOT-AI pipeline for analyzing the dynamics of cell shape over time in a high-throughput manner.

(B) Definition of the cell-shape descriptors, aspect ratio, and solidity, with example cells and associated measurements.

(C and D) IVM timelapses with MaSCOT-AI segmentations of cells representative of the upper range of population elongation and complexity (10-min steps). Scale bars, 10 μm.

(E and F) Visualization of the range and dynamics over time for (E) aspect ratio and (F) solidity. Two representative cell tracks for each percentile are plotted with smoothing by a running average across three time points.

(G and H) The maximum aspect ratio (G) and minimum solidity (H) of each cell track. Gray dots are individual cells.

(I) Multiple variables graph plotting the cumulative frequency of large, long aspect ratio peaks, and solidity minima (>40 min). Dot size increases with the mean complexity maximum per cell (1/solidity), and shade darkens with increasing mean aspect ratio maximum per cell. Significance was determined by one-way ANOVA between selected pairs with Sidak correction for multiple comparisons: $^*p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$. Data are shown as median with quartiles (G and H).

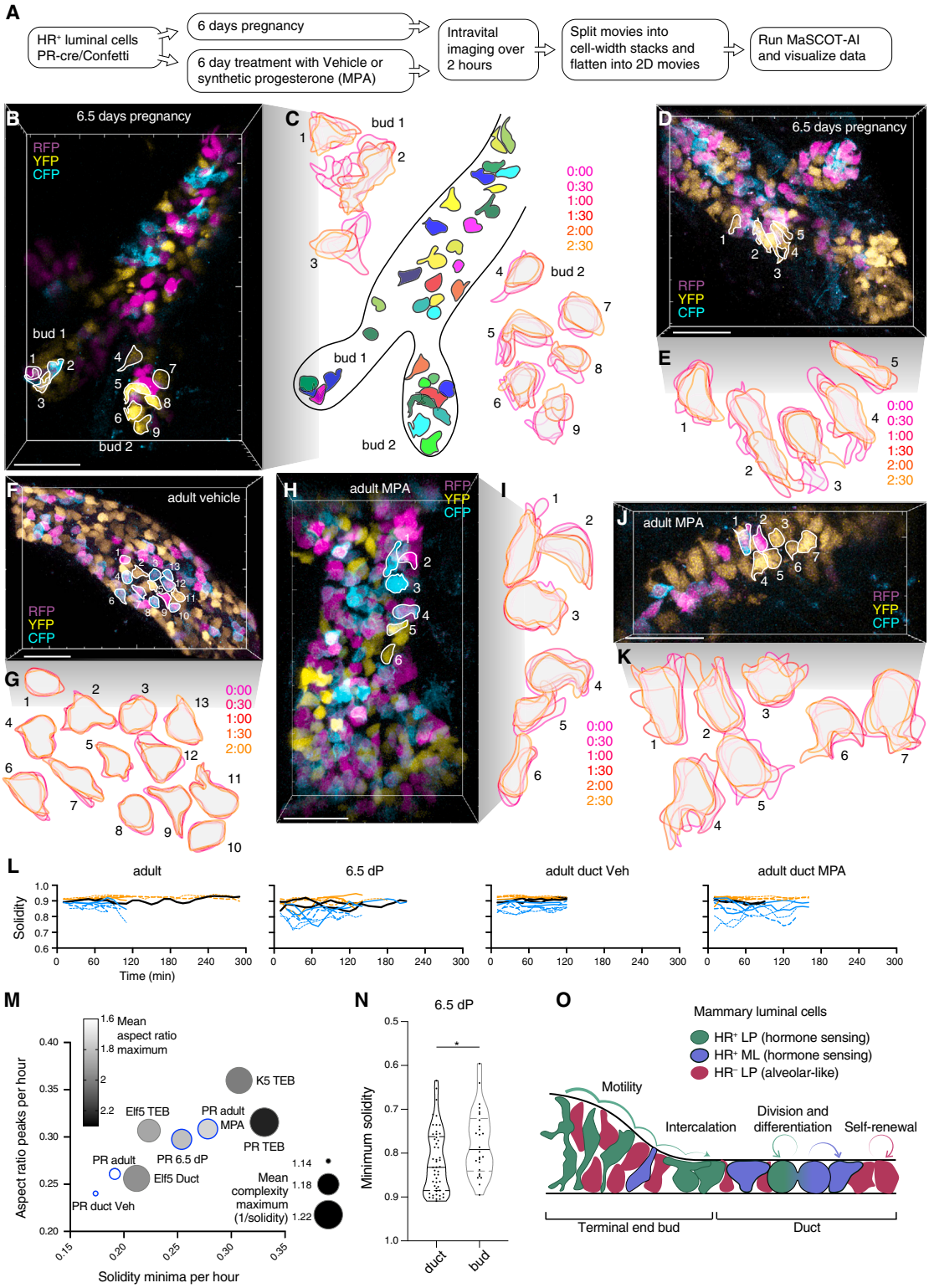


Figure 7. Pregnancy hormones induce dynamic HR+ cell behavior in developing alveolar buds

(A) Flow diagram of multicolor fluorescent labeling of HR+ cells and intravital microscopy during pregnancy and hormone treatment for MaSCOT-AI cell dynamics analysis.

(B and D) 3D intravital snapshots of PR-cre/Confetti epithelium at 6.5 days of pregnancy (6.5 dP) (n = 5). Scale bars, 50 μm.

(legend continued on next page)

luminal cells, on the other hand, have emerged as a major potential driving force in this process as they transition from a highly elaborate and loose structure in TEBs into the rigid layer of conoid cells within ducts (Figure 7O). Furthermore, our findings in pregnancy indicate that HR^+ cells respond to hormones not only as a signaling hub but also as drivers of the physical shaping of alveolar buds. These fascinating functions have implications for breast cancer, where the role of HR^+ cell behavior during pathogenic tissue invasion should be explored. Gaining an understanding of the linkages between hormone signaling and dynamic cell behavior that drive tissue invasion will be crucial for understanding the role of these cells in normal development and disease.

Limitations of the study

Although we have refined markers to enable the resolution of HR^+ LP, HR^- LP, and HR^+ ML cells by flow cytometry and their cell-cycle status, it remains to be formally proved that HR^+ LP cells serve as precursors for the ML population. Moreover, HR^+ LPs do not expand in the mammary gland despite being the most proliferative population, implying that they undergo rapid asymmetric cell division and differentiation, although this is yet to be determined. IVM has proved essential in describing cellular behavior within the TEB; however, this technique can only be applied over several hours. TEB branching likely occurs on a timescale beyond this, which prevented observation of how *in vivo* lineage-specific dynamics contribute to TEB branching. Lineage-specific imaging of single-cell behavior in *ex vivo* organoids could provide information on compartmentalized roles during branching that would complement our work *in vivo*.

RESOURCE AVAILABILITY

Lead contact

Further information regarding the resources and reagents used in this study should be directed to Jane Visvader (visvader@wehi.edu.au).

Materials availability

The study did not generate any new unique reagents.

Data and code availability

- Raw experimental data reported in this paper will be shared by the lead contact upon request.
- The accession number for the RNA-seq datasets is GSE268949.
- The original codes for the MaSCOT-AI analysis are publicly available at GitHub: https://github.com/cadaws/MaSCOT-AI.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are grateful to A. Rios and B. Pal for discussions. We thank the Imaging, Bioservices, FACS, Histology, and Genomics facilities at WEHI. This work was supported by the Australian National Health and Medical Research Council (NHMRC), NHMRC IRIISS, the Victorian State Government Operational Infrastructure Support, the Australian Cancer Research and Ian Potter Foundations, and the Breast Cancer Research Foundation (US). M.J.G.M. was supported by a VCA early-career fellowship (ECRF19011); Y.C. was supported by an MRFF Investigator Grant (#1176199); and G.J.L., G.K.S., J.E.V., and C.A.D. were supported by NHMRC Investigator Grants (G.J.L., #1175960; G.K.S., #2025645; J.E.V., #1037230 and 1102742; C.A.D., #2018105).

