Mammary Stem Cells: Premise, Properties and Perspectives

Bethan Lloyd-Lewis[1,♯*] Olivia B. Harris[1,2], Christine J. Watson[1,2], Felicity M. Davis[3,4*]

[1]Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK, [2]Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, CB2 1QR, UK, [3]School of Pharmacy, The University of Queensland, Brisbane, 4102, Australia, [4]Mater Research Institute, The University of Queensland, Brisbane, 4102, Australia. [♯]Current address: Department of Genetics and Developmental Biology, Institut Curie, Paris, France. *Correspondence: f.davis@uq.edu.au (F.M. Davis) and bethan.lloyd-lewis@curie.fr (B. Lloyd-Lewis).

Keywords

Adult mammary stem cells, mammary gland development, lineage-tracing, breast cancer, cancer stem cells.

Abstract

Adult mammary stem cells (MaSCs) drive postnatal organogenesis and remodeling in the mammary gland, and their longevity and potential have important implications for breast cancer. However, despite intense investigation, the identity, location and differentiation potential of MaSCs remains subject to deliberation. The application of genetic lineage-tracing models, combined with quantitative 3-dimensional imaging and biophysical methods, has provided new insights into the mammary epithelial hierarchy that challenges classical definitions of MaSC potency and behaviors. Herein, we review recent advances—discussing fundamental unresolved properties of MaSC potency, dynamics and plasticity—and point to evolving technologies promising to shed new light on this intractable debate. An elucidation of the physiological mammary differentiation hierarchy is paramount to understanding the complex heterogeneous breast cancer landscape.

Adult mammary stem cells: concepts and challenges

Adult stem cells exist in various organs, such as the intestine, skin and skeletal muscle [1,2]. In these tissues, their primary role is homeostatic, that is, to replenish cells lost to attrition or injury. However, unlike many other organs, the mammary gland primarily develops postnatally [3,4] (Figure 1), and thus stem cells in the adult mammary gland serve both developmental and homeostatic functions.

Construction of the branching ductal epithelium during puberty is driven by hormones, growth factors and local signaling cues, and proceeds via proliferation of **mammary stem cells** (MaSCs, see Glossary) and their progeny within bulbous distal structures known as **terminal end buds** (TEBs) (Figure 1b) [5,6]. By the end of puberty, ductal morphogenesis is complete and the TEBs have fully regressed (Figure 1c) [4]. Although it is generally accepted that stem cells persist in the adult mammary gland following the demise of the TEBs—where they have essential roles in the generation and regeneration of the alveolar (milk-producing) epithelium during pregnancy and lactation (Figure 1c-f)—the location of these cells within the complex ductal epithelium remains elusive [6,7]. Additionally, despite intense investigation and debate, the differentiation potential of adult MaSCs (i.e. their ability to generate one or both of the mammary epithelial cell lineages) remains contentious [7–18]. The longevity and extensive self-renewal properties of these cells, however, place them as probable candidates for oncogenic transformation in some breast cancers [19,20]. Moreover, some breast cancers may be hierarchically-organized and contain a pool of cancer stem cells that drive their precipitous and long-term growth and regrowth [19–21]. Thus, a greater understanding of the identity, plasticity and differentiation potential of adult MaSCs, and the specific pathways that regulate their self-renewal and fate, may also provide important insights into the heterogeneity and treatment resistance of this intractable disease.

Recent studies, using single cell lineage-tracing approaches, have revealed the immense capacity of a single MaSC to contribute to the formation of the ductal epithelium during puberty [16,18] and the alveolar epithelium during pregnancy/lactation [16]. These studies also highlight considerable redundancy within this system [16,18], positing that several hundred lineage-restricted MaSCs actively and stochastically contribute to ductal and alveolar morphogenesis under physiological conditions. This is not entirely surprising, given that lactation is an evolutionarily essential aspect of mammalian survival that demands functional stem cells. However, if MaSCs are the cell-of-origin in some breast cancers, then this superfluity brings with it a heighted opportunity for oncogenic transformation. Regardless, the inextricable connections between MaSCs and breast cancer warrants further investigation, to achieve a unified and enduring characterization of their potential, anatomical location and molecular profile.

Here, we discuss recent insights into the **mammary epithelial cell hierarchy**, addressing unanswered questions relating to MaSC potency, dynamics and plasticity. We discuss the unique challenges in elucidating

the mammary epithelial cell hierarchy and highlight evolving technologies that promise to shed new light on these difficult questions.

The epithelial cell hierarchy: an evolving paradigm

In 1959, a seminal study published in *Cancer Research* demonstrated that fragments of mammary tissue could be transplanted into the epithelium-divested fat pad of a recipient mouse, successfully engraft, and generate an entire ductal epithelium anew [22]. What followed was a divisive pursuit to identify and characterize the cells responsible for the development, maintenance and regeneration of the mammary epithelium (i.e. adult MaSCs) that has lasted for more than 50 years (Figure 2). Transformative advances came on the back of at least three key enabling methodologies: the isolation of cells with enhanced repopulating and self-renewal properties upon transplantation; population-based genetic fate-mapping; and stochastic, single cell genetic lineage-tracing. The ability to image ducts and alveoli in three- [14–16,18,23,24] or four-dimensions [18] (Box 1), combined with quantitative image analysis and biostatistical modeling [5,15,18], has also provided important insights into clonal dynamics and dispersion patterns that could not have been attained through the examination of thin tissue sections [25]. Here, we broadly examine these techniques, summarizing key findings with a retrospective wisdom.

