Integrating single-cell RNA-sequencing and functional assays to decipher mammary cell states and lineage hierarchies

Joseph L. Regan* and Matthew J. Smalley

The identification and molecular characterization of cellular hierarchies in complex tissues is key to understanding both normal cellular homeostasis and tumorigenesis. The mammary epithelium is a heterogeneous tissue consisting of two main cellular compartments, an outer basal layer containing myoepithelial cells and an inner luminal layer consisting of estrogen receptor-negative (ER−) ductal cells and secretory alveolar cells (in the fully functional differentiated tissue) and hormone-responsive estrogen receptor-positive (ER+) cells. Recent publications have used single-cell RNA-sequencing (scRNA-seq) analysis to decipher epithelial cell differentiation hierarchies in human and murine mammary glands, and reported the identification of new cell types and states based on the expression of the luminal progenitor cell marker KIT (c-Kit). These studies allow for comprehensive and unbiased analysis of the different cell types that constitute a heterogeneous tissue. Here we discuss scRNA-seq studies in the context of previous research in which mammary epithelial cell populations were molecularly and functionally characterized, and identified c-Kit+ progenitors and cell states analogous to those reported in the recent scRNA-seq studies.

Previous studies to elucidate the cellular identities of mammary epithelial subpopulations have involved functional and molecular characterization by flow cytometric and functional (down to single cell) transplantation assays1–14, as well as, more recently, lineage-tracing studies15–26. Transplantation experiments have generally supported a model in which facultative MaSCs, cells capable of regenerating the epithelium when injected into a cleared mammary fat pad (one free of endogenous epithelium)1,27, are localized to the basal cell layer28,29,30. Progenitor cells, which are functionally defined by high colony-forming and proliferative potential in vitro and limited repopulating ability when transplanted into cleared fat pads, are localized to the luminal layer31,32,36. Differentiated cells do not transplant or generate colonies in vitro. The molecular profiling of mammary epithelial subpopulations functionally defined by their transplantation potential has been extensive9,17,31–40.

Supporting this model, in situ evidence, including lineage-tracing studies from early mammary development, puberty, and alveolagenesis during pregnancy, has shown that basal cells can contribute to the luminal layer39,41–43. We previously proposed, based on in situ analysis, that basal MaSCs located in the cap cell layer of terminal end buds (TEBS), the outermost cell layer of the specialized growth structure that drives ductal growth during puberty, are bipotent and produce daughter cells that contribute to both the basal and luminal cell lineages43. Lineage-tracing experiments from Rios et al.16 and Wang et al.18 were in agreement with transplantation data and our in situ analysis, suggesting that MaSCs in the developing postnatal gland are bipotent18,16,43. However, more recently, it has been shown that, rather than a transcriptionally defined bipotent TEB MaSC, a group of transcriptionally heterogeneous lineage-committed MaSCs mediate development of the pubertal mammary gland and contribute transiently to ductal expansion39, mirroring the organization and neutral drift of adult stem cells observed in the intestine44,45. This model of postnatal mammary gland development is in agreement with saturation, single-cell genetic, and neutral lineage-tracing studies demonstrating that bipotent fetal MaSCs (MaSCs), first functionally and molecularly characterized (including single-cell gene expression analysis demonstrating molecular heterogeneity) by Spike et al.57, exist in the embryo, but that in the postnatal gland, basal and luminal lineages are maintained by separate lineage-committed stem/progenitor populations16–24,26,28–40. During oncogenic transformation, basal and luminal cell populations may lose this restricted lineage potential and acquire multipotency20,24,49,50.

Recent studies have used scRNA-seq, which unlike functional and population-based sequencing studies, allows for unbiased analysis of individual cells in a heterogeneous tissue, to decipher lineage hierarchies and cell states in the mammary epithelium51–54. To investigate cellular heterogeneity and lineage relationships in the human breast, Nguyen et al.51 performed scRNA-seq analysis on fluorescence-activated cell-sorted (FACS) breast epithelial cells and reported the identification of additional cell types within the three main mammary epithelial cell populations, previously identified as basal (B: CD49f(High) EPCAM−, K14+), luminal progenitors (L1: CD49f(High) EPCAM+, ER−, K8/18−), and mature luminal (L2: CD49f− EPCAM+, ER+, K8/18+) cells8,10,51. Significantly, the authors detected replicating KIT+ cells in all three main populations (Basal, L1, and L2), suggesting that each cluster may be maintained by its own KIT+ progenitor cell population, and proposed a continuous lineage hierarchy connecting the basal lineage to the two luminal branches via a bipotent MaSC. Furthermore, the authors highlight adult luminal cells that co-express both luminal (KRT8/18) and basal (KRT14) markers in situ. The receptor tyrosine kinase KIT (c-Kit) has previously been identified as a defining marker of mammary epithelial progenitor cells (summarized in Table 1) and of the cells of origin of BRCA1-mutation breast cancer, luminal ER+ cells7,17,28,34,40,55,56. Similar to Nguyen et al.51, in Regan et al.28, we identified in the mouse, and also functionally tested via in vitro colony-forming assays and...
### Table 1. Studies demonstrating that luminal ER<sup>−</sup> cells are enriched for c-Kit and that c-Kit identifies progenitor cells in the mammary epithelium