AUTHOR CONTRIBUTIONS

C.A.D., M.J.G.M., and J.E.V. conceptualized the study. C.A.D., M.J.G.M., B.D.C., F.V., L.P., and X.S. performed biological experiments. C.A.D. performed AI image analysis. R.K.H.Y. performed EdU image quantification. Y.C. and G.K.S. performed bioinformatic analyses. F.C.J. provided mouse expertise. C.A.D., M.J.G.M., J.E.V., and G.J.L. interpreted data. C.A.D., M.J.G.M., and J.E.V. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Mouse models: Induction and treatment
  - Isolation of mammary epithelial cells
  - Mammary epithelial colony forming assays
  - EdU cell labeling and estrus staging
  - Intracellular FACS analysis
  - RNA-sequencing
  - Differential gene expression and gene ontology analysis
  - 3D confocal imaging of whole-mount tissue
  - Intravital imaging

(C and E) Manually outlined cells in alveolar buds over time as numbered from (B) and (D) (overlapping cells were separated), and (C) cell outlines from the whole structure in (B), randomly colored to facilitate differentiation. Time indicated in hours:minutes.
(F) 3D intravital snapshot of a duct in a 9-week-old PR-cre/Confetti mouse treated with vehicle for 5 days before imaging on the sixth day ($n = 2$). Scale bar, 50 μm.
(G) Manually outlined cells over time as numbered from (F).
(H and J) 3D intravital snapshot of ducts from 9-week-old PR-cre/Confetti mice treated with medroxyprogesterone acetate (MPA) for 5 days before imaging on the sixth day ($n = 4$). Scale bars, 50 μm.
(I and K) Manually outlined cells over time as numbered from (H and J).
(L) Visualization of the range and dynamics of cell solidity over time measured by MaSCOT-AI; percentiles shown as in Figures 6E and 6F.
(M) Multiple variables graph from Figure 6I with added datapoints for PR 6.5 dP, PR adult vehicle treatment, and PR adult MPA treatment.
(N) Minimum solidity of PR-cre/Confetti cells at 6 dP from Figure S7F separated into cells in ducts and buds.
(O) Summary diagram: HR^- LPs maintain similar columnar morphology throughout development with short projections, while HR^+ LPs uniquely display extensive motility and migration during collective cell migration in puberty, and motility during alveolar bud formation downstream of progesterone signaling. HR^- LPs self-renew, whereas HR^+ LP cells appear to rapidly differentiate upon cell division, so that tracked dividing cells only accumulate within HR^- LP and HR^+ ML populations. Significance was determined by an unpaired two-tailed t test (N): $^*p < 0.05$. Data are shown as snapshots in hours:minutes (B–K) and median with quartiles (N).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.115073.

Received: June 22, 2024
Revised: October 18, 2024
Accepted: November 25, 2024
Published: December 17, 2024

REFERENCES

1. Brisken, C., and O'Malley, B. (2010). Hormone Action in the Mammary Gland. Cold Spring Harbor Perspect. Biol. *2*, a003178. https://doi.org/10.1101/cshperspect.a003178.
2. Macias, H., and Hinck, L. (2012). Mammary Gland Development. Wiley Interdiscip. Rev. Dev. Biol. *1*, 533–557. https://doi.org/10.1002/wdev.35.
3. Paine, I.S., and Lewis, M.T. (2017). The Terminal End Bud: the Little Engine that Could. J. Mammary Gland Biol. Neoplasia *22*, 93–108. https://doi.org/10.1007/s10911-017-9372-0.
4. Asselin-Labat, M.L., Sutherland, K.D., Barker, H., Thomas, R., Shackleton, M., Forrest, N.C., Hartley, L., Robb, L., Grosveld, F.G., van der Wees, J., et al. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat. Cell Biol. *9*, 201–209. https://doi.org/10.1038/ncb1530.
5. Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I.A., Avril, S., Prater, M., Eirew, P., Caldas, C., Watson, C.J., and Stingl, J. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. *14*, R134. https://doi.org/10.1186/bcr3334.
6. Tharmapalan, P., Mahendralingam, M., Berman, H.K., and Khokha, R. (2019). Mammary stem cells and progenitors: targeting the roots of breast cancer for prevention. EMBO J. *38*, e100852. https://doi.org/10.15252/embj.2018100852.
7. Fu, N.Y., Nolan, E., Lindeman, G.J., and Visvader, J.E. (2020). Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. Physiol. Rev. *100*, 489–523. https://doi.org/10.1152/physrev.00040.2018.
8. Asselin-Labat, M.-L., Sutherland, K.D., Vaillant, F., Gyorki, D.E., Wu, D., Holroyd, S., Breslin, K., Ward, T., Shi, W., Bath, M.L., et al. (2011). Gata-3 Negatively Regulates the Tumor-Initiating Capacity of Mammary Luminal Progenitor Cells and Targets the Putative Tumor Suppressor Caspase-14. Mol. Cell Biol. *31*, 4609–4622. https://doi.org/10.1128/MCB.05766-11.
9. Van Keymeulen, A., Fioramonti, M., Centonze, A., Bouvencourt, G., Achouri, Y., and Blanpain, C. (2017). Lineage-Restricted Mammary Stem Cells Sustain the Development, Homeostasis, and Regeneration of the Estrogen Receptor Positive Lineage. Cell Rep. *20*, 1525–1532. https://doi.org/10.1016/j.celrep.2017.07.066.
10. Wang, C., Christin, J.R., Oktay, M.H., and Guo, W. (2017). Lineage-Biased Stem Cells Maintain Estrogen-Receptor-Positive and -Negative Mouse Mammary Luminal Lineages. Cell Rep. *18*, 2825–2835. https://doi.org/10.1016/j.celrep.2017.02.071.
11. Rodilla, V., Dasti, A., Huyghe, M., Lafkas, D., Laurent, C., Reyal, F., and Fre, S. (2015). Luminal progenitors restrict their lineage potential during mammary gland development. PLoS Biol. *13*, e1002069. https://doi.org/10.1371/journal.pbio.1002069.
12. Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913. https://doi.org/10.1038/nm.2000.
13. Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J.T., and Eaves, C.J. (2008). A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat. Med. *14*, 1384–1389. https://doi.org/10.1038/nm.1791.
14. Clarke, R.B., Howell, A., Potten, C.S., and Anderson, E. (1997). Dissociation between steroid receptor expression and cell proliferation in the human breast. Cancer Res. *57*, 4987–4991.
15. Zeps, N., Bentel, J.M., Papadimitriou, J.M., D'Antuono, M.F., and Dawkins, H.J. (1998). Estrogen receptor-negative epithelial cells in mouse mammary gland development and growth. Differentiation *62*, 221–226. https://doi.org/10.1046/j.1432-0436.1998.6250221.x.
16. Brisken, C., Park, S., Vass, T., Lydon, J.P., O'Malley, B.W., and Weinberg, R.A. (1998). A paracrine role for the epithelial progesterone receptor in mammary gland development. Proc. Natl. Acad. Sci. USA *95*, 5076–5081. https://doi.org/10.1073/pnas.95.9.5076.
17. Russo, J., Ao, X., Grill, C., and Russo, I.H. (1999). Pattern of distribution of cells positive for estrogen receptor α and progesterone receptor in relation to proliferating cells in the mammary gland. Breast Cancer Res. Treat. *53*, 217–227. https://doi.org/10.1023/A:1006186719322.
18. Mallepell, S., Krust, A., Chambon, P., and Brisken, C. (2006). Paracrine signaling through the epithelial estrogen receptor α is required for proliferation and morphogenesis in the mammary gland. Proc. Natl. Acad. Sci. USA *103*, 2196–2201. https://doi.org/10.1073/pnas.0510974103.
19. Giraddi, R.R., Shehata, M., Gallardo, M., Blasco, M.A., Simons, B.D., and Stingl, J. (2015). Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. Nat. Commun. *6*, 8487. https://doi.org/10.1038/ncomms9487.
20. Shehata, M., Waterhouse, P.D., Casey, A.E., Fang, H., Hazelwood, L., and Khokha, R. (2018). Proliferative heterogeneity of murine epithelial cells in the adult mammary gland. Commun. Biol. *1*, 111. https://doi.org/10.1038/s42003-018-0114-7.
21. Huebner, R.J., Lechler, T., and Ewald, A.J. (2014). Developmental stratification of the mammary epithelium occurs through symmetry-breaking vertical divisions of apically positioned luminal cells. Development *141*, 1085–1094. https://doi.org/10.1242/dev.103333.
22. Nguyen-Ngoc, K.-V., Cheung, K.J., Brenot, A., Shamir, E.R., Gray, R.S., Hines, W.C., Yaswen, P., Werb, Z., and Ewald, A.J. (2012). ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. Proc. Natl. Acad. Sci. USA *109*, E2595–E2604. https://doi.org/10.1073/pnas.1212834109.
23. Neumann, N.M., Perrone, M.C., Veldhuis, J.H., Huebner, R.J., Zhan, H., Devreotes, P.N., Brodland, G.W., and Ewald, A.J. (2018). Coordination of Receptor Tyrosine Kinase Signaling and Interfacial Tension Dynamics Drives Radial Intercalation and Tube Elongation. Dev. Cell *45*, 67–82.e6. https://doi.org/10.1016/j.devcel.2018.03.011.
24. Huebner, R.J., Neumann, N.M., and Ewald, A.J. (2016). Mammary epithelial tubes elongate through MAPK-dependent coordination of cell migration. Development *143*, 983–993. https://doi.org/10.1242/dev.127944.
25. Neumann, N.M., Kim, D.M., Huebner, R.J., and Ewald, A.J. (2023). Collective cell migration is spatiotemporally regulated during mammary epithelial bifurcation. J. Cell Sci. *136*, jcs259275. https://doi.org/10.1242/jcs.259275.
26. Myllymäki, S.-M., Kaczyńska, B., Lan, Q., and Mikkola, M.L. (2023). Spatially coordinated cell cycle activity and motility govern bifurcation of mammary branches. J. Cell Biol. *222*, e202209005. https://doi.org/10.1083/jcb.202209005.
27. Myllymäki, S.-M., and Mikkola, M.L. (2019). Inductive signals in branching morphogenesis – lessons from mammary and salivary glands. Curr. Opin. Cell Biol. *61*, 72–78. https://doi.org/10.1016/j.ceb.2019.07.001.
28. Goodwin, K., and Nelson, C.M. (2021). Uncovering cellular networks in branching morphogenesis using single-cell transcriptomics. Curr. Top. Dev. Biol. *143*, 239–280. https://doi.org/10.1016/bs.ctdb.2020.09.004.

