Stem Cell Reports
Article

ISSCR

OPEN ACCESS

A Long-Lived Luminal Subpopulation Enriched with Alveolar Progenitors Serves as Cellular Origin of Heterogeneous Mammary Tumors

Luwei Tao,[1,2,3] Maaike P.A. van Bragt,[1,2,3] and Zhe Li[1,2,*]
[1]Division of Genetics, Brigham and Women's Hospital (BWH), Boston, MA 02115, USA
[2]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
[3]Co-first author
*Correspondence: zli4@rics.bwh.harvard.edu
http://dx.doi.org/10.1016/j.stemcr.2015.05.014

SUMMARY

It has been shown that the mammary luminal lineage could be maintained by luminal stem cells or long-lived progenitors, but their identity and role in breast cancer remain largely elusive. By lineage analysis using *Wap-Cre* mice, we found that, in nulliparous females, mammary epithelial cells (MECs) genetically marked by *Wap-Cre* represented a subpopulation of $CD61^+$ luminal progenitors independent of ovarian hormones for their maintenance. Using a pulse-chase lineage-tracing approach based on *Wap-Cre* adenovirus (*Ad-Wap-Cre*), we found that *Ad-Wap-Cre*-marked nulliparous MECs were enriched with $CD61^+$ alveolar progenitors (APs) that gave rise to $CD61^-$ alveolar luminal cells during pregnancy/lactation and could maintain themselves long term. When transformed by different oncogenes, they could serve as cells of origin of heterogeneous mammary tumors. Thus, our study revealed a type of long-lived AP within the luminal lineage that may serve as the cellular origin of multiple breast cancer subtypes.

INTRODUCTION

Breast cancer is heterogeneous and can be divided into at least six molecular subtypes with distinct clinical outcomes (Perou et al., 2000; Prat et al., 2010; Sorlie et al., 2003). The current hypothesis is that this heterogeneity mainly results from the interplay of different cells of origin and transformation events (Visvader, 2011). Studies in both human cells and murine models have implied that basal mammary epithelial cells (MECs), luminal progenitors (LPs), and more mature luminal MECs may be cells of origin of the Claudin-low, Basal-like, and Luminal subtypes of breast cancers, respectively (Skibinski and Kuperwasser, 2015). However, definitive evidence to support this remains largely elusive. As target cells of breast tumorigenesis, MECs are composed of the luminal and basal/myoepithelial lineages. The luminal lineage can be further divided into a ductal and an alveolar sub-lineage. Studies based on cleared mammary fat pad transplantation have shown that a single multipotent mammary stem cell (MaSC), which is a basal MEC, could reconstitute the entire mammary epithelium in the recipient (Shackleton et al., 2006; Stingl et al., 2006). Transplantation studies based on limiting dilutions of MECs also revealed the presence of duct-limited progenitors, which produced both ductal luminal and ductal myoepithelial cells as well as lobule-limited progenitors that gave rise to secretory alveolar luminal cells (ALs) and alveolar myoepithelial cells (Kordon and Smith, 1998; Smith, 1996). In support of the presence of lobule-limited biopotent progenitors, a population of lobule-restricted cells termed parity-induced MECs (PI-MECs) was defined previously based on genetic marking using the *Wap-Cre* (*Whey Acidic Protein* [*Wap*] promoter-driven Cre expression) transgenic mouse and a conditional Cre-reporter (*Rosa26-loxP-Stop-loxP-lacZ* [*R26lz*]) (Figure 1A; Boulanger et al., 2005; Wagner et al., 2002). Specifically, PI-MECs were identified in parous females as $lacZ^+$ alveolar cells that survived involution and could serve as multipotent mammary stem/progenitor cells to produce new secretory acini, composed of both secretory luminal and myoepithelial cells, in the subsequent pregnancies; furthermore, transplantation assays revealed that PI-MECs were multipotent and could self-renew over multiple rounds of transplantations (Boulanger et al., 2005).