| Study (year) | Method(s) | Cells/Progenitor cell marker(s) | Results |
|-------------|-----------|---------------------------------|---------|
| Natali et al. (1992)<sup>56</sup> | Immunohistochemistry | Normal human breast tissue | High levels of c-Kit protein detected in the luminal alveolar/ductal epithelium but not in the basal/myoepithelial layer. |
| Matsuda et al. (1993)<sup>55</sup> | Immunostaining | Mouse mammary cell populations | Luminal ER<sup>−</sup> cells are in vitro progenitors and possess limited mammary gland repopulation potential. |
| Hines et al. (1995)<sup>50</sup> | Immunostaining | Basal CD24<sup>+/−</sup>, CD49f/CD29<sup>−</sup>/High Sca-1<sup>−</sup>/CD61<sup>−</sup> | Basal cells contain facultative MaSCs. |
| Ulivi et al. (2004)<sup>54</sup> | Transcriptome analysis | Basal CD24<sup>+/−</sup>, CD49f/CD29<sup>−</sup>/High Sca-1<sup>−</sup>/CD61<sup>−</sup> | Basal cells contain facultative MaSCs. |
| Westbury et al. (2009)<sup>54</sup> | Transcriptome analysis | Basal CD24<sup>+/−</sup>, CD49f/CD29<sup>−</sup>/High Sca-1<sup>−</sup>/CD61<sup>−</sup> | Basal cells contain facultative MaSCs. |
| Kendrick et al. (2008)<sup>58</sup> | FACS | Transcriptome analysis | Functional testing of isolated c-Kit<sup>+</sup> cells was not carried out. |
| Sleeman et al. (2007)<sup>6</sup> | Immunostaining | Mouse mammary cell populations | c-Kit is an in vitro and in vivo functional marker of mammary progenitors and lineage-primed cell states in basal, luminal, ER<sup>+</sup>, and luminal ER<sup>−</sup> cell populations. Facultative MaSCs are CD24<sup>+/−</sup>, Sca-1<sup>−</sup>/CD49f/CD29<sup>−</sup>/High c-Kit<sup>+</sup>. |
| Stingl et al. (2006)<sup>7</sup> | Colony-forming assays | Mouse mammary cell populations | c-Kit is highly expressed in mouse and human luminal progenitor cells. Functional testing of isolated c-Kit<sup>+</sup> cells was not carried out in these studies. |
| Asselin-Labat et al. (2007)<sup>79</sup> | Gene expression analysis | Mouse mammary cell populations | c-Kit<sup>−</sup> luminal cells expand during early pregnancy and are in vitro colony-forming progenitors. In vivo functional testing of isolated c-Kit<sup>−</sup> cells was not carried out. |
| Kim and Villadsen (2018)<sup>83</sup> | Immunohistochemistry | Mouse mammary cell subpopulations | Identified luminal ER<sup>−</sup> and luminal ER<sup>+</sup> progenitor cells in mouse and human. Detected c-Kit<sup>+</sup> cells in the luminal populations of FVB/N mice but not in C57Bl6J mice. Functional testing of isolated c-Kit<sup>+</sup> cells was not carried out in this study. |
| Asselin-Labat et al. (2011)<sup>85</sup> | Colony-forming assays | Mouse mammary cell subpopulations | Identified luminal ER<sup>−</sup> and luminal ER<sup>+</sup> progenitor cells in mouse and human. Detected c-Kit<sup>+</sup> cells in the luminal populations of FVB/N mice but not in C57Bl6J mice. Functional testing of isolated c-Kit<sup>+</sup> cells was not carried out in this study. |
| Shehata et al. (2012)<sup>90</sup> | Colony-forming assays | Mouse mammary cell subpopulations | Detected c-Kit<sup>+</sup> cells in the luminal populations of FVB/N mouse and human. Functional testing of isolated c-Kit<sup>+</sup> cells was not carried out in this study. |
| Pal et al. (2017)<sup>92</sup> | scRNA-Seq | Mouse mammary cell populations | Hierarchical clustering revealed that luminal progenitors are enriched for c-Kit. Transcriptome mapping identified rare c-Kit<sup>−</sup> lineage-primed basal cells. |
| Bach et al. (2017)<sup>93</sup> | scRNA-Seq | Mouse mammary cell populations | Identified c-Kit<sup>+</sup> luminal progenitor cells that give rise to intermediate, alveolar, and hormone-sensitive progenitors. |
| Chung et al. (2019)<sup>99</sup> | scRNA-Seq | Mouse mammary cell populations | Identified KIT<sup>+</sup> progenitor cells in each mammary population, including L1.1 luminal (ER<sup>−</sup> KIT<sup>−</sup>/High) and L1.2 luminal (ER<sup>−</sup> KIT<sup>−</sup>/Low) progenitors. |
| Giraddi et al. (2018)<sup>94</sup> | snATAC-seq | Embryonic and postnatal mouse mammary cells | c-Kit is most highly expressed and chromatin accessible in luminal progenitor cells. |
| Giraddi et al. (2018)<sup>95</sup> | scRNA-Seq | Embryonic and postnatal mouse mammary cells | c-Kit is most highly expressed and chromatin accessible in luminal progenitor cells. |