Transplantation

The observation that any fragment of mammary tissue has the potential to regenerate the entire bilayered mammary epithelium upon serial transplantation provided strong evidence that **mammary repopulating cells** (believed to be bona fide MaSCs) were distributed throughout the length of the adult ductal epithelium [26–28]. Subsequent work using retroviral-tagged mammary tissue fragments [29] and limiting dilutions of heterogeneous cell suspensions confirmed these results [30,31], and these studies were in-turn refined and expanded by the identification and purification of a subset of cells with superior repopulating capacity [8–10] (Figure 2). Collectively, these analyses supported the notion that adult MaSCs were **bi/multipotent**. The demonstration that lineage-restricted cells could be forced to adopt a multipotent fate under "regenerative conditions" [11–13] challenged this dogma. It is now widely accepted that mammary repopulating cells, identified by transplantation, are distinct from stem cells that exist under physiological conditions. Nevertheless, this technique has provided some important insights into qualities of self-renewal and regeneration, with enduring relevance.

Population-based genetic fate-mapping

The application of genetic **lineage-tracing** techniques to mammary tissue has enabled temporal examination of lineage relationships under physiological conditions. These studies have utilized tamoxifen- or doxycycline- responsive transgenic mouse models to induce the expression of reporter genes in predefined cohorts of cells [11–15,17]. The genetic label, typically a fluorescent or histochemical reporter, is permanently

expressed by the original cell and is transmitted to all of its progeny. An analysis of reporter expression through time can be used to determine whether the original labeled population contained lineage-restricted stem cells or cells with multi-lineage differentiation potential (Figure 2). In its original application in the mammary gland [11], this approach was used to track the fate of **luminal cells** (e.g., cytokeratin (K)8-expressing) and **basal cells** (e.g., K14-expressing), demonstrating that lineage-restricted MaSCs drive postnatal mammary gland development and maintenance. Subsequent lineage-tracing studies have provided evidence in support of both **unipotent** and bi/multipotent adult MaSCs [12–15,17,32] (Figure 2). Lineage-restricted cell populations have also been shown to convert to multipotency *in vivo* by oncogenic PI3KCA signaling, suggesting that there is scope for plastic transformation and thereby adding further complexity to this system [33,34].

Inconsistencies in recent lineage-tracing studies in the normal mammary gland may be in-part attributable to the temporal expression of pathway-specific promoters [12,13] or the fidelity of pan-lineage promoters. Given that a single mammary stem/progenitor cell is capable of producing many hundred progeny [16], the promiscuous labeling of even a small number of cells of the opposing lineage could significantly confound downstream lineage analysis in this model [15,16]. A second limitation relates to the power of population-based labeling approaches to accurately detect single **clone** expansion, which is a function of both the method of detection and the initial labeling density (Figure 3). To overcome this problem, as well as potential tracing artefacts associated with the preferential labeling of specific (and potentially non-representative) cell sub-populations, a recent study has mapped the fate of all basal cells (a technique termed saturation lineage-tracing) [15]. If rare bipotent MaSCs do reside in the basal compartment and contribute even minimally to mammary gland morphogenesis and homeostasis [1], this could be detected by an increase in the number of fluorescently-labeled luminal cells, observed using both fluorescence activated cell sorting or 3D image quantification. No population flux was detected using either method of analysis in these studies, suggesting that basal MaSCs are indeed lineage-restricted [15]. A subsequent report [35], however, demonstrated that enzymatic digestion prior to 3D imaging [15,18,36,37] can deplete or structurally damage basal cells, postulating that rare bi-lineage clones are not detected under these conditions [35]. Recently described methods for non-proteolytic 3D imaging [16,24], together with quantitative platforms for image analysis, which consider tissue architecture, cell morphology, chimerism and Cre-specificity [15,35], will undoubtedly aid future lineage tracing studies in the mammary gland.

Stochastic, single cell genetic lineage-tracing

Lineage-tracing has facilitated *in situ* examination of MaSC properties under conditions of minimal interference. However, unlike transplantation assays, these studies have been unable to map the fate of a single labeled cell [9]. Obstacles to single cell genetic lineage-tracing have, however, been mitigated in-part by advances in whole-organ clearing [24] and high-resolution 3D imaging [14] (Box 1).

Recently, $R26^{[CA]30}$ mice [38] have been used to achieve unbiased labeling of single proliferating cells in the mammary gland [16]. Genetic labeling in this model is exceedingly rare, and thus it can be combined with 3D imaging to track the fate of a single labeled cell with confidence (Figure 2). A similar approach to achieve low-density, unbiased labeling involves the use of mice that express inducible Cre-recombinase in all cells ($R26^{CreERT2}$). Neutral, multi-color labeling is achieved by crossing these mice with $R26^{Confetti}$ animals, and sparse reporter induction is attained using low doses of tamoxifen [16,18]. Recent application of these models has provided further evidence that unipotent MaSCs drive ductal morphogenesis during puberty [16,18] and alveolar morphogenesis during gestation [16]. However, whilst the small number of cells initially labeled in these models permit the indisputable analysis of clonal progeny, it also limits their power to detect and characterize the full spectrum of stem and progenitor cells present in the mammary epithelium. For example, quiescent bi/multipotent MaSCs, if they exist, would not be detected by this approach [16].