Transplantation assays, which measure the potential of cells under investigation, traditionally were used as the gold standard for characterizing tissue stem cells. However, recent studies in epithelial tissues based on pulse-chase lineage-tracing approaches, which measure the activity of primary cells in their native habitats, often revealed a discrepancy in the phenotype of an epithelial stem cell population when different assays were used (Choi et al., 2012; Van Keymeulen et al., 2011; Wang et al., 2013). In the mammary gland (MG), lineage-tracing studies based on MEC lineage-specific *CreER* (Cre-estrogen receptor fusion) mice demonstrated that adult luminal and basal lineages are largely self-sustained (Prater et al., 2014; van Amerongen et al., 2012; Van Keymeulen et al., 2011); in particular, in unperturbed tissues, the luminal lineage appears to be maintained largely by its own lineage-restricted luminal stem cells (LuSCs), rather than by basal MaSCs (Van Keymeulen et al., 2011). However, more recent

lineage-tracing studies provided new evidence of the existence of both bipotent basal MaSCs and distinct long-lived LPs to support homeostasis of the luminal lineage in the physiological setting (Rios et al., 2014; Wang et al., 2015). Collectively, these studies reflect the intrinsic complexity of the MEC hierarchy that operates in vivo and a need to further dissect the MEC hierarchy in the physiological setting. Furthermore, since most promoters used in previous lineage-tracing studies either target a broad range of MECs or are subject to developmental stage-dependent regulation, it is difficult to use these genetic tools to directly determine the cell of origin of breast cancer and how the cellular origin contributes to breast cancer heterogeneity.

In this study, we analyzed a narrower MEC lineage (i.e., the alveolar luminal sub-lineage) defined based on genetic marking by Cre expression controlled by the same *Wap* promoter (Wagner et al., 1997), a MEC-specific promoter frequently used in breast cancer mouse models. We provide evidence for a long-lived luminal MEC subpopulation enriched with alveolar-committed progenitors that may serve as the cell of origin of multiple breast cancer subtypes.

RESULTS

Wap-Cre-Marked Cells in Nulliparous MGs Are Luminal MECs Expressing LP Markers

PI-MECs, as the name suggests, were originally identified in parous females as *Wap-Cre*-marked multipotent mammary stem/progenitor cells that might be derived from the de-differentiation of secretory alveolar cells (Boulanger et al., 2005; Wagner et al., 2002). Later it was found that such cells (i.e., *lacZ*$^+$ MECs genetically marked by *Wap-Cre*) also were present in nulliparous (virgin) MGs (thus, the term "PI-MECs" now refers to parity-*identified* MECs) (Booth et al., 2007). However, these *lacZ*$^+$ cells from the virgin gland were only detected in mammary explant cultures in the presence of insulin, so it was uncertain whether they were also present in nulliparous *Wap-Cre*;*R26lz* MGs in vivo. It has been well documented that the *Wap* promoter is active at estrus, but not at diestrus, in virgin females (e.g., based on a *Wap-lacZ* transgenic mouse model [Robinson et al., 1995]). However, *lacZ*$^+$ MECs detected in *Wap-lacZ* females are *not equivalent* to *lacZ*$^+$ MECs found in *Wap-Cre*;*R26lz* females, as the latter only require *transient* expression of Cre (to irreversibly turn on the *lacZ* reporter) from the *Wap* promoter, and they can be either long-lived cells (with transient *Wap* expression) or progeny cells that inherit the activated *lacZ* allele from their *Wap*$^+$ parental MECs. We reasoned that if some of these *Wap-Cre*-marked MECs in virgin glands are stem cells or long-lived progenitors, they might be detected by more sensitive genetic-marking approaches (due to their long-lived nature).