*Published in partnership with the Breast Cancer Research Foundation*
Gene expression was only detected (at very low levels) in two of three isolates of the luminal Sca-1+ cells, luminal genes Esr1 and Krt18 (b, c, RHS), and basal gene Krt14 (d RHS) relative to comparator luminal Sca-1− c-Kit− cells, in murine mammary epithelium. Data are presented as fold expression levels ±95% confidence intervals (n = three independently harvested isolates of each cell population). *Gene expression was undetectable in these populations in all three independent isolates. **Gene expression was only detected (at very low levels) in two of three isolates of the luminal Sca-1− c-Kit− population. Therefore, no error bars are shown for this sample. Images used with permission under a CC-BY 4.0 license from Nguyen et al.51 and Regan et al.28. cleared mammary fat pad transplantation, c-Kit+ and c-Kit− cell states within each of the mammary epithelial basal (CD24−/Low Sca-1− CD49f+/High c-Kit+ and c-Kit−), myoepithelial (CD24+/Low Sca-1− CD49f+/Low c-Kit+ and c-Kit−), luminal ER− (CD24+/High Sca-1− c-Kit+/Low and c-Kit+/High), and luminal ER+ (CD24+/Low Sca-1− c-Kit−/CD49f+/Low Sca-1+ c-Kit+ and c-Kit+) cellular compartments36. The expression of KIT, as well as the luminal markers KRT8/18 and ESR1 and basal marker KRT14, in each of Nguyen et al.’s human breast populations of B, Myo, L1.1, L1.2, and L2, are consistent with the expression levels reported in Regan et al.28 in the corresponding murine basal, myoepithelial, luminal ER− c-Kit−/High, luminal ER+ c-Kit−/Low, and luminal ER+ cells, respectively (Fig. 1). The KIT+ cells identified by Nguyen et al.51 are therefore likely equivalent to the c-Kit+ progenitor cells previously reported in Regan et al.28, which was the first study to functionally characterize c-Kit as a progenitor marker in the mammary gland (Table 1). When discussing KIT as a progenitor cell marker, Nguyen et al. incorrectly cite Stingl et al.29 and Shehata et al.60. These papers, respectively, did not investigate or functionally test c-Kit as a progenitor marker in the mammary gland.