Single cell lineage-tracing has unquestionably demonstrated the immense capacity of unipotent stem cells to contribute to the development of the adult mammary epithelium, whilst at the same time revealing significant redundancy in the construction of each major duct [16,18] and lobuloalveolar structure [16]. Whether adult stem cells work cooperatively or competitively to achieve developmental and morphogenetic outcomes in the mammary gland is an area of active investigation and is discussed in more detail later in this review.

Multiplicity in the mammary gland: roles for potential and quiescent stem cells

In addition to the cells that are responsible for the genesis and expansion of the mammary epithelium (known as professional, functional or bona fide stem cells), there may also exist a population of cells in the adult breast with the capacity to behave as stem cells under certain conditions (i.e. facultative or **potential stem cells**) [1,39]. This may include 1) a subset of cells that remain quiescent during normal tissue development, and 2) cells that are recruited under regenerative conditions [9,11,40] or in cancer [33,34]. Support for a cellular arrangement in the breast that departs from a unidirectional, top-down model is given by transplantation studies. Although it is now generally accepted that mammary repopulating cells are activated under non-homeostatic conditions [11–13], the underlying experimental observation (i.e. that not all cells are capable of repopulating the empty fat-pad [8–10]) points to the existence of a population of cells that have an intermediate or plastic nature. The physiological and pathological role of these cells, and their relationship to putative populations of quiescent MaSCs, is not immediately apparent (Figure 4). However, the notion that fate decisions within the hierarchy are not strictly unidirectional, and in some conditions could be reversed, has wide-reaching implications for oncology and regenerative medicine.

A putative population of quiescent MaSCs

A pool of quiescent stem cells, which have temporarily and reversibly exited the cell cycle, has been observed in various self-renewing tissues, including the skin [41–43] and intestine [44]. These cells may be able to re-enter the cell cycle when required, for example upon injury [45] or homeostasis [46]. Quiescent stem cells are unlikely to be detected by conventional lineage-tracing approaches, which require proliferation for clone identification [47]. As such, label-retention assays have been developed for the analysis of slow-cycling and quiescent cells [48]. DNA-intercalating nucleosides (e.g., BrdU/EdU and [^{3}H]-thymidine) can be used to label cells that are in cycle at the time of the pulse [47]. Alternatively, a GFP-labeled histone H2B model could be used to label specific populations of cells, with expression of H2B-GFP temporally-moderated by administration of doxycycline [43,48,49]. Cells that remain labeled after a pre-determined chase, known as **label-retaining cells**, are presumed to be slow-cycling/quiescent stem cells, but may also be long-lived terminally-differentiated cells [48]. Application of the H2B-GFP model to the mammary gland has identified a novel population of Cd1d$^+$ cells with enhanced repopulating ability upon transplantation [49]. Cd1d$^+$ mammary repopulating cells are also enriched for Bcl11b expression, a C2H2 zinc finger transcription factor that has independently been shown to be associated with physiological quiescence and superior repopulating activity under transplantation conditions [50]. Interestingly, neither Cd1d nor Bcl11b mRNAs are enriched in the recently-identified quiescent basal cell population defined by Lgr5 and Tspan8 expression [51]. These Lgr5$^+$Tspan8^{hi} basal cells, located within the proximal ductal tree, were also demonstrated to have enhanced repopulating activity in limiting dilution transplantation assays [51]. Thus, these data suggest significant multiplicity, even within the putative subset of quiescent mammary repopulating cells.

Unanswered questions: organization, function and recruitment

The proliferative demand on mammary stem and progenitor cells throughout reproductive life is substantial (Figure 1) [3,4]. Thus, the relative importance of quiescent MaSCs in normal development and homeostasis is unclear. How quiescent and potential stem cells may be recruited by specific signals in the microenvironment, and their hierarchical relationship to functional stem cells is also shrouded in uncertainty. In light of the ongoing debate regarding the identity and potency of MaSCs [11–18], the fundamental requirement for proliferation for clone detection in lineage-tracing studies [47], and the idea that quiescent stem cells may reside at the apex of tissue hierarchies [49], one could reasonably suggest that there may be a residual population of quiescent bi/multipotent MaSCs that remain in the postnatal mammary gland after embryonic development (Figure 4). *In utero* DNA-labeling has provided some support for this hypothesis, identifying long-lived label-retaining cells that are able to reversibly re-enter the cell cycle and contribute to tissue development and maintenance [46]. More-recent saturation lineage-tracing, which has been able to label more than 95% of all cells within a single lineage, however, indicates that quiescent MaSCs (if they exist and participate in any way to tissue development and/or homeostasis), are lineage-restricted [15]. Analysis of cell division kinetics and telomere lengths in mammary epithelial populations also suggests that that each lineage is maintained by its own precursors throughout reproductive life [52].

A number of important questions in this area remain unanswered. However, given the complex cellular heterogeneity in breast cancer, a long-lived and highly plastic stem cell could serve as a potential cell-of-origin for this disease. This highlights the importance of determining the full landscape of MaSC populations and the factors regulating their recruitment.