To test this hypothesis, we crossed *Wap-Cre* mice to a conditional Cre-reporter, *Rosa26-loxP-Stop-loxP-YFP* (*R26YFP*) (Figure 1A), which allows robust lineage-tracing and clonal analysis of MECs in vivo by immunostaining and fluorescence-activated cell sorting (FACS) (Tao et al., 2014; Van Keymeulen et al., 2011). In *Wap-Cre*;*R26YFP* (but not in *R26YFP*-only) adult virgin females, we could consistently detect a small population of YFP$^+$ MECs, although under the FVB background the percentage of YFP$^+$ MECs was at least 10-fold higher than that under the C57Bl/6 background (Figures 1B, S1A, and S1B). This profound strain difference might partially explain the difficulty in consistently detecting the *Wap-Cre*-marked MEC population in virgin females previously (Booth et al., 2007; Boulanger et al., 2005; Matulka et al., 2007; Wagner et al., 2002). Upon gating for Lin$^-$YFP$^+$ (Lin: lineage markers) cells, we found that these YFP$^+$ cells were almost entirely within the Lin$^-$CD24^{hi}CD29^{med} luminal gate (Figures 1C and S1C); the majority of them were positive for all known LP markers, including CD61, CD14, C-KIT, and CD49B (Asselin-Labat et al., 2007, 2011; Li et al., 2009), and they were negative for mature luminal cell (ML, i.e., estrogen receptor [ER]-positive cells) markers, including SCA-1 and CD133 (Prominin 1) (Sleeman et al., 2007; Figures 1C and S1D). Thus, these YFP-marked MECs appear to largely overlap with the SCA-1$^-$CD49B$^+$ LP subpopulation described recently, which is analogous to the ALDH$^+$ LP subpopulation in the human breast (Shehata et al., 2012).

Immunofluorescence (IF) staining for YFP revealed that these YFP$^+$ cells were indeed K8$^+$K14$^-$ luminal MECs (Figure 1D). Immunohistochemistry (IHC) staining and clonal analysis further revealed that most YFP$^+$ cells were localized in mammary ducts and alveoli as single-cell or two-cell clones interspersed by YFP$^-$ luminal MECs (Figures 1E, 1F, and S1E), whereas larger clones with more than two YFP$^+$ cells were often found in alveoli (Figure 1F). FACS isolation of these YFP$^+$ cells and subsequent IF staining further confirmed that they were exclusively K8$^+$K14$^-$K5$^-$ luminal MECs (Figure 1G). Since Lin$^-$CD24$^+$CD133$^+$ (or SCA-1$^+$) MECs were shown enriched with ER$^+$ cells (Sleeman et al., 2007), and a small portion of YFP$^+$ cells from *Wap-Cre*;*R26YFP* virgin females were CD133$^+$ (and SCA-1$^+$), we sorted the CD133$^+$ or CD133$^-$ subsets from YFP$^+$ or YFP$^-$ MECs from the same mice, and, by qRT-PCR analysis, we found that YFP$^+$CD133$^+$ MECs exhibited a low level of *Esr1* expression comparable to that of CD133$^-$ MECs, which are known to be enriched in ER$^-$ MECs (Sleeman et al., 2007; Figure 1H). Together, these data suggest that *Wap-Cre* genetically marks a subpopulation of MECs in nulliparous mammary ducts and alveoli that are exclusively luminal cells, express all known cell surface markers of LPs, and are ER negative.