Nguyen et al.51 observed fractions of cells that co-express both luminal K8 and basal K14 markers, and report that such K8+ K14+ cells had previously been observed in mouse fMASCs by Spike et al.37 (such fetal cells were also previously described by Sun et al.35), but not in adult human tissue in homeostasis. However, while the canonical view among mouse mammary developmental biologists is that the K5/14 pair is a basal marker and the K8/18 pair is a luminal marker38–40, breast pathologists have known for many years that keratins 5 and 14 (and indeed another “basal” keratin, 17) are in fact expressed in basal cells of human breast ducts and in the luminal cells of the terminal ductal lobuloalveolar units (TDLUs)58–61. Indeed, K5/K18 and K14/K18 double-positive cells are not uncommon in human TDLUs61. More recently, Boecker et al.62 identified K5+ K18/19+ and K5+ K18/K19− populations in the luminal layer of ductal and TDLU breast tissue in situ62, while in human breast epithelial populations isolated by flow cytometry, the progenitor populations (Lin− CD49f− EpCAM−) include cells double-positive for K5/6 and K14 — and notably are also c-Kit+62. To add to the complexity of these marker patterns, K19 has been described both as a marker of progenitors66–68 and highly expressed in differentiated luminal ER+ cells69. Boecker et al.62 termed the populations they identified as progenitors and intermediary cells, respectively, but it is difficult to definitively assign such functions purely on the basis of marker expression, or indeed ex vivo assays. Of course, human breast tissue cannot be lineage-traced through transgene activation as one can in the mouse, but use of cytochrome C oxidase (CCO) mutations in the mitochondrial genome has proven feasible as an approach. Cereser et al.70 report the presence of CCO-deficient clonal expansions in both ducts and TDLUs of normal breast70. Notably, the expansions were limited to the luminal layers, and...
they found no evidence of luminal CCO-deficient clones contributing to the basal layer. Therefore, if the K5/K14/c-Kit⁺ luminal cells of the human breast are indeed progenitors, they are lineage-restricted.

Keratin expression patterns in the mouse mammary epithelium are somewhat easier to define, but also not as straightforward as often suggested. Unlike in the human, when analyzed in situ, K14 and K8/18 in the mouse appear to be restricted to the basal and luminal cell layers, respectively. Indeed, we have rarely (if ever) observed a luminal cell in the normal resting adult mammary gland we could confidently say is K14 positive, or a basal cell that is K8/18 positive, by immunofluorescence in situ, and this is in agreement with most studies. However, immunohistochemical analysis of the mouse mammary gland by Mikaelian et al. has detected rare weak K14 staining of luminal cells from birth to puberty and weak K8/18 labeling of basal cells during mammary morphogenesis, which were most easily visualized during lactation. As an added complication, it should be noted that in the mammary alveoli, the basal/myoepithelial cells form a classic “basket-like network” around the secretory cells, and in that location, the “luminal” cells are in fact touching the basement membrane through the gaps between the myoepithelial cells. Interestingly, therefore, in agreement with Mikaelian et al., when basal and luminal subpopulations were isolated by flow cytometry and stained by immunofluorescence, we found that c-Kit⁺ luminal cells (which were approx. 50% of the total mammary epithelium) were all strongly K18⁺ but also weakly K14⁺, and that c-Kit⁺ basal cells were strongly K14⁺ and weakly K18⁺ (Fig. 2b)[38]. c-Kit-negative single luminal and basal cells prepared and stained at the same time were respectively K18⁺ K14− and K14⁺ K18−, suggesting that we were not seeing background staining in the c-Kit-positive cells. This discrepancy is likely due to signal/noise ratio when using in situ immunofluorescence approaches — enhancing the K14 staining to a level where it can be detected in luminal cells would result in a huge excess of staining from the basal cells as well as background signal from other cell types in the mammary gland (and likewise for K18 detection in basal cells), which is notorious for background fluorescence coming from adipocytes. Thus, only approaches based on single-cell separation will accurately detect mouse cells expressing the “luminal” keratin 18 and the “basal” keratin 14, and as we report using such approaches, such cells express the c-Kit marker[39]. Note that the scRNA-seq analysis of mouse mammary epithelium by Bach et al. shows that a subset of luminal cells have Krt14 expression levels equivalent to the mean expression level of Krt14 in basal cells. Their differentiation trajectory maps show that the Krt14-expressing luminal cells are enriched in a progenitor population that is also c-Kit-positive[55].

In contrast, we find that cells double-positive for “basal” keratin 5 and “luminal” keratin 19 are readily detectable in the mouse luminal epithelium in situ (Fig. 2c, d). Interestingly, K19 has been proposed to be a neutral switch keratin that permits the changeover of one type of cytoskeleton to the other[48,71]. We have particularly noted KS-positive cells in the body cell region of terminal end buds in situ (Fig. 2c). The origin of these cells is unclear. Rios et al. reported that using a Krt5-promoter-driven cell-labeling approach, labeled cells were only observed in the basal compartment, but generated both luminal and basal daughter clones, and hence proposed the existence of bipotent basal stem cells arising from the basal layer of the TEBs[46]. However, the work of Scheele et al. and others[18,23,46,47] suggests that cap cells (the basal cell layer of the TEBs) do not contribute to the luminal layer of the subducting duct; therefore KS-positive body cells, if they are cap cell-derived, are unlikely to contribute to outgrowth of the ducts. In contrast, if these cells are derived from the body cells, they are switching on high levels of KS expression, but whether this is only transient — perhaps a temporary failure of lineage specification in a newly established daughter cell that is later corrected — is unclear.