29. Sumbal, J., Fre, S., and Sumbalova Koledova, Z. (2024). Fibroblast-induced mammary epithelial branching depends on fibroblast contractility. PLoS Biol. *22*, e3002093. https://doi.org/10.1371/journal.pbio.3002093.

30. Scheele, C.L.G.J., Hannezo, E., Muraro, M.J., Zomer, A., Langedijk, N.S.M., van Oudenaarden, A., Simons, B.D., and van Rheenen, J. (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. Nature *542*, 313–317. https://doi.org/10.1038/nature21046.

31. Messal, H.A., van Rheenen, J., and Scheele, C.L.G.J. (2021). An Intravital Microscopy Toolbox to Study Mammary Gland Dynamics from Cellular Level to Organ Scale. J. Mammary Gland Biol. Neoplasia *26*, 9–27. https://doi.org/10.1007/s10911-021-09487-2.

32. Lloyd-Lewis, B., Gobbo, F., Perkins, M., Jacquemin, G., Huyghe, M., Faraldo, M.M., and Fre, S. (2022). In vivo imaging of mammary epithelial cell dynamics in response to lineage-biased Wnt/β-catenin activation. Cell Rep. *38*, 110461. https://doi.org/10.1016/j.celrep.2022.110461.

33. Dawson, C.A., Mueller, S.N., Lindeman, G.J., Rios, A.C., and Visvader, J.E. (2021). Intravital microscopy of dynamic single-cell behavior in mouse mammary tissue. Nat. Protoc. *16*, 1907–1935. https://doi.org/10.1038/s41596-020-00473-2.

34. Paine, I., Chauviere, A., Landua, J., Sreekumar, A., Cristini, V., Rosen, J., and Lewis, M.T. (2016). A Geometrically-Constrained Mathematical Model of Mammary Gland Ductal Elongation Reveals Novel Cellular Dynamics within the Terminal End Bud. PLoS Comput. Biol. *12*, e1004839. https://doi.org/10.1371/journal.pcbi.1004839.

35. Williams, J.M., and Daniel, C.W. (1983). Mammary ductal elongation: Differentiation of myoepithelium and basal lamina during branching morphogenesis. Dev. Biol. *97*, 274–290. https://doi.org/10.1016/0012-1606(83)90086-6.

36. Kendrick, H., Regan, J.L., Magnay, F.A., Grigoriadis, A., Mitsopoulos, C., Zvelebil, M., and Smalley, M.J. (2008). Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC Genom. *9*, 591. https://doi.org/10.1186/1471-2164-9-591.

37. Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.-L., Vaillant, F., Yagita, H., Lindeman, G.J., Smyth, G.K., and Visvader, J.E. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. *12*, R21. https://doi.org/10.1186/bcr2560.

38. Welm, B.E., Tepera, S.B., Venezia, T., Graubert, T.A., Rosen, J.M., and Goodell, M.A. (2002). Sca-1pos Cells in the Mouse Mammary Gland Represent an Enriched Progenitor Cell Population. Dev. Biol. *245*, 42–56. https://doi.org/10.1006/dbio.2002.0625.

39. Soyal, S.M., Mukherjee, A., Lee, K.Y.-S., Li, J., Li, H., DeMayo, F.J., and Lydon, J.P. (2005). Cre-mediated recombination in cell lineages that express the progesterone receptor. Genesis *41*, 58–66. https://doi.org/10.1002/gene.20098.

40. Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J. Cell Biol. *176*, 19–26. https://doi.org/10.1083/jcb.200604065.

41. Fernandez-Valdivia, R., Mukherjee, A., Creighton, C.J., Buser, A.C., DeMayo, F.J., Edwards, D.P., and Lydon, J.P. (2008). Transcriptional Response of the Murine Mammary Gland to Acute Progesterone Exposure. Endocrinology *149*, 6236–6250. https://doi.org/10.1210/en.2008-0768.

42. Kunasegaran, K., Ho, V., Chang, T.H.-T., De Silva, D., Bakker, M.L., Christoffels, V.M., and Pietersen, A.M. (2014). Transcriptional Repressor Tbx3 Is Required for the Hormone-Sensing Cell Lineage in Mammary Epithelium. PLoS One *9*, e110191. https://doi.org/10.1371/journal.pone.0110191.

43. Milevskiy, M.J.G., Coughlan, H.D., Kane, S.R., Johanson, T.M., Kordafshari, S., Chan, W.F., Tsai, M., Surgenor, E., Wilcox, S., Allan, R.S., et al. (2023). Three-dimensional genome architecture coordinates key regulators of lineage specification in mammary epithelial cells. Cell Genom. *3*, 100424. https://doi.org/10.1016/j.xgen.2023.100424.

44. Pal, B., Chen, Y., Milevskiy, M.J.G., Vaillant, F., Prokopuk, L., Dawson, C.A., Capaldo, B.D., Song, X., Jackling, F., Timpson, P., et al. (2021). Single cell transcriptome atlas of mouse mammary epithelial cells across development. Breast Cancer Res. *23*, 69. https://doi.org/10.1186/s13058-021-01445-4.

45. Pal, B., Chen, Y., Vaillant, F., Jamieson, P., Gordon, L., Rios, A.C., Wilcox, S., Fu, N., Liu, K.H., Jackling, F.C., et al. (2017). Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. Nat. Commun. *8*, 1627. https://doi.org/10.1038/s41467-017-01560-x.