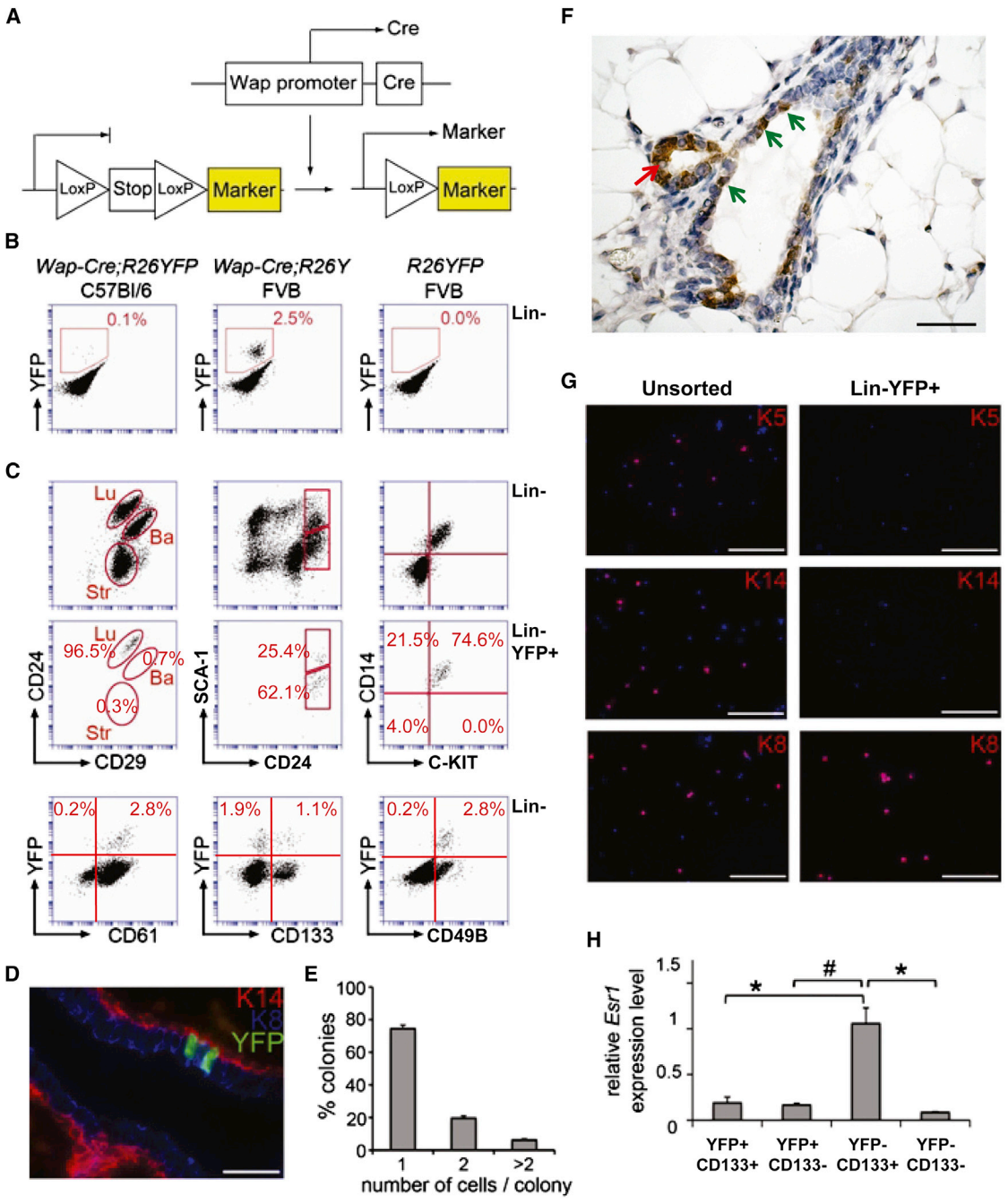


Figure 1. WVs Are Luminal Cells Expressing LP Markers

(A) Schematic diagram shows Cre-mediated excision of a floxed *Stopper* cassette (Stop) in a conditional reporter leads to expression of the reporter (Marker, *YFP* or *lacZ*); the cell-type specificity is achieved through the promoter (e.g., *Wap*) that controls Cre expression.

(B) FACS plots show percentages of YFP-marked Lin$^-$ cells in MGs isolated from *Wap-Cre;R26YFP* virgin females on distinct genetic backgrounds (C57Bl6 [left], FVB [middle], or *R26YFP* control [FVB, right]). See also Figure S1B.

(C) FACS profiles of YFP-marked Lin$^-$ cells isolated from adult *Wap-Cre;R26YFP* virgin females (FVB) for the indicated markers are shown. Lu, luminal gate; Ba, basal gate; Str, stromal cell gate. FACS gating strategy is shown in Figure S1C; detailed quantification of the percentage of each gate is shown in Figure S1D.

(D) IF staining for YFP on MG sections of adult *Wap-Cre;R26YFP* virgin females is shown. Scale bar, 50 μm.