Therefore, while use of keratins as basal/luminal lineage markers is more robust in the mouse mammary epithelium than in the human, single-cell analysis approaches have demonstrated that even the mouse has a more promiscuous pattern of keratin expression than previously suspected, and that this promiscuous expression of keratins is seen in c-Kit⁺ stem/progenitor cells. Plasticity in the expression of keratins and other genes within c-Kit⁺ luminal progenitors may relate to their potential to contribute to multiple cell lineages during epithelial remodeling, e.g., at involution of the mammary gland after weaning[72]. In addition, the phenotypic plasticity and multilineage differentiation potential of these luminal progenitors is consistent with their ability to give rise to tumors with basal features[40,50], as well as lineage switching in response to injury and oncogene activation[29,24,29]. It is clear, therefore, that a great deal of caution must be used when keratin promoters are being used for lineage-tracing studies in the mouse or for assigning luminal/basal identity in human cells. Indeed, in a dissociated human breast epithelial cell population, keratin expression levels alone cannot be used to assign basal/luminal identity to a cell with any confidence.

To address the debate as to whether homeostasis and development in the postnatal mammary gland are maintained by bipotent MaSCs[15,16,43] or lineage-restricted basal and luminal cells[48,49,52], Nguyen et al. performed pseudotemporal reconstruction-based lineage hierarchy analysis. This analysis identified a continuous lineage connecting the basal lineage, via a bipotent MaSC, to the two luminal branches. These results agree with previous models of mammary differentiation wherein a bipotent basal MaSC generates daughter cells that differentiate into myoepithelial and luminal cell lineages[15,16,43]. However, Nguyen et al. propose that their results differ from previous studies in that L1.2 cells (luminal ER⁺ c-kit⁻/low cells) are progenitors to L1.1 cells (luminal ER⁻ c-kit⁻/high cells), and that c-kit⁺/high L1.1 cells are another type of mature differentiated luminal cell rather than a luminal progenitor upstream of luminal ER⁺ L2 cells. Based on this pseudotemporal analysis, the authors suggest that KIT is not a marker of luminal progenitor cells. This is a surprising conclusion considering that L1.2 progenitor cells do express KIT (Fig. 1), which as well as being a defining marker of mouse and human progenitor cell gene expression signatures[17,24,40,52,53], has been functionally demonstrated as a progenitor cell marker (Table 1).

Similar to Nguyen et al[51], Pal et al. used scRNA-seq to identify lineage relationships in the mouse mammary gland, and also suggested that bipotent basal MaSCs give rise to basal and luminal lineages[52]. Supporting our previous assessment of intermediate cells in the luminal lineage[28], the authors also described the identification of intermediate luminal cells. Significantly, Pal et al. report the identification of rare mixed-lineage or “lineage-primed” c-Kit-expressing basal cells in the adult mammary gland and state, “It is presumed that these cells represent a transient population that is poised for commitment to the luminal lineage, reminiscent of “lineage-primed” stem and progenitor cells initially reported in the hematopoietic system.” These lineage-primed c-Kit⁺ basal cells comprised ~5% of the basal compartment and expressed luminal genes such as Esr1, Pshr, Csn2, and Areg in addition to basal genes. Pal et al. state, “these data suggest that the basal state may predate commitment to a luminal cell fate in the post-natal mammary gland.”

In Regan et al., we also identified cells that we described as lineage-primed basal cells (CD24⁺/low Sca1⁻ CD49f⁺/high c-Kit⁻) in the adult mammary gland that expressed luminal genes, including those described by Pal et al. (Esr1, Pshr, Csn2, and Areg), but that clustered with the basal facultative MaSCs[28]. Significantly, we functionally tested these cells by single-cell cleared mammary fat pad transplantation and demonstrated that they can
reconstitute an entire ductal tree, although at a lower frequency (1 in 8 ± 95% CI 1 in 3–1 in 21.3) than facultative c-Kit\(^+\) MaSCs (1 in 3 ± 95% CI 1 in 1.69–1 in 6.27), the highest enrichment of facultative MaSCs reported to date and potentially a pure facultative MaSC population. Based on these data, we came to the same conclusion as Pal et al.\(^{52}\) and described these c-Kit\(^+\) basal cells as intermediate MaSCs that were undergoing “lineage priming,” in which stem cells express genes associated with their
differentiated daughter populations. This was the first time that lineage-primed basal cells in the adult mammary gland had been reported and functionally tested. In contrast to Nguyen et al. and Pal et al., scRNA-Seq by Bach et al. on mouse mammary epithelial cells at nulliparous, mid gestation, lactation, and post involution concluded that, rather than clearly defined clusters maintained by their own stem/progenitor population, a continuous spectrum of differentiation exists. In this model, a common luminal progenitor cell, which notably expressed c-Kit at high levels, gives rise to intermediate, restricted alveolar, and hormone-sensitive progenitors.