46. Zeng, Y.A., and Nusse, R. (2010). Wnt Proteins Are Self-Renewal Factors for Mammary Stem Cells and Promote Their Long-Term Expansion in Culture. Cell Stem Cell *6*, 568–577. https://doi.org/10.1016/j.stem.2010.03.020.

47. Brisken, C., Heineman, A., Chavarria, T., Elenbaas, B., Tan, J., Dey, S.K., McMahon, J.A., McMahon, A.P., and Weinberg, R.A. (2000). Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. Genes Dev. *14*, 650–654.

48. Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M.L., Oakes, S.R., Lindeman, G.J., and Visvader, J.E. (2008). Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell Stem Cell *3*, 429–441. https://doi.org/10.1016/j.stem.2008.08.001.

49. Gu, B., Watanabe, K., Sun, P., Fallahi, M., and Dai, X. (2013). Chromatin Effector Pygo2 Mediates Wnt-Notch Crosstalk to Suppress Luminal/Alveolar Potential of Mammary Stem and Basal Cells. Cell Stem Cell *13*, 48–61. https://doi.org/10.1016/j.stem.2013.04.012.

50. Smith, P., Godde, N., Rubio, S., Tekeste, M., Vladar, E.K., Axelrod, J.D., Henderson, D.J., Milgrom-Hoffman, M., Humbert, P.O., and Hinck, L. (2019). VANGL2 regulates luminal epithelial organization and cell turnover in the mammary gland. Sci. Rep. *9*, 7079. https://doi.org/10.1038/s41598-019-43444-8.

51. Abe, T., Sakaue-Sawano, A., Kiyonari, H., Shioi, G., Inoue, K.i., Horiuchi, T., Nakao, K., Miyawaki, A., Aizawa, S., and Fujimori, T. (2013). Visualization of cell cycle in mouse embryos with Fucci2 reporter directed by *Rosa26* promoter. Development *140*, 237–246. https://doi.org/10.1242/dev.084111.

52. Wagner, K.U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Li, M., Furth, P.A., and Hennighausen, L. (1997). Cre-mediated gene deletion in the mammary gland. Nucleic Acids Res. *25*, 4323–4330.

53. Pauklin, S., and Vallier, L. (2014). The Cell-Cycle State of Stem Cells Determines Cell Fate Propensity. Cell *156*, 1338. https://doi.org/10.1016/j.cell.2014.02.044.

54. Singh, A.M., Chappell, J., Trost, R., Lin, L., Wang, T., Tang, J., Matlock, B.K., Weller, K.P., Wu, H., Zhao, S., et al. (2014). Cell-Cycle Control of Developmentally Regulated Transcription Factors Accounts for Heterogeneity in Human Pluripotent Cells. Stem Cell Rep. *2*, 398. https://doi.org/10.1016/j.stemcr.2014.02.009.

55. Goldman, J.M., Murr, A.S., and Cooper, R.L. (2007). The rodent estrous cycle: characterization of vaginal cytology and its utility in toxicological studies. Birth Defects Res. B Dev. Reprod. Toxicol. *80*, 84–97. https://doi.org/10.1002/bdrb.20106.

56. Joshi, P.A., Jackson, H.W., Beristain, A.G., Di Grappa, M.A., Mote, P.A., Clarke, C.L., Stingl, J., Waterhouse, P.D., and Khokha, R. (2010). Progesterone induces adult mammary stem cell expansion. Nature *465*, 803–807. https://doi.org/10.1038/nature09091.

57. Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322–327. https://doi.org/10.1038/nature12948.

58. Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., and Clevers, H. (2010). Intestinal Crypt Homeostasis Results from

Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. Cell *143*, 134–144. https://doi.org/10.1016/j.cell.2010.09.016.

59. Lafkas, D., Rodilla, V., Huyghe, M., Mourao, L., Kiaris, H., and Fre, S. (2013). Notch3 marks clonogenic mammary luminal progenitor cells in vivo. J. Cell Biol. *203*, 47–56. https://doi.org/10.1083/jcb.201307046.
60. Pachitariu, M., and Stringer, C. (2022). Cellpose 2.0: how to train your own model. Nat. Methods *19*, 1634–1641. https://doi.org/10.1038/s41592-022-01663-4.
61. Ershov, D., Phan, M.-S., Pylvänäinen, J.W., Rigaud, S.U., Le Blanc, L., Charles-Orszag, A., Conway, J.R.W., Laine, R.F., Roy, N.H., Bonazzi, D., et al. (2022). TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. Nat. Methods *19*, 829–832. https://doi.org/10.1038/s41592-022-01507-1.
62. Lakhani, S., Ellis, I., Schnitt, S., Tan, P., and van de Vijver, M. (2012). WHO Classification of Tumours of the Breast, 4th ed. (IARC Press).
63. Kaneko, T., Endo, M., Tsunoda, S., Nakagawa, Y., and Abe, H. (2020). Simple induction of pseudopregnancy by artificial stimulation using a sonic vibration in rats. Sci. Rep. *10*, 2729. https://doi.org/10.1038/s41598-020-59611-1.
64. Stone, B.J., and Srodulski, S.J. (2023). Inducing Pseudopregnancy in Female Mice Without the Need for Vasectomized Males Prior to Non-Surgical Embryo Transfer or Artificial Insemination. J. Vis. Exp. *197*, e65477. https://doi.org/10.3791/65477.
65. Srinivas, S., Watanabe, T., Lin, C.-S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev. Biol. *1*, 4. https://doi.org/10.1186/1471-213X-1-4.
66. Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84–88. https://doi.org/10.1038/nature04372.
67. McLean, A.C., Valenzuela, N., Fai, S., and Bennett, S.A.L. (2012). Performing Vaginal Lavage, Crystal Violet Staining, and Vaginal Cytological Evaluation for Mouse Estrous Cycle Staging Identification. J. Vis. Exp. *67*, e4389. https://doi.org/10.3791/4389.
68. Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. *47*, e47. https://doi.org/10.1093/nar/gkz114.
69. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.
70. Chen, Y., Lun, A.T.L., and Smyth, G.K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Res. *5*, 1438. https://doi.org/10.12688/f1000research.8987.2.
71. Sheridan, J.M., Ritchie, M.E., Best, S.A., Jiang, K., Beck, T.J., Vaillant, F., Liu, K., Dickins, R.A., Smyth, G.K., Lindeman, G.J., and Visvader, J.E. (2015). A pooled shRNA screen for regulators of primary mammary stem and progenitor cells identifies roles for Asap1 and Prox1. BMC Cancer *15*, 221. https://doi.org/10.1186/s12885-015-1187-z.
72. Rios, A.C., Capaldo, B.D., Vaillant, F., Pal, B., van Ineveld, R., Dawson, C.A., Chen, Y., Nolan, E., Fu, N.Y., et al.; 3DTCLSM Group (2019). Intraclonal Plasticity in Mammary Tumors Revealed through Large-Scale Single-Cell Resolution 3D Imaging. Cancer Cell *35*, 953. https://doi.org/10.1016/j.ccell.2019.05.011.
73. Dawson, C.A., Pal, B., Vaillant, F., Gandolfo, L.C., Liu, Z., Bleriot, C., Ginhoux, F., Smyth, G.K., Lindeman, G.J., Mueller, S.N., et al. (2020). Tissue-resident ductal macrophages survey the mammary epithelium and facilitate tissue remodelling. Nat. Cell Biol. *22*, 546–558. https://doi.org/10.1038/s41556-020-0505-0.
74. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat FITC anti-mouse CD29 (clone HMβ1-1)	Biolegend	Cat# 102206; RRID: AB_312882
Rat APC-Cy7 anti-mouse CD29 (clone HMβ1-1)	Biolegend	Cat# 102226; RRID: AB_2128076
Armenian hamster Pacific blue anti-mouse CD24 (clone M1/69)	Biolegend	Cat# 101820; RRID: AB_572010
Armenian hamster PE anti-mouse CD61 (clone HMβ3-1)	Biolegend	Cat# 104308; RRID: AB_313084
rat IgG2b PE anti-mouse Tspan8 (clone 657909)	R&D	Cat# FAB6524P; RRID: AB_10891351
Rat IgG2b APC anti-mouse Tspan8 (clone 657909)	R&D	Cat# FAB6524A; RRID: AB_10889851
Rat IgG2a PE anti-mouse CD14 (clone Sa14-2)	Biolegend	Cat# 123310; RRID: AB_940582
Rat IgG2a PE-Cy7 anti-mouse CD14 (clone Sa14-2)	Biolegend	Cat# 123316; RRID: AB_10645329
Rat IgG2a PE-Cy7 anti-mouse Prom1 (CD133) (Clone 315-2C11)	Biolegend	Cat# 141210; RRID: AB_2564068
Armenian hamster APC-Cy7 anti-mouse Podoplanin (Pdpn) (Clone 8.1.1)	Biolegend	Cat# 127418; RRID: AB_2629803
Rat BV786 IgG2a anti-mouse Sca-1 (Ly-6A/E) (Clone E13–161.7)	BD Biosciences	Cat# 744327; RRID: AB_2742155
Rat IgG2a BV785 anti mouse Sca-1 (Ly-6A/E) (Clone D7)	Biolegend	Cat# 108139; RRID: AB_2565957
Rat PE anti-mouse CD49b (Clone HMα2)	Biolegend	Cat# 103506; RRID: AB_313029
Rat APC anti-mouse CD31 (clone 390)	Biolegend	Cat# 102410; RRID: AB_312905
Rat IgG2a Biotin anti-mouse CD31 (Clone MEC13.3)	Biolegend	Cat# 102504; RRID: AB_312910
Rat IgG2b APC anti-mouse CD45 (clone 30-F11)	Biolegend	Cat# 103112; RRID: AB_312977
Rat IgG2b Biotin anti-mouse CD45 (Clone 30-F11)	Biolegend	Cat# 103104; RRID: AB_312969
Rat IgG2b APC anti-mouse TER-119 (clone TER-119)	Biolegend	Cat# 116212; RRID: AB_313713
Rat IgG2b Biotin TER-119 (clone TER-119)	Biolegend	Cat# 116204; RRID: AB_313704
Streptavidin BV650	BD Biosciences	Cat# 563855; RRID: AB_2869528
Rabbit anti-mouse Keratin 5 (polyclonal)	Biolegend	Cat# 905503; RRID: AB_2734679
Sheep anti-human progesterone receptor (polyclonal)	R&D Systems	Cat# AF5415; RRID:AB_2299396
Sheep anti-human estrogen receptor α (polyclonal)	R&D Systems	Cat# AF5715; RRID:AB_2102093
Chicken anti-GFP (polyclonal)	Abcam	Cat# ab13970; RRID: AB_300798
Rat IgG2a monoclonal anti-E-cadherin (clone ECCD-2)	Thermo Fisher Scientific	Cat# 13–1900; RRID: AB_2533005
Goat anti-chicken IgY (H + L) Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11039; RRID: AB_2534096
donkey anti-chicken IgY (H + L) Alexa Fluor 488	Jackson ImmunoResearch Laboratories Inc.	Cat# 703-545-155; RRID: AB_2340375
Donkey anti-Rabbit IgG (H + L) Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-31573; RRID: AB_2536183
Donkey anti-sheep IgG (H + L) Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-21436; RRID: AB_2535857
Donkey anti-sheep IgG (H + L) Alexa Fluor 594	Thermo Fisher Scientific	Cat# A-11016; RRID: AB_2534083
Donkey anti-sheep IgG (H + L) Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21448; RRID: AB_2535865
Goat anti-Rat IgG (H + L) secondary antibody, Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-21434; RRID: AB_2535855
Goat anti-Rat IgG (H + L) secondary antibody, Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21247; RRID: AB_141778
Chemicals, peptides, and recombinant proteins		
Trypsin (2.5%)	Thermo Fisher Scientific	Cat# 15090-046