(legend continued on next page)

Wap-Cre-Marked Nulliparous MECs Represent a Subpopulation of LPs Independent of Ovarian Hormones for Their Maintenance

To further determine the identity of *Wap-Cre*-marked MECs in virgin females (referred to as "WVs") at the molecular level, we sorted YFP$^+$ WVs as well as MaSCs, LPs, and MLs (MEC subsets defined based on Lim et al. [2010]) from *Wap-Cre*;*R26YFP* virgin females; as a control, we also sorted YFP$^+$ MECs from *Wap-Cre*;*R26YFP* females at mid-gestation, which are enriched with differentiating CD61$^-$ ALs (Figures S2A and S2B). We then profiled these sorted MEC subsets by microarray. By principal-component analysis (PCA), we found that the cluster of WVs exhibited the closest distance to that of LPs (Figure 2A). To validate this observation, we selected several marker genes that were shown previously to mark different MEC subsets and are conserved between human and mouse (Lim et al., 2010), and we confirmed by qRT-PCR that WVs exhibited an expression pattern similar to that of the bulk of LPs (Figure 2B). Since LPs were sorted based on CD61 (Lim et al., 2010), they appeared to represent a mixed population of heterogeneous progenitor cells for ductal luminal cells and ALs (Shehata et al., 2012). To determine whether WVs represent a subpopulation of the heterogeneous CD61$^+$ LPs, we compared the microarray expression profile of WVs to that of LPs. By gene set enrichment analysis (GSEA) (Subramanian et al., 2005), we found that, compared to the bulk of CD61$^+$ LPs, gene ontology (GO) terms related to activities of receptors (in plasma membrane) and responses to external stimuli, as well as those related to amino acid metabolism, were enriched in WVs (Figure 2C), whereas those related to RNA processing were downregulated in WVs (i.e., WVs may be less active compared to the bulk of LPs) (Figure 2D). Together, these data suggest that WVs may represent a subpopulation of CD61$^+$ LPs.

Our strategy to visualize WVs is based on Cre expression from the *Wap-Cre* transgene in nulliparous MGs; once they are genetically marked by YFP, the *Wap* promoter activity is not needed any more. Although the *Wap* promoter is a well-known promoter that responds to steroid hormones (Robinson et al., 1995), this does not necessarily indicate that maintenance of WVs also is dependent on ovarian hormones. To test this notion, we performed bilateral ovariectomy on *Wap-Cre*;*R26YFP* adult virgin females. We found that, after ovariectomy, a population of YFP$^+$ MECs similar to that in intact females could still be detected, and these ovarian hormone-independent YFP$^+$ MECs remained in the luminal gate (Figures S2C–S2E). These data suggest that ovarian hormones are not required for the maintenance of WVs. A previous study demonstrated that PI-MECs analyzed in parous females also share this property (Matulka et al., 2007).

Ad-Wap-Cre-Marked Nulliparous MECs Contain Long-Lived Progenitors that Exhibit Spontaneous Alveolar Differentiation in Virgin Females

In *Wap-Cre*;*R26YFP* females, since the *Wap-Cre* transgene is constitutive, it is not possible to perform longitudinal studies to determine the lineage relationship between ductal and alveolar localized WVs in nulliparous MGs, as well as that between WVs and YFP$^+$ CD61$^-$ ALs in pregnant and lactating females. A pulse-chase lineage-tracing system that can allow us to control the timing of Cre expression from the *Wap* promoter is necessary. To achieve this, we utilized a pulse-chase lineage-tracing system we established recently based on intraductal injection of lineage-specific Cre-expressing adenovirus (*Ad-Cre*) (Tao et al., 2014). Adenovirus is a DNA virus that does not integrate into the host genome and the adenoviral vector we use only leads to transient protein expression; thus, it serves a similar purpose as an inducible Cre system for pulse-chase lineage tracing.

We generated an *Ad-Cre* virus using the same *Wap* promoter as in *Wap-Cre* mice (Wagner et al., 1997) (*Ad-Wap-Cre*). We injected *Ad-Wap-Cre* into mammary ducts of *R26YFP* females at the virgin stage (pulse labeling at ~2 months of age) and analyzed YFP-marked MECs in virgin females at several time points thereafter (Figure 3A). To determine the specificity of genetic marking by *Ad-Wap-Cre*, we sorted YFP$^+$ MECs 3 days after injection (i.e., freshly labeled WVs). As an internal control, we also sorted CD61$^+$ LPs (i.e., the bulk of the LP population) from the same animals. By qRT-PCR, we found that WVs expressed a higher level of the endogenous *Wap* than LPs (Figure 3B),