The mammary stem cell niche: an elusive entity or dynamic force?

The ability of MaSCs to rapidly and faithfully respond to developmental and homeostatic demands throughout reproductive life may be attributable to their intimate association with a specific cellular microenvironment, known as the mammary **stem cell niche**. Stem cell niches can embody discrete and highly-specialized sites in certain tissues, e.g., the crypt base of the small intestine and the hair follicle [2]. Other tissues, including the post-pubescent mammary gland, prostate and lung, lack an easily-discernable niche, and stem cells in these organs may instead respond to more-ubiquitous tissue signals [2]. In any case, reciprocal interactions between MaSCs and their mature epithelial progeny, neighboring stromal cells and the supporting extracellular matrix, undoubtedly provide the autocrine, juxtracrine and paracrine signals that direct and adjust cell fate [19]. Extrinsic regulatory cues may include diffusible molecules (e.g., growth factors and cytokines) as well as mechanical forces (e.g., cell-cell and cell-matrix interactions) [53,54]. In this section, we outline designs of mammary stem and progenitor cell distribution in the pubescent, mature and secretory epithelium, discussing how the spatial arrangement of these cells may underpin the development and integrity of this highly dynamic tissue.

Architectural conceptions of a MaSC niche

The absence of a definitive molecular portrait of MaSCs, combined with uncertainties regarding their precise location within the post-pubescent mammary epithelium, has greatly impeded the analysis of prospective MaSC niches. Cell surface signatures that facilitate the isolation of mammary repopulating cells also provide little insight into the tissue-positional cues that direct cell behavior. Early transplantation and ultrastructural studies, however, did imply that mammary repopulating cells were distributed throughout the ductal epithelium [26–28,55], positing that MaSC niches may reside in a “suprabasal” location in the epithelial bilayer [27,56,57].

Although the precise location of stem cells within the post-pubescent breast remains unclear (Figure 1c), it is generally accepted that the TEBs of elongating ducts serve as a transient niche during puberty (Figure 1b) [5,19]. Thus, a comprehensive examination of signaling events in TEB-resident stem cells is expected to yield important insights into the pathways directing MaSC activity and fate, which may also be relevant in the post-pubescent gland. TEBs consist of an outer layer of cap cells that envelop multiple layers of inner body cells [3]. Cap and body cells are generally considered to be the precursors of mature basal and luminal epithelial

lineages, respectively [19]. Cap cells have also long been hypothesized to represent an enriched population of bi/multipotent MaSCs [27,58,59]. Indeed, the stem cell associated phosphatase gene *s-Ship*, which is exclusively expressed in cap cells during puberty, correlates with enhanced mammary repopulating capacity in limiting dilution transplantation assays [6]. In addition, *s-Ship*-expressing cap cells are strongly associated with the expression of Par3L, a protein related to the cell polarity regulator Par3, which is required for MaSC maintenance and ductal morphogenesis [60]. Recent mathematical modeling of mammary ductal elongation, however, suggests that inwardly-migrating cap cells do not contribute to the luminal epithelial lineage, as previously hypothesized [5]. Therefore, the precise contribution of these anatomically-distinct cells to ductal morphogenesis requires further investigation. The relationship between cap cells in the TEB and unipotent MaSCs, identified by genetic lineage-tracing [11,15,16,18], is also unclear. An answer to these important questions, and a potential unifying definition of physiological MaSC potency, awaits future inducible fate-mapping studies using transgenic s-SHIP and/or Par3L reporter models.

In the post-pubescent mammary gland, where TEBs have fully regressed, the location of MaSCs and their niche constituents is more ambiguous (Figure 1c). It is presumed that MaSCs, left behind by elongating TEBs during pubertal growth, are dispersed throughout the adult epithelial network. Here, hormonal cues stimulate further branching and the formation of alveolar-like buds and lobuloalveoli during estrous cycling and in pregnancy, respectively [61]. The notable absence of hormone receptors in mammary repopulating [62] and MaSC-enriched basal cell populations [63] implies that paracrine interactions between hormone receptor-expressing cells and stem cells guide tissue development and homeostasis [64–68]. Multiple paracrine signaling pathways, including Wnt, EGFR, IGFR and RANK signaling, are reported to regulate MaSC function downstream of hormone action. In addition, FGF, Hedgehog and Notch signaling have also been implicated in modulating MaSC fate during different stages of mammary gland development. How the local activities of these pathways are controlled by systemic changes in hormone levels, however, remains unknown [69,70]. Nevertheless, the widespread distribution of hormone receptor-positive cells throughout the adult mammary epithelial tree [16,71], suggests that MaSCs would be able to receive and integrate these paracrine signals at most architectural locations within the ductal epithelium. Moreover, alterations in the abundance and distribution of hormone receptor-positive cells with age [71], may reflect lifetime-dependent variations in a putative MaSC niche.