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dispase II (neutral protease, grade II)	Sigma Aldrich	Cat# 4942078001; CAS Number: 9001-92-7
Gibco DMEM/F12, GlutaMAX Supplement	Thermo Fisher Scientific	Cat# 10565018
Gibco Penicillin-Streptomycin	Thermo Fisher Scientific	Cat# 15140122
Insulin (Roche)	Sigma Aldrich	Cat# 11376497001; CAS Number: 11061-68-0
Cyclodextrin-encapsulated hydrocortisone	Sigma Aldrich	Cat# H0396; PubChem Substance ID: 24895401
Epidermal growth factor (EGF)	Sigma Aldrich	Cat# E9644; CAS Number: 62229-50-9
Cholera Toxin	Sigma Aldrich	Cat# C-8052; CAS Number: 9012-63-9
Trypan Blue 0.4%	Thermo Fisher Scientific	Cat# T10282; CAS Number: 72-57-1
Deoxyribonuclease I (DNAse I)	Worthington Biochemical Corp	Cat# LS002140; CAS Number: 9003-98-9
Clostridiopeptidase A (Collagenase)	Sigma Aldrich	Cat# C9891; CAS Number: 9001-12-1
Hyaluronate 4-glycanohydrolase (Hyaluronidase)	Sigma Aldrich	Cat# H3506; CAS Number: 37326-33-3
EGTA	Sigma Aldrich	Cat# E0396; CAS Number: 67-42-5
7-Aminoactinomycin D (7AAD)	Sigma Aldrich	Cat# A9400; CAS Number: 7240-37-1
DPBS, no calcium, no magnesium	Gibco	Cat# 14190144
UltraPure™ DNase/RNase-Free Distilled Water	Invitrogen	Cat# 10977015
HEPES pH 7.5	Sigma Aldrich	Cat# H3375
AMPure® XP Beads	Beckman Coulter	Cat# A63881
Glycerol	Ajax Finechem	Cat# 242-500mL
Sodium dodecyl sulfate	Sigma Aldrich	Cat# L6026; CAS Number: 151-21-3
Sodium dodecyl sulfate solution	Sigma Aldrich	Cat# 71736; CAS Number: 151-21-3
Triton X-100	Sigma Aldrich	Cat# T9284; CAS Number: 9036-19-5
Tween 20	Sigma Aldrich	Cat# P7949; CAS Number: 9005-64-5
Bovine Serum Albumin	Sigma Aldrich	Cat# A7906; CAS Number: 9048-46-8
D(−)Fructose	Sigma Aldrich	Cat# F0127; CAS Number: 57-48-7
Urea	Sigma Aldrich	Cat# U5378; CAS Number: 57-13-6
Trizma (Tris) base (2-Amino-2-(hydroxymethyl)-1,3-propanediol)	Sigma Aldrich	Cat# T1699; CAS Number: 77-86-1
Ethylenediaminetetraacetic acid, disodium salt dihydrate (EDTA)	Sigma Aldrich	Cat# 03690; CAS Number: 6381-92-6
Paraformaldehyde (PFA) powder	Sigma Aldrich	Cat# P6148; CAS Number: 30525-89-4
Hair removal cream	Veet	Cat# EA_3019439
Isoflurane	Pharmachem	Cat# FGISO0250
Silicone grease	Dow Corning	Cat# DC976VF
Doxycycline hyclate	Sigma	Cat# D9891; CAS Number: 24390-14-5
Medroxyprogesterone 17-acetate (MPA)	Sigma	M1629-1G; CAS Number: 71-58-9
Sunflower seed oil from *Helianthus annuus*	Sigma	Cat# S5007-1L; CAS Number: 8001-21-6
5-ethynyl-2′-deoxyuridine	Thermo Fisher Scientific (Invitrogen)	Cat# A10044; CAS Number: 61135-33-9
Matrigel	Corning	Cat# 354234
4′,6-Diamidino-2-phenylindole dihydrochloride (DAPI)	Thermo Fisher Scientific	Cat# 62248; CAS Number: 28718-90-3
Phalloidin Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-22287
Phalloidin Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-34055
Phalloidin Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-12379
Critical commercial assays		
miRNeasy Micro Kit	Qiagen	Cat# 217084
TruSeq RNA Library Preparation Kit v2	Illumina	Cat# RS-122-2001, Cat#RS-122-2002