More recently, Giraddi et al. used scRNA-seq and transposase-accessible chromatin sequencing (ATAC-seq), which examines global chromatin accessibility of embryonic, postnatal, and adult mouse mammary epithelia, to elucidate the lineage hierarchies and biological programs that generate mature cell types from their embryonic precursors. This work was more consistent with the conclusions of Bach et al. than Nguyen et al. and Pal.
et al.\textsuperscript{52,} as well as the lineage-tracing studies showing that while embryonic mammary cells are bipotent, in the adult gland, basal and luminal cell lineages are derived from and maintained by separate lineage-committed progenitor populations\textsuperscript{18–24,42,46–48}. Similar to Pal et al.\textsuperscript{52} Giradddi et al.\textsuperscript{53} also identified rare c-Kit\textsuperscript{+} basal cells, although they did not occur at a frequency greater than the expected doublet frequency (~1%) of the 10X Genomics Chromium System sequencing platform\textsuperscript{54,}, a frequency similar to the c-Kit\textsuperscript{+} basal cells that Pal et al.\textsuperscript{52} also detected using the 10X platform. In contrast, the lineage-primed c-Kit\textsuperscript{+} basal cells that we identified in our 2012 study were visually confirmed to be single cells prior to performing the single-cell transplants, in which they displayed a transplantation-frequency intermediary to facultative c-Kit\textsuperscript{+} MaSCs and c-Kit\textsuperscript{+} luminal progenitor cells. In addition, immunofluorescence staining of single c-Kit\textsuperscript{+} basal cells demonstrated that they expressed both K14 and K18 (Fig. 2b)\textsuperscript{26,27,28}.

Transcriptional profiling by Giradddi et al.\textsuperscript{54} did not detect any distinct adult basal stem cell subpopulation. However, ATAC-seq revealed that adult basal cells display an embryonic MaSC-type chromatin accessibility at luminal gene loci, which the authors speculate allows for lineage plasticity\textsuperscript{54,73,77}. Such plasticity may account for acquisition of multilineage potential upon perturbation of a homeostatic niche environment, such as during cell isolation and ex vivo culture, transplantation assays, wounding, and cancer\textsuperscript{17,54,77–80}. The performance of a particular cell type during functional assays may therefore be a product of both their transcriptional heterogeneity and the context in which they are challenged\textsuperscript{49}. Similar functional stem cell capacities have also been described in embryonic tissue, intestine, bone marrow, skin, and lung\textsuperscript{81–83}. These observations challenge the concept of fixed-cell identities in complex tissues, and suggest a more fluid concept of cell state (for a more detailed discussion of this concept see Wahl and Spike\textsuperscript{49}). With this in mind, a potential mammary epithelial cell hierarchy based on lineage tracing, functional assay, and recent scRNA-seq and snATAC-seq studies is shown in Fig. 3.

Future studies that aim to map fluid cell-state dynamics and their regulatory mechanisms will require the use of single-cell and single-molecule epigenomic technologies that reveal a cell's regulatory potential, rather than its current state, as indicated by its transcriptome\textsuperscript{80,85}. Indeed, Chung et al.\textsuperscript{73} recently demonstrated that single-cell chromatin accessibility mapping of mammary gland development using single-nucleus ATAC-seq (snATAC-seq) enables greater resolution of cell-state heterogeneity, and to be a better indicator of cell state during development than scRNA-seq\textsuperscript{72}. The lineage relationships delineated in this study were consistent with those of Bach et al.\textsuperscript{53} and Giradddi et al.\textsuperscript{54,} and also found c-Kit to be most highly expressed and chromatin accessible in luminal progenitor cells.