MaSCs are thought to survive tissue remodeling during post-lactational involution, enabling further cycles of expansion with each subsequent pregnancy (Figure 1c-f). It is therefore tempting to speculate that MaSCs reside in the vicinity of epithelial branch points, poised to generate the lateral branches and lobuloalveolar structures required for lactation. Fate-mapping studies using an alveolar-specific whey acidic protein (WAP)-driven Cre have also identified a population of long-lived parity induced-mammary epithelial cells (PI-MECs) that are sustained through multiple reproductive cycles [72]. These cells reside at ductal extremities in the

post-parous mammary gland, and contribute exclusively to the hormone receptor-negative luminal lineage in subsequent pregnancies [7,72]. Intriguingly, a recent single cell lineage-tracing study has revealed unequal distribution of MaSC progeny between lobuloalveolar units in lactating mammary tissue [16]. Thus, these striking observations also support a model whereby an alveolar stem cell niche is positioned near bifurcation sites in the mature ductal epithelium. Interestingly, increased MaSC activity during pregnancy correlates with the re-expression of *s-Ship* specifically in basal cells at the tips of alveolar buds, suggesting the emergence of a transient stem cell niche during lobuloalveologenesis [6].

MaSC niche dynamics

As described earlier in this review, distinct adult MaSCs are postulated to fulfil the proliferative and homeostatic demands of the mammary gland (Figure 4) [19]. The degree to which the heterogeneity in the MaSC compartment is intrinsic or a result of microenvironmental cues, however, is not known. A recent single cell lineage-tracing study, which employed quantitative volumetric analysis to determine the contribution of a single labeled MaSC to ductal morphogenesis, estimated that at least 35 lineage-restricted MaSCs actively and stochastically contribute to the development of each major duct during puberty [16]. A subsequent study, also using quantitative lineage-tracing at clonal density, put this number at 260 lineage-restricted MaSCs per TEB, leading to the suggestion that most TEB cells can function as lineage-committed MaSCs [18]. Discrepancies between these two studies may reflect differing functional definitions of MaSCs, and the quantitative and mathematical platforms and assumptions for analysis. Quantitative lineage-tracing studies also suggest that molecularly heterogeneous populations of TEB-resident MaSCs function as single equipotent pools, colonizing ductal branches through stochastic neutral drift dynamics [18]. Random segregation during successive rounds of TEB bifurcation mediates the unequal distribution of MaSC progeny between adjacent ductal structures, leading to clonal enrichment or extinction over time [18], supporting previous observations of clonal labeling patterns [16]. Furthermore, single cell lineage-tracing has shown that most lactational alveoli are comprised of the progeny of more than a single unipotent MaSC, indicating that a pool of lineage-restricted alveolar MaSCs also contribute to alveolar morphogenesis during pregnancy and lactation [16]. These early applications of quantitative and single-cell lineage-tracing approaches in the mammary gland [16,18] have provided unprecedented insights into clonal dynamics and stem/progenitor heterogeneity and multiplicity, heralding a new era in our investigation and understanding of normal and malignant stem cells in the breast.

Concluding Remarks

In this review we examined properties of potency, dynamics and plasticity in adult MaSCs, and the respective technologies that have underpinned key experimental observations. Whilst this area has received considerable attention over the last decade, many questions remain unanswered (see Outstanding Questions). At the center of this enquiry is whether MaSCs in the adult breast are unipotent, bipotent or something less discordant.

Stem cells are defined by their functional abilities, that is: proliferation, self-maintenance, production of a large number of differentiated progeny, tissue regeneration/repair, and a flexibility within these states [39]. The challenge thus far has been how to study a cell's functionality without inadvertently altering its function. Lineage-tracing has come a long way in this respect [11–16,18]. The refinement of lineage-tracing approaches and the application of other novel experimental models and methods for marking, visualizing and profiling individual cells (Box 1) will continue to provide important insights in this field. The question then becomes, what level of evidence is required to achieve a consensus?

Acknowledgements

This work was supported by the Medical Research Council (MR/J001023/1 to C.J.W. and B.L.-L.), Wellcome Trust (105377/Z/14/Z to O.B.H.), the National Health and Medical Research Council (1071074 to F.M.D.) and a University of Queensland Early Career Researcher Grant (UQECR1718865 to F.M.D.). We thank Dr. Silvia Fre for her valuable feedback on this manuscript, and apologize to all investigators whose work could not be cited owing to space limitations.

Glossary

Basal cell: one of the two main cell lineages in the mammary gland; basal cells surround the luminal cell layer and typically express cytokeratin-5, -14 and smooth muscle actin.

Bi/multipotent: able to give rise to more than one cell lineage, e.g., a bipotent MaSC may be able to give rise to both basal and luminal progeny.

Clone: All of the progeny of a single parent cell.

Label retaining cell: a cell that is able to retain a label (be it a lipophilic dye, DNA intercalating nucleoside or regulated-expression of a fluorescently-tagged histone) over a defined chase period. Cells that remain in cycle dilute the label, whereas slow-cycling or quiescent cells remain labeled at the end of the assay.

Lineage-tracing: a technique to identify the progeny of a single cell; the phrase "population-based lineage-tracing" has been used here to distinguish techniques that trace the progeny of specific populations of cells (e.g., cytokeratin-14-expressing cells) generally at levels higher than clonal density.

Luminal cell: one of the two main cell lineages in the mammary gland; luminal cells line the lumen of ducts and alveoli; they typically express cytokeratin-8 and may be hormone receptor positive or negative.

Mammary epithelial cell hierarchy: the organization of stem, progenitor and differentiated cells in the mammary gland.

Mammary repopulating cells: Cells enriched for the ability to regenerate the mammary epithelium upon serial transplantation at limiting dilution into the cleared fat pad of a recipient mouse.