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
On-column DNase	Qiagen	Cat# 79256
Click-iT Plus EdU Flow Cytometry Assay kit, Alexa Fluor 594 dye	Thermo Fisher Scientific	Cat# C10646
Click-iT EdU Cell Proliferation Kit for Imaging, Alexa Fluor 647 dye	Thermo Fisher Scientific	Cat# C10340
BD Perm/Wash	BD Biosciences	Cat# 554723
Deposited data		
RNA-seq	Milevskiy et al.[43]	GSE227750
RNA-seq	This paper	GSE268949
scRNA-seq mouse atlas	Pal et al.[41]	GSE164017
MaSCOT-AI Cellpose model	This paper	https://github.com/cadaws/MaSCOT-AI
Experimental models: Organisms/strains		
Mouse: FVB/NJ	The Jackson Laboratory	Cat# JAX:001800; RRID: IMSR_JAX:001800
Mouse: C57BL/6J	The Jackson Laboratory	Cat# JAX:000664; RRID: IMSR_JAX:000664
Mouse: Krt5-rtTA-IRES-GFP	WEHI	Rios et al.[55]
Mouse: Elf5-rtTA-IRES-GFP	WEHI	Rios et al.[55]
Mouse: PR-cre	John Lydon, Baylor College of Medicine	Soyal et al.[37]
Mouse: R26R-Confetti	Hans Clevers, Hubrecht Institute	Snippert et al.[56]
Mouse: R26R-EYFP	The Jackson Laboratory	Cat# JAX: 006148, RRID: IMSR_JAX:006148
Mouse: Ai9 or Ai9(RCL-tdT)	The Jackson Laboratory	Cat# JAX: 007909 RRID: IMSR_JAX:007909
Mouse: B6.Cg-Tg(tetO-cre)	Jackson Laboratory	Cat# JAX: 006234 RRID: IMSR_JAX:006234
Software and algorithms		
R	R Project for Statistical Computing	RRID: SCR_001905
R packages Rsubread, edgeR, limma, Glimma, plyranges, Iranges, rtracklayer, BSgenome.Mmusculus.UCSC.mm10	Bioconductor	RRID: SCR_006442
R packages gplots, ComplexUpset, Pheatmap, viridisLite, dplyr, readr, writexl, purrr, ggplot2, tools, nplyr, magrittr, stringr, tidyverse	CRAN	RRID: SCR_003005
Mus musculus gene information	NCBI	https://ftp.ncbi.nlm.nih.gov
Cellpose 2	Pachitariu et al.[58] https://github.com/MouseLand/cellpose	RRID: SCR_021716
Trackmate	Ershov et al.[59]	N/A
FluoView	Olympus	N/A
Prism 10	GraphPad	N/A
FIJI/ImageJ	https://imagej.nih.gov/ij/	N/A
ZEISS ZEN Microscopy Software	Zeiss	RRID: SCR_013672
Leica Application Suite X	Leica	RRID: SCR_013673
Python Programming Language 3	Python Software Foundation	RRID: SCR_008394
Miniconda	https://docs.anaconda.com/miniconda/	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Imaris 8.2, 8.4 and 9.5 with XT	Oxford Instruments	RRID: SCR_007370
FlowJo (Version 10.10)	BD Biosciences	RRID: SCR_008520
OMIQ	Dotmatics, https://app.omiq.ai	N/A
Photoshop	Adobe	RRID: SCR_014199
Other		
BD FACSymphony A3	BD Biosciences	N/A
BD FACSAria Fusion	BD Biosciences	N/A
LowBind DNA tubes	Eppendorf	Cat# 0030108051
Tapestation loading tips	Agilent Technologies	Cat# 5067-5153
D1000 DNA ScreenTape	Agilent Technologies	Cat# 5067- 5582
D1000 DNA Reagents	Agilent Technologies	Cat# 5067- 5583
Upright multiphoton	Olympus	FVMPE-RS
Mai Tai laser	Spectra-Physics	HPDS-OL 690–1040 nm
Insight laser	Spectra-Physics	X3-OL 680–1300 nm
25 × 1.05 NA objective	Olympus	FV30-AC25W
Fluorescence stereo microscope	Olympus	SZX16
Dichroic filters	Semrock	FF458-Di02, FF520-Di02 and FF555-Di03
Bandpass filters	Semrock	FF01-417/60, FF01-488/50, FF01-539/30 and FF01-650/200
Zeiss LSM 880 Fast Airyscan Confocal microscope with 25x/0.8 multi-immersion, 40x/1.20 Oil DIC or 63x/1.40 Oil DIC objectives	Zeiss	N/A
Zeiss LSM 980 Fast Airyscan Confocal microscope with 25x/0.8 multi-immersion, 40x/1.20 Oil DIC or 63x/1.40 Oil DIC objectives	Zeiss	N/A
Leica SP8 Resonant Scanning Confocal microscope with 20x/0.75 multi-immersion, 40x/1.3 OIL DIC or 63x/1.40 Oil DIC objectives	Leica Microsystems GmbH	N/A
Nikon Eclipse Ts2 microscope	Nikon	N/A
Quadro P4000 8GB GGDR5 graphics card	NVIDIA	N/A
Illumina Nextseq 500	Illumina	RRID: SCR_014983

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

C57BL/6, FVB/NJ and Rosa26-LOX-stop-LOX-eYFP (YFP)[65] mice were maintained and provided by the WEHI animal facility. PR-cre knockin mice were kindly provided by J. Lydon[39] and Rosa26-rtTA3-Confetti mice were a gift from H. Clevers. K5-rtTA3-IRES-GFP and Elf5-rtTA3-IRES-GFP mice were described in Rios et al. 2014.[57] All mice were bred and maintained in the WEHI animal facility according to institutional guidelines. All experiments were approved by the WEHI Animal Ethics Committee.

METHOD DETAILS

Mouse models: Induction and treatment

Recombination of the Confetti allele in Elf5-rtTA or K5-rtTA/Tet-Cre/Confetti mice was induced by intraperitoneal (i.p.) injection of 1 mg doxycycline (10 mg/mL in PBS) (Sigma) at 28, 29 and 30 days of age and mice were analyzed at 5 weeks (TEBs) or 6 weeks (ducts). For lineage tracing at the adult stage, mice at 8.5 weeks were administered a doxycycline-laced diet for a period of one week and collected after a chase of 5 weeks. PR-cre/Confetti ducts were analyzed at 6 and 9 weeks. PR-cre/YFP mice or Elf5-rtTA3-IRES-GFP mice were collected at 5 weeks; TEBs and subtending ducts were microdissected for RNA-seq analysis. For analysis of pregnancy, 9-week-old adult mice were mated and females scored for the presence of vaginal plugs. The day of the observed plug was considered 0.5 dP. For medroxyprogesterone acetate (MPA) treatment, 9-week-old mice were injected subcutaneously with 8 mg MPA in 200 μL 10% ethanol +90% sunflower oil or vehicle alone for 5 days then analyzed on the 6th day.