**CONCLUDING REMARKS**

Taken together, the weight of evidence supports c-Kit as a progenitor marker in the mammary epithelium and, more importantly, one that is functionally characterized and can be used to enrich stem/progenitor cells. Indeed, we have already begun to understand the signaling pathways downstream of c-Kit in mammary progenitor cells\textsuperscript{86–90}. scRNA-seq studies, which allow for comprehensive and unbiased analysis of the different cell types that constitute a heterogeneous tissue\textsuperscript{57}, have been extremely valuable in contributing to our understanding of lineage relationships and cell-state heterogeneity in the mammary gland. However, in order to fully understand the significance of these studies, it is essential to link them to functional data, in particular where such data already exist, and future studies should aim to do so. The evidence from lineage tracing, scRNA-seq, and snATAC-seq studies currently supports a model in which MaSCs in the embryo are bipotent, whereas in the adult gland, stem/progenitor cells are lineage-restricted, and facultative MaSCs (defined by functional studies) are induced to acquire multilineage potential upon loss of homeostasis/injury. Bipotent fetal MaSCs are described as fMaSCs to differentiate them from adult facultative MaSCs. However, the scientific literature up to now continues to refer to adult cells with facultative stem cell potential simply as MaSCs or, in a handful of publications, adult MaSCs (aMaSCs\textsuperscript{57,69}), which is no longer an accurate or apt description. We therefore propose the renaming of MaSCs in the postnatal gland as “inducible mammary stem cells” (iMaSCs). This new definition will help to more clearly define the status and stem cell potential of functionally defined iMaSCs in the era of large-scale single-cell molecular profiling.

**DATA AVAILABILITY**

Source data for all figures and tables are provided in the paper. No new datasets have been generated or analyzed for this article.

Received: 10 September 2019; Accepted: 2 July 2020; Published online: 29 July 2020

**REFERENCES**

1. Regan, J. & Smalley, M. Prospective isolation and functional analysis of stem and differentiated cells from the mouse mammary gland. Stem Cell Res. 3, 124–136 (2007).
2. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. Nature 439, 84–88 (2006).
3. Plaks, V. et al. Lgr5 expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep. 3, 70–78 (2013).
4. Visser, K. E. et al. Developmental stage-specific contribution of LGR5+ cells to basal and luminal epithelial lineages in the postnatal mammary gland. J. Pathol. 228, 300–309 (2012).
5. Sleeper, K. E. et al. Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J. Cell Biol. 176, 19–26 (2007).
6. Wilson, N. K. et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. Cell Stem Cell 16, 712–724 (2015).
7. Smalley, M. J. et al. Isolation of mouse mammary epithelial subpopulations: a comparison of leading methods. J. Mammary Gland Biol. Neoplasia 17, 91–97 (2012).
8. Stingl, J. et al. Purification and unique properties of mammary epithelial stem cells. Nature 439, 993–997 (2006).
9. Shehata, M. et al. Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. 14, R134 (2012).
10. Alvi, A. J. et al. Functional and molecular characterisation of mammary side population cells. Breast Cancer Res. 5, R1–8 (2003).
11. Britt, K. L. et al. Pregnancy in the mature adult mouse does not alter the proportion of mammary epithelial stem/progenitor cells. Breast Cancer Res 11, R20 (2009).
12. Nguyen, L. V. et al. Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells. Cell Stem Cell 14, 253–263 (2014).
13. Ginestier, C. et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of a poor clinical outcome. Cell Stem Cell 1, 555–567 (2007).
14. Wang, D. et al. Identification of multipotent mammary stem cells by protein C receptor expression. Nature 517, 81 (2014).
15. Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. In situ identification of bipotent stem cells in the mammary gland. Nature 506, 322 (2014).
16. Lim, E. et al. Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. 12, R21–R21 (2010).
17. Chang, T. H.-T. et al. New insights into lineage restriction of mammary gland epithelium using parity-identified multipotent epithelial cells. Breast Cancer Res. 16, R1–R1 (2014).
18. Van Keymeulen, A. et al. Distinct stem cells contribute to mammary gland development and maintenance. Nature 479, 189 (2011).
20. Van Keymeulen, A. et al. Lineage-restricted mammary stem cells sustain the development, homeostasis, and regeneration of the estrogen receptor positive lineage. *Cell Rep.* **20**, 1525–1532 (2017).

21. Davis, F. M. et al. Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat. Commun.* **7**, 13053 (2016).

22. Lloyd-Lewis, B., Davis, F. M., Harris, O. B., Hitchcock, J. R. & Watson, C. J. Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells. *Development* **145**, dev164079 (2018).

23. Scheele, C. L. G. J. et al. Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* **542**, 313–317 (2017).

24. Koren, S. et al. PIK3CAH1047R induces multipotency and multi-lineage mammary tumours. *Nature* **525**, 114 (2015).

25. van Amerongen, R. in *Mammary Stem Cells. Methods in Molecular Biology.* (ed. Vivanco, M. del M.) 187–211 (Springer New York, 2015).