Mammary stem cells (MaSCs): undifferentiated cells in the mammary gland that are capable of giving rise indefinitely to more stem cells (self-renewal) as well as to more-differentiated daughters through symmetric and asymmetric divisions. Uncertainties surrounding the identity, differentiation potential and plasticity of these cells has generated semantic debate, and MaSCs are also referred to more conservatively as "stem/progenitor cells".

Potential stem cell: a more differentiated cell that is able to re-acquire stem-like properties under regenerative/wounding conditions. Also known as facultative stem cells.

Stem cell niche: the specialized microenvironment in which a stem cell resides that can regulate stem cell self-renewal, differentiation and longevity.

Terminal end bud (TEB): bulbous proliferative structures at the ends of each main duct during puberty; the presumptive location of pubertal MaSCs.

Unipotent: able to give rise to one main cell lineage, e.g., a unipotent luminal stem cell is able to give rise only to luminal progeny and a unipotent basal stem cell is able to give rise only to basal progeny.

Figure Legends

Figure 1: **Postnatal mammary gland development in mice. a**) Mammals are born with only a rudimentary ductal structure (see [3] and [4] for a description of embryonic mammary gland development), which begins to elongate and invade the empty fat pad at puberty (**b**). By the end of puberty (**c**), the ductal structures have reached the boundaries of the mammary fat-pad and the TEBs have fully regressed. Mammary ducts are comprised of two epithelial cell lineages arranged into distinct cell layers; luminal cells line the lumen of each duct and are surrounded by an outer layer of basal cells (depicted inset). Whether MaSCs in the adult mammary gland are lineage-restricted or can give rise to both luminal and basal cells is area of contention. **c-e**) Resident MaSCs in the mature mammary epithelium are responsible for the generation of milk-producing alveoli during pregnancy and lactation. **f**) Stem cells are likely to survive post-lactational regression (involution) to enable successive pregnancies. The mouse is an excellent model for studying processes regulating human mammary gland development and tumorigenesis, however, key differences exist (see [83]). Notably, the human mammary gland is arranged in distinct lobes, each with a separate ductal structure and outlet.

Figure 2: A summary of the key discoveries in the field and the methodologies that enabled these advances. This timeline focuses on discoveries made within the last decade, using transplantation or genetic lineage-tracing assays. For a more detailed historical review see [19,20,84]. Schematic diagrams summarizing each *in vivo* methodology are depicted at puberty, however, these techniques have also been utilized to assess cell fate at other developmental stages, and in some cases their use has also been extended to investigate cellular dynamics in mammary tumorigenesis.

Figure 3: Limitations of population-based lineage-tracing studies. a) Clonal patterns arising from the genetic labeling of a single cell (purple). These studies demonstrate that progeny of a single marked cell can be distributed throughout the length of the ductal epithelium in a stochastic, interspersed labeling pattern. These patterns are likely to be caused by the proliferation of both labeled and unlabeled TEB-resident stem cells, which deposit their progeny throughout the epithelium during ductal elongation. Labeling patterns can extend more than 8 mm in linear length and comprise many side branches, highlighting the importance of performing 3D imaging and/or macro clone analysis. Scale bar: 0.2 mm. Adapted from Lloyd-Lewis et al. Breast Cancer Research (Springer Nature) [24]. A schematic representation of these labeling patterns in luminal and basal clones is shown in (**b**). The extensive and stochastic dispersion of stem cell progeny increase the likelihood of clone convergence in studies where labeling is performed above clonal density. Clone convergence is particularly evident when using a multi-color reporter gene. In the example here (b, bottom panel), it is difficult to distinguish whether luminal and basal blue cells came from a single bipotent precursor, or whether they arose from separate labeling events. Other technical limitations of population-based lineage-tracing approaches include periodic and promiscuous labeling by pathway-specific or pan-lineage promoters.

Figure 4: A working model of the mammary epithelial cell hierarchy. Multipotent MaSCs are present in the embryo. Although the exact stage of lineage-specification is not clear, postnatal mammary gland development (i.e. ductal and alveolar morphogenesis) is principally driven by unipotent luminal and basal MaSCs. Luminal stem cells give rise to ductal and alveolar cells that can be estrogen receptor (ER) positive or negative. The extent of sub-lineage diversity in the basal compartment, and whether there are distinct ductal and alveolar basal cells, is not yet clear. In addition to the cells responsible for building mammary ducts and alveoli under physiological conditions (left panel), various studies indicate that quiescent and potential stem cells may also reside within the adult mammary gland (right panel). Quiescent bi/multipotent MaSCs (not detected by quantitative or single cell lineage-tracing approaches) may remain in the mammary gland after embryonic development. Additionally, a plastic, intermediate cell type with properties similar to the basal cell lineage may be capable of reverting to a multipotent state under regenerative conditions. Lineage-restricted luminal and basal progenitors have also been shown to reacquire multipotency with oncogenic reprogramming. A holistic description of the cellular differentiation hierarchy in the mammary gland may need to accommodate aspects of plasticity.