Isolation of mammary epithelial cells

To prepare mammary epithelial cell suspensions, mammary glands were dissected, lymph nodes removed and glands chopped with a McIllwain Tissue Chopper.[66] Cells were then dissociated with collagenase/hyaluronidase/DNase solution for 1 h at 37°C, Trypsin/EGTA solution for 5 min at 37°C, Dispase/DNase for 5 min at 37°C then filtered through a 70 μm filter. Single-cell suspensions were then blocked for 10 min in 2% FCS/PBS, anti-FCR 1:40, anti-rat IgG 1:80 and DNase 1:10. Blocked cell suspensions were then stained using a combination of the following primary antibodies: FITC/APC-Cy7 anti-mouse CD29 (rat, clone HMβ1-1, BioLegend Cat#102206 and #102226, 1/200 dilution), Pacific Blue anti-mouse CD24 (Armenian Hamster, clone M1/69, BioLegend Cat#101820, 1/200), PE anti-mouse CD61 (Armenian Hamster, clone HMβ3-1, BioLegend Cat#104308, 1/100), PE/APC anti-mouse Tspan8 (rat, clone 657909, R&D Cat#FAB6524P and #FAB6524A, 1/200), PE/PE-Cy7 anti-mouse CD14 (rat, clone Sa14-2, BioLegend Cat#123310 and #123316, 1/200), PE-Cy7 anti-mouse Prom1 (CD133) (rat, clone 315-2C11, BioLegend Cat#141210, 1/200), APC-Cy7 anti-mouse Podoplanin (Pdpn) (Armenian Hamster, clone 8.1.1, BioLegend Cat#127418, 1/200), BV786 anti-mouse Sca-1 (Ly-6A/E) (rat, clone E13–161.7, BD Cat#744327, 1/200 and rat, clone D7, BioLegend Cat#108139, 1/200), PE anti-mouse CD49b (rat, clone HMα2, BioLegend Cat#103506, 1/200), APC anti-mouse CD31 (rat, clone 390, BioLegend Cat#102410, 1/40), APC/Biotin anti-mouse CD45 (rat, clone 30-F11, BioLegend, Cat#103112 and #103104, 1/100), and APC/Biotin anti-mouse TER-119/erythroid cell (rat, clone TER-119, BioLegend Cat#116212 and #116204, 1/80 dilution), Biotin anti-mouse CD31 (rat, MEC13.3, BioLegend Cat#102504, 1/40). Biotin-conjugated antibodies were stained using BV650 Streptavidin (BD, Cat#563855, 1/500). Following washing of primary antibodies from cell suspensions, cells were resuspended in 2% FCS/PBS with 7-AAD (1/500 dilution) to exclude dead and dying cells. For cells used for RNA-sequencing, glands from 2 to 4 mice were pooled and dissociated as described above with a red-cell lysis performed immediately after Dispase digestion.

For FACS analysis, stained mammary single cell suspensions were analyzed on a BD Biosciences FACSymphony A3. FCS files were exported and subsequently analyzed using either FlowJo (BD, version 10.10) or OMIQ (Dotmatics, https://app.omiq.ai).

Mammary epithelial colony forming assays

Sorted mammary epithelial cells were washed in 2% FCS/PBS, resuspended in MEC media[66] and counted: 5,000 cells were combined with Matrigel (Corning Cat#354234) resulting in 20% cells 80% Matrigel, then 1000 cells were plated into a chamber well-slide and incubated at 37°C, 5% CO_2 and 5% O_2 for up to 14 days, with media changes performed every 3 days. Colonies were imaged with a Nikon Eclipse Ts2 microscope. 3x z stack images were taken for each Matrigel drop and auto-aligned and auto-blended in Photoshop (Adobe) to merge into one image and to capture all colonies in focus.

EdU cell labeling and estrus staging

5-Ethynyl-2′-deoxyuridine (EdU) labeling and staining was performed using the Click-iT Plus EdU Flow Cytometry Assay kit from Invitrogen (ThermoFisher #C10646). Mice were i.p. injected with 300 μg EdU and collected at various times following labeling, as indicated. Mammary epithelial single-cell suspensions were isolated from glands as indicated above. Following primary antibody staining, $1x10^6$ cells were fixed in 4% paraformaldehyde in PBS for 15 min at room temperature and washed using 1x saponin-based permeabilization buffer (Invitrogen) in 1% bovine-serum albumin (BSA) (Permwash), removing dead and red blood cells. EdU was detected with Alexa Fluor 594 picolyl azide for 30 min at room temperature (RT). Cells were washed, resuspended in Permwash and analyzed on a BD FACSymphony A3 Cell Analyzer. To determine mouse estrus stage, mice were first euthanized, then using a 20 μL pipette and tip, 20 μL of PBS was flushed through the vagina before spreading onto a microscope slide and air-drying. Slides where stained with haematoxylin and eosin for staging.[67]

Intracellular FACS analysis

Mouse mammary glands from PR-cre/YFP adult mice were harvested, isolated and prepared as described above. Following primary cell antibody staining, $1x10^6$ cells were fixed in 1x PBS 4% paraformaldehyde at RT for 15 min followed by 2x washes using Permwash. Fixed cells were then blocked for intracellular staining, using 2% FCS/PBS, 1% BSA, anti-FCR 1:40 and anti-rat IgG 1:80 for 15 min on ice. Blocked cells were stained with anti-PR antibody (clone SP2, Invitrogen, 1:200 dilution) for 30 min on ice followed by secondary antibody (anti-Rabbit Alexa Fluor 594, 1:1000) for 15 min on ice. Cells were resuspended in Permwash and analyzed on a BD FACSSymphony A3 Cell Analyzer.

RNA-sequencing

Sorted mammary epithelial cells were snap-frozen on dry ice. Pellets were resuspended in 350 μL of QIAzol and RNA extracted using Qiagen's miRNeasy micro-kit (Qiagen Cat#217084), following the recommended guidelines, including on-column DNase digestion (Qiagen Cat#79256). 50–100 ng of total RNA was used to produce sequencing libraries with Illumina's TruSeq RNA v2 kit (Illumina #RS-122-2001, #RS-122-2002). RNA libraries were sequenced on the Illumina Nextseq500, producing >30 million paired-end reads of 80 bp in length.

Differential gene expression and gene ontology analysis

RNA-seq sequencing reads were aligned to the mouse genome mm10 using the Rsubread package v1.29.6.[68] The number of reads mapped to each gene was quantified using featureCounts.[69] Lowly expressed genes were filtered if their total number of counts

across all the samples is less than 10. Compositional differences between libraries were normalized using the trimmed mean of M-values (TMM) method. Differential gene expression analysis was performed using the edgeR package[70] with quasi negative binomial generalized linear models fitted to each gene using the glmQLFit function in edgeR. Differential expression was assessed using the quasi-likelihood F-test. Genes were considered to be differentially expressed if they achieved a false discovery rate of 5%. Gene ontology and KEGG pathway analyses were performed using the goana and kegga function in edgeR.[70] Heatmaps were produced using the coolmap function in the limma package. Gene signatures of luminal progenitor (LP) and mature luminal (ML) were obtained using the bulk RNA-seq data from our previous work.[71] The average log-reads per kilobase per million (RPKM) values of those gene signatures were calculated and used as the LP and ML signature scores for each sample. The t-SNE visualization and single-cell clustering were performed in the same way as described in the previous single-cell study.[44]

3D confocal imaging of whole-mount tissue

3D wholemount fluorescence imaging was performed on fresh tissue[72] and following IVM.[33] At 4°C for all steps, tissue was collected and fixed for 30 mins–2 hrs in 4% paraformaldehyde (PFA, Merck) in PBS pH 7.4, washed in 0.1% Tween 20 (Sigma) in PBS for 30 min, then washed in 0.2% Tween 20, 0.2% Triton X-100 (Sigma), 0.02% SDS (Sigma) and 0.2% BSA (Sigma) in PBS overnight. Tissue was incubated with primary antibodies overnight in 0.1% Triton X-100, 0.02% SDS and 0.2% BSA, washed 3 × 2 h in staining buffer, then incubated with secondary antibodies overnight. Primary antibodies included: rabbit polyclonal anti-Keratin 5 (Biolegend, Cat#905503), sheep polyclonal anti-PR (R&D Systems, Cat#AF5415), chicken polyclonal anti-GFP (Abcam, Cat#ab13970), rat monoclonal anti-Ki67 (ThermoFisher, Cat#56-5698-82) and sheep polyclonal anti-ERα (R&D Systems, Cat#AF5715). Secondary antibodies: goat anti-chicken IgY Alexa Fluor 488, donkey anti-sheep IgG Alexa Fluor 555, donkey anti-sheep IgG Alexa Fluor 647, donkey anti-chicken IgG Alexa Fluor 488, donkey anti-rat Alexa Fluor 647 and donkey anti-rabbit IgG Alexa Fluor 647. EdU labeling was performed after secondary antibody staining by overnight incubation with Click-iT EdU Imaging Kit (Invitrogen). During secondary staining, tissue was stained with DAPI (Thermo Fisher, 4 μg/mL) and/or F-actin was labeled with Phalloidin Alexa Fluor 488, 555 or 647 (Invitrogen, 1:25). Tissue was then cleared with FUnGI[72] for 2 h at RT or at 4°C overnight. Cleared tissue was dissected under a fluorescence stereomicroscope (Leica) and mounted on a glass slide with slight compression. For confocal microscopy image acquisition, tile scans of z-stacks were acquired at minimum optical section resolution of 1,024 × 1,024 using software optimized Z-resolution. Samples were imaged on a laser scanning confocal microscope: Leica SP8, Zeiss 880 or Zeiss 980. Zeiss microscopes were equipped with 25x/0.8 multi-immersion, 40x/1.20 Oil DIC or 63x/1.40 Oil DIC objectives and the Leica SP8 equipped with 20x/0.75 multi-immersion, 40x/1.3 OIL DIC or 63x/1.40 Oil DIC objectives and PMT/HyD detectors. Optimal Airyscan imaging was performed on a Zeiss 880 using an ×63 oil objective. Images were stitched in LasX (Leica) or Zen (Zeiss) Software and image processing was conducted using Imaris (Bitplane). Images shown are representative of n = 3–5 mice.