26. de Moosdijk, A. A. A., Fu, N. Y., Rios, A. C., Visvader, J. E. & van Amerongen, R. in *Mammary Gland Development. Methods in Molecular Biology.* Vol. 1501. (eds. Martin, F., Stein, T. & Howlin, J.) 291–308 (Humana Press, New York, 2017).

27. DeOme, K. B., Faulkin, L. J. Jr., Bern, H. A. & Blair, P. B. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* **19**, 515–520 (1959).

28. Regan, J. L. et al. c-Ki-t is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. *Oncogene* **36**, 819 (2017).

29. Asselin-Labat, M. L. et al. Gata-3 is an essential regulator of mammary gland development and homeostasis in mice. *Cell Stem Cell* **10**, 62–73 (2012).

30. Jones, C. et al. Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer. *Cancer Res.* **64**, 3037–3045 (2004).

31. Santagata, S. et al. Spatially correlated phenotyping reveals K5-positive luminal progenitors and cell type-specific gene expression programs in early mammary gland development. *Nat. Commun.* **9**, 4609–4622 (2018).

32. Spike, B. T. et al. A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* **10**, 183–197 (2012).

33. Giraddi, R. R. et al. Single-cell transcriptomes distinguish stem cell state changes and lineage specification programs in early mammary gland development. *Cell Rep.* **24**, 1653–1666.e7 (2018).

34. Grigoriadis, A. et al. Establishment of the epithelial-specific transcriptome of normal and malignant human breast cells based on MPSS and array expression data. *Breast Cancer Res.*** **8**, R56 (2006).

35. Pecce, S. et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell** **140**, 62–73 (2010).

36. Wansbury, O. et al. Transcriptome analysis of embryonic mammary cells reveals insights into mammary lineage establishment. *Breast Cancer Res.* **13**, R79–R79 (2011).

37. Clarke, R. B. et al. A putative human breast stem cell population is enriched for steroid receptor-positive cells. *Dev. Biol.* **277**, 443–56 (2005).

38. Asselin-Labat, M.-L. et al. 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 (2011).

39. Bale, K. et al. Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat. Commun.* **8**, 1627 (2017).

40. Bach, K. et al. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat. Commun.* **8**, 2128 (2017).

41. Girardi, R. R. et al. Single-cell transcriptomes distinguish stem cell state changes and lineage specification programs in early mammary gland development. *Cell Rep.* **24**, 1653–1666.e7 (2018).

42. Driessens, J. G. et al. Identification of mammary gland progenitors and not from basal stem cells. *Cell Stem Cell* **7**, 403–410 (2017).

43. Molyneux, G. et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **7**, 403–410 (2017).

44. Nguyen, Q. H. et al. Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nat. Commun.* **9**, 2028 (2018).

45. Spike, B. T. et al. Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat. Commun.* **8**, 1627 (2017).

46. Ritsma, L. et al. Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. *Nature*** **507**, 362–365 (2014).

47. Wang, C. et al. Single-cell chromatin analysis of mammary gland development reveals cell-state transcriptional regulators and lineage relationships. *Cell Rep.* **29**, 495–510.e6 (2019).

48. Wu, S., Guo, Y. P., May, G. & Enver, T. Bifurcation dynamics in lineage commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007).

49. Wagner, K. L. et al. An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. *Development* **129**, 1377–1386 (2002).

50. Chung, C. Y. et al. Single-cell chromatin analysis of mammary gland development reveals cell-state transcriptional regulators and lineage relationships. *Cell Rep.* **29**, 495–510.e6 (2019).

51. Huang, S., Guo, Y. P., May, G. & Enver, T. Bifurcation dynamics in lineage commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007).

52. Figueiredo, I. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transcriptional programs in early mammary gland development. *Cell Stem Cell* **10**, 1213 (2013).

53. Dravis, C. et al. Epigenetic and transcriptomic profiling of mammary gland development and tumor models disclose regulators of cell state plasticity. *Cell Cancer Cell* **34**, 466–482.e6 (2018).

Published in partnership with the Breast Cancer Research Foundation
ACKNOWLEDGEMENTS

The authors thank Geoffrey M. Wahl (The Salk Institute for Biological Studies) for his critical reading of the paper and helpful comments.

AUTHOR CONTRIBUTIONS

Conceptualization and writing of the original draft, J.L.R.; review and editing, J.L.R. and M.J.S. All the authors read and approved the final version of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to J.L.R.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020