References

1 Visvader, J.E. and Clevers, H. (2016) Tissue-specific designs of stem cell hierarchies. *Nat. Cell Biol.* 18, 349–355

2 O'Brien, L.E. and Bilder, D. (2013) Beyond the niche: tissue-level coordination of stem cell dynamics. *Annu. Rev. Cell Dev. Biol.* 29, 107–36

3 Macias, H. and Hinck, L. (2012) Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 533–557

4 Watson, C.J. and Khaled, W.T. (2008) Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* 135, 995–1003

5 Paine, I. *et al.* (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

6 Bai, L. and Rohrschneider, L.R. (2010) s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev.* 24, 1882–1892

7 Chang, T.H.-T. *et al.* (2014) New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. *Breast Cancer Res.* 16, R1

8 Stingl, J. *et al.* (2006) Purification and unique properties of mammary epithelial stem cells. *Nature* 439, 993–7

9 Shackleton, M. *et al.* (2006) Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84–8

10 Sleeman, K.E. *et al.* (2006) CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res.* 8, R7

11 Van Keymeulen, A. *et al.* (2011) Distinct stem cells contribute to mammary gland development and maintenance. *Nature* 479, 189–193

12 de Visser, K.E. *et al.* (2012) Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J. Pathol.* 228, 300–9

13 van Amerongen, R. *et al.* (2012) Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* 11, 387–400

14 Rios, A.C. *et al.* (2014) In situ identification of bipotent stem cells in the mammary gland. *Nature* 506, 322–7

15 Wuidart, A. *et al.* (2016) Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* 30, 1261–77

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

17 Wang, D. *et al.* (2015) Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* 517, 81–4

18 Scheele, C. *et al.* (2017) Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature*

19 Sreekumar, A. *et al.* (2015) The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. *Endocr. Relat. Cancer* 22, T161-76

20 Visvader, J.E. and Stingl, J. (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* 28, 1143–1158

21 Nassar, D. and Blanpain, C. (2016) Cancer stem cells: basic concepts and therapeutic implications. *Annu. Rev. Pathol. Mech. Dis.* 11, 47–76

22 DeOme, K.B. *et al.* (1959) Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.* 19, 515–20

23 Rios, A.C. *et al.* (2016) Essential role for a novel population of binucleated mammary epithelial cells in lactation. *Nat. Commun.* 7, 11400

24 Lloyd-Lewis, B. *et al.* (2016) Imaging the mammary gland and mammary tumours in 3D: optical tissue clearing and immunofluorescence methods. *Breast Cancer Res.* 18, 127

25 Sale, S. and Pavelic, K. (2015) Mammary lineage tracing: The coming of age. *Cell. Mol. Life Sci.* 72, 1577–1583

26 Hoshino, K. (1964) Regeneration and growth of quantitatively transplanted mammary glands of normal female mice. *Anat. Rec.* 150, 221–36

27 Smith, G.H. and Medina, D. (1988) A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. *J. Cell Sci.* 90 (Pt 1), 173–183

28 Daniel, C. *et al.* (1968) The in vivo life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. *Proc. Natl. Acad. Sci. U. S. A.* 61, 53–60

29 Kordon, E.C. and Smith, G.H. (1998) An entire functional mammary gland may comprise the progeny from a single cell. *Development* 125, 1921–30

30 Ehmann, U.K. *et al.* (1987) Cultured mouse mammary epithelial cells: normal phenotype after implantation. *J. Natl. Cancer Inst.* 78, 751–7.

31 Smith, G.H. (1996) Experimental mammary epithelial morphogenesis in an in vivo model: Evidence for distinct cellular progenitors of the ductal and lobular phenotype. *Breast Cancer Res. Treat.* 39, 21–31

32 Prater, M.D. *et al.* (2014) Mammary stem cells have myoepithelial cell properties. *Nat. Cell Biol.* 16, 942–50, 1–7

33 Van Keymeulen, A. *et al.* (2015) Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* 525, 119–123

34 Koren, S. *et al.* (2015) PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. *Nature* 525, 114–118

35 Rios, A.C. *et al.* (2016) The complexities and caveats of lineage tracing in the mammary gland.

Breast Cancer Res. 18, 116

36 Rodilla, V. *et al.* (2015) Luminal Progenitors Restrict Their Lineage Potential during Mammary Gland Development. *PLoS Biol.* 13,

37 Lafkas, D. *et al.* (2013) Notch3 marks clonogenic mammary luminal progenitor cells in vivo. *J. Cell Biol.* 203, 47–56

38 Kozar, S. *et al.* (2013) Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* 13, 626–33

39 Potten, C.S. and Loeffler, M. (1990) Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* 110, 1001–1020

40 Aloia, L. *et al.* (2016) Cellular plasticity in the adult liver and stomach. *J. Physiol.* 594, 4815–25

41 Cotsarelis, G. *et al.* (1990) Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* 61, 1329–37

42 Horsley, V. *et al.* (2008) NFATc1 balances quiescence and proliferation of skin stem cells. *Cell* 132, 299–310

43 Tumbar, T. *et al.* (2004) Defining the epithelial stem cell niche in skin. *Science* 303, 359–63

44 Buczacki, S.J.A. *et al.* (2013) Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495, 65–69

45 Ito, M. *et al.* (2005) Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nat. Med.* 11, 1351–1354

46 Boras-Granic, K. *et al.* (2014) Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Res.* 16, 487

47 Li, L. and Clevers, H. (2010) Coexistence of quiescent and active adult stem cells in mammals. *Science* 327, 542–5

48 Fuchs, E. and Horsley, V. (2011) Ferreting out stem cells from their niches. *Nat. Cell Biol.* 13, 513–518

49 dos Santos, C.O. *et al.* (2013) Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 110, 7123–30