Intravital imaging

For IVM of mammary glands,[33,73] mice were anesthetized with 3% Isoflurane in 20% O_2 in N_2, which was then decreased to 1–1.5% to maintain stable anesthesia. Temperature was maintained at 37°C with a thermostatic heat pad (FHC, Cat. 40-90-8D, 40-90-5D-02 and 40-90-2) and rectal probe. The right flank of the mouse was shaved and hair removal cream (Veet) applied for 1 min before removing hair. The 4th mammary gland was exposed by making a skin incision along the midline and right hind leg, followed by cauterizing the 5th mammary gland. Exposed tissue was constantly irrigated with PBS. The resulting skin flap was gently separated from the abdomen and secured on a small, raised platform. Typical and superficial TEBs were found under a fluorescence dissection microscope (Leica) and further exposed for IVM by micro-dissection of connective tissue from the top of the fat pad and separation of fat lobes if required. The exposed torso was covered with plastic wrap to prevent dehydration. The mammary gland was surrounded by a ring of silicone grease (Dow Corning, Cat#DC976VF) and sealed with a small coverslip. Hydration was maintained by i.p. injection of saline at a rate of 100 μL per hour. Intravital imaging was performed using the FVMPE-RS two-photon system (Olympus), equipped with MaiTai DeepSee and InSight DeepSee lasers (Spectra-Physics), 2 GaAsP and 2 Multialkali PTM detectors, SIM scanner and a 25x/1.05 NA water objective (Olympus). Three-dimensional image stacks were acquired with 1–2 μm z-step size every 10 min for up to 12 h. Laser power was kept below 0.1W for imaging (~2.5% MaiTai 830 nm, ~6% InSight 970 nm). To separate the Confetti signals, YFP and RFP were excited at 970 nm and collected by bandpass filters 539/30 and 650/200 (Semrock) respectively and detected by the Multialkali PTMs. CFP and GFP (plus second harmonic generation) were both collected by a 488/50 filter with sequential excitation at 830 and 970 nm and detected by GaAsP dectector. SHG stimulated at 830 nm was collected by a 417/60 filter and this signal was subtracted from the GFP/SHG channel post-acquisition in Imaris. The dichroic mirrors used were longpass 458, 520 and 555 nm. In a few experiments, the bandpass filters 475/42 (CFP), 525/50 (GFP+YFP) and 596/83 (RFP) were used with 488 and 560 nm dichroics.

Image analysis

3D confocal images were visualized in Imaris (Oxford Instruments). For EdU imaging analysis, the frequency of markers of interest (i.e., GFP, YFP, EdU) was computed using the Imaris spot measurements, signal colocalization, channel arithmetic and surface function, and further refined by manual curation across optical sections of TEBs and subtending ductal regions. For quantification of PR protein on 3D images, a surface was created on the PR-YFP channel; masks were used on the DAPI and PR channels with the spot feature (5 μm) to detect YFP^+ and PR^+ cells, respectively. Erroneous and stromal spots were removed manually.

Cell migration in IVM was quantified in Imaris by manual cell tracking utilizing cyan/red 3D visualization.[33] 3D drift was corrected by the 'Correct 3D drift' function in Imaris using cell tracks. Residual tissue warping was accounted for in migration speed and displacement by calculating the running average of cell xyz position over three time-points (30 min). mCFP, nGFP, YFP and RFP-positive cells were tracked to measure migration. Average speed (μm/hr) was calculated by averaging the distance per 10 min (minimum 1 h) then multiplying this by 6. Hourly displacement was calculated by averaging the distance traveled between 0 and 1 h, 1–2 h, 2–3 h etc. (minimum 3 h but occasionally 2 h in short movies). Cap cell measurements excluded cap-in-body cells, dividing cells, and cells that eventually migrated into the body. All graphs were generated in Prism (Graphpad).

MaSCOT-AI

3D movies of Confetti-labeled glands were trimmed to 5 h then cropped into z-stacks in Imaris to capture whole cell volumes while minimizing cell overlap, typically 10 μm and up to 30 μm with sparse labelling. nGFP was not analyzed as it gives no morphological information. The cropped z stack movies were flattened by maximum projection in FIJI,[74] then the 5th time point was extracted from all movies for AI model training. The AI model was trained using Cellpose 2.0[60] with initial segmentation by the built-in CP model with automatic cell diameter estimation. Segmentation was manually corrected for approximately 5 images before retraining the model and this was repeated over 150 images, when consistently accurate segmentation was achieved. A trackmate-Cellpose[61] python script was then written in FIJI incorporating the final MaSCOT-AI model (https://github.com/cadaws/MaSCOT-AI). Settings were as follows. Object filter: Cell area <250 μm^2. Tracking settings with SparseLAPTracker: No track splitting or merging, max linking distance 3 μm, max frame gap 3 (2 frames missing), max gap closing distance 5 μm, high track linkage penalties for variation in area (5) and mean intensity (5). Statistics were exported to .csv and the trackmate files exported as .xml, several of which were opened to confirm segmentation and tracking accuracy. Next, an R script was written using dplyr to extract the statistics of interest and compile these into large tables of 5619 cell tracks with sample labels and the measurement at each timepoint. Graphs tracking representative cell measurements over time were generated by first taking a running average across three timepoints, then filtering for tracks with >5 time points (50 min), then calculating the average and dynamic range (standard deviation) of each cell track over time. Cells at each percentile of average measurement for the sample type were then selected and plotted. Max and min measurement graphs were generated by selecting the max or min measurement across each cell track. Short peaks in aspect ratio were counted as a single increase followed by a decrease with differences >1%. Short solidity minima were counted as inverse changes with differences >0.33%, which is approximately 1% of the data range. Long peaks were counted as two increases followed by two decreases (or the inverse) with the thresholds as above. Long peaks were excluded from the count of small peaks. Peak/minima frequency per population was calculated by dividing these counts by the cumulative time every cell in a population was imaged for, minus the start and end periods for each cell that were ineligible (minus 10 min ×2 for short peaks and 20 min ×2 for long peaks per track).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details of experiments can be found in the figure legends, including the statistical tests used and the exact value of n. n represents the number of mice analyzed unless otherwise specified in the figure legend. Statistical tests between two groups used a two-tailed T test with Welch's corrections, unless otherwise specified. Statistical tests between multiple pairs used one-way ANOVA with Sidak correction. Statistical analysis was performed in R v4.2.0 and GraphPad Prism 9. P-values of statistical tests are indicated in figures as follows: n.s. = not significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ and ****$p < 0.0001$. Data as shown as median with Tukey whiskers, mean ± SEM or median with quartiles as indicated in figure legends.