50 Cai, S. *et al.* (2017) A quiescent Bcl11b high stem cell population is required for maintenance of the mammary gland. *Cell Stem Cell* 20, 247–60

51 Fu, N. *et al.* (2017) Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. *Nat. Cell Biol.* 19, 164–76

52 Giraddi, R.R. *et al.* (2015) Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat. Commun.* 6, 8487

53 Inman, J.L. *et al.* (2015) Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* 142, 1028–1042

54 Howard, B.A. and Lu, P. (2014) Stromal regulation of embryonic and postnatal mammary epithelial

development and differentiation. *Semin. Cell Dev. Biol.* 25–26, 43–51

55 Young, L.J.T. *et al.* (1971) The influence of host and tissue age on life span and growth rate of serially transplanted mouse mammary gland. *Exp. Gerontol.* 6, 49–56

56 Chepko, G. and Smith, G.H. (1997) Three division-competent, structurally-distinct cell populations contribute to murine mammary epithelial renewal. *Tissue Cell* 29, 239–253

57 Chepko, G. and Dickson, R.B. (2003) Ultrastructure of the putative stem cell niche in rat mammary epithelium. *Tissue Cell* 35, 83–93

58 Williams, J.M. and Daniel, C.W. (1983) Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.* 97, 274–90

59 Srinivasan, K. *et al.* (2003) Netrin-1/neogenin interaction stabilizes multipotent progenitor cap cells during mammary gland morphogenesis. *Dev. Cell* 4, 371–382

60 Huo, Y. and Macara, I.G. (2014) The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. *Nat. Cell Biol.* 16, 529–537

61 Brisken, C. and O'Malley, B. (2010) Hormone action in the mammary gland. *Cold Spring Harb. Perspect. Biol.* 2, a003178

62 Sleeman, K.E. *et al.* (2007) Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. *J. Cell Biol.* 176, 19–26

63 Asselin-Labat, M.-L. *et al.* (2006) Steroid hormone receptor status of mouse mammary stem cells. *J. Natl. Cancer Inst.* 98, 1011–1014

64 Beleut, M. *et al.* (2010) Two distinct mechanisms underlie progesterone-induced proliferation in the mammary gland. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2989–94

65 Asselin-Labat, M.-L. *et al.* (2010) Control of mammary stem cell function by steroid hormone signalling. *Nature* 465, 798–802

66 Cai, C. *et al.* (2014) R-spondin1 is a novel hormone mediator for mammary stem cell self-renewal. *Genes Dev.* 28, 2205–2218

67 Rajaram, R.D. *et al.* (2015) Progesterone and Wnt4 control mammary stem cells via myoepithelial crosstalk. *EMBO J.* 34, 641–652

68 Joshi, P. a *et al.* (2010) Progesterone induces adult mammary stem cell expansion. *Nature* 465, 803–7

69 Rosen, J.M. and Roarty, K. (2014) Paracrine signaling in mammary gland development: what can we learn about intratumoral heterogeneity? *Breast Cancer Res.* 16, 202

70 Brisken, C. and Ataca, D. (2015) Endocrine hormones and local signals during the development of the mouse mammary gland. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 181–195

71 Ismail, P.M. *et al.* (2002) A novel LacZ reporter mouse reveals complex regulation of the progesterone receptor promoter during mammary gland development. *Mol. Endocrinol.* 16, 2475–89

72 Wagner, K.-U. *et al.* (2002) An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. *Development* 129, 1377–86

73 Tainaka, K. *et al.* (2014) Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* 159, 911–924

74 Ke, M.-T. *et al.* (2013) SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nat. Neurosci.* 16, 1154–61

75 Davis, F.M. *et al.* (2015) Essential role of Orai1 store-operated calcium channels in lactation. *Proc. Natl. Acad. Sci. U. S. A.* 112, 5827–32

76 Ritsma, L. *et al.* (2013) Surgical implantation of an abdominal imaging window for intravital microscopy. *Nat. Protoc.* 8, 583–94

77 Zomer, A. *et al.* (2013) Intravital imaging of cancer stem cell plasticity in mammary tumors. *Stem Cells* 31, 602–6

78 Kumar, P. *et al.* (2017) Understanding development and stem cells using single cell-based analyses of gene expression. *Development* 144, 17–32

79 McKenna, A. *et al.* (2016) Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* 353, aaf7907

80 Perli, S.S.D. *et al.* (2016) Continuous genetic recording with self-targeting CRISPR-Cas in human cells. *Science* 511, 53058

81 Clevers, H. (2016) Modeling Development and Disease with Organoids. *Cell* 165, 1586–97

82 Jardé, T. *et al.* (2016) Wnt and Neuregulin1/ErbB signalling extends 3D culture of hormone responsive mammary organoids. *Nat. Commun.* 7, 13207

83 Cardiff, R.D. and Wellings, S.R. (1999) The comparative pathology of human and mouse mammary glands. *J. Mammary Gland Biol. Neoplasia* 4, 105–122

84 Oakes, S.R. *et al.* (2014) The mammary cellular hierarchy and breast cancer. *Cell. Mol. Life Sci.* 71, 4301–4324