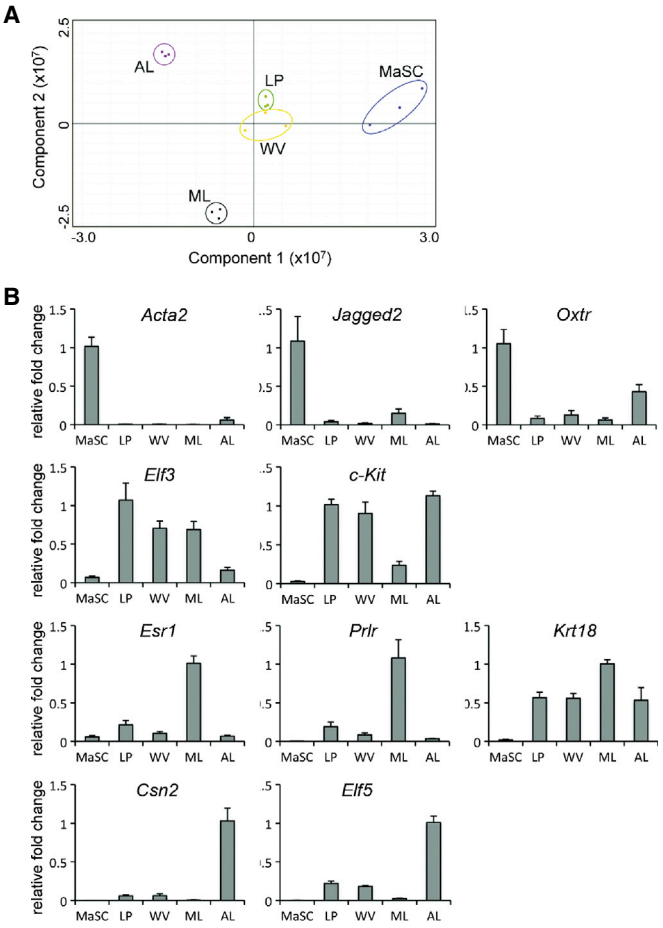
(E) Percentages of YFP-marked clones composed of single cell, doublet, or more than two cells quantified from YFP-IHC-stained MG sections of *Wap-Cre*;*R26YFP* females are shown (n = 4 mice). Data represent mean ± SEM.

(F) IHC (brown cells, YFP$^+$ cells) staining for YFP on MG sections of adult *Wap-Cre*;*R26YFP* virgin females. Red arrow indicates an alveolar structure budding out from a mammary duct, and most luminal MECs in it are YFP$^+$. Green arrows indicate interspersed YFP$^+$ luminal cells. Scale bar, 20 μm.

(G) IF staining for K5, K14, and K8 on Lin$^-$YFP$^+$ cells sorted from *Wap-Cre*;*R26YFP* virgin females is shown; unsorted cells from the same MGs were used as the control. Scale bars, 100 μm.

(H) Expression levels of *Esr1* in the indicated populations sorted from adult *Wap-Cre*;*R26YFP* virgin females are shown (n = 3 mice). Data represent mean ± SEM (*$p < 0.001$, $^{\#}p < 0.005$).

See also Figure S1.



C

RANK	NAME	NES	NOM p-val	FDR q-val
1	RESPONSE_TO_WOUNDING	2.12	0	0.0086
2	HEMATOPOIETIN_INTERFERON_CLASSD200_DOMAIN_CYTOKINE_RECEPTOR_ACTIVITY	2.12	0	0.0049
3	IMMUNE_RESPONSE	2.10	0	0.0057
4	RESPONSE_TO_EXTERNAL_STIMULUS	2.10	0	0.0043
5	INFLAMMATORY_RESPONSE	2.05	0	0.0102
6	DEFENSE_RESPONSE	2.03	0	0.0128
7	AMINO_ACID_METABOLIC_PROCESS	2.00	0	0.0163
8	NITROGEN_COMPOUND_CATABOLIC_PROCESS	2.00	0	0.0144
9	IMMUNE_SYSTEM_PROCESS	1.99	0	0.0147
10	OXYGEN_BINDING	1.97	0	0.0206
11	INTEGRAL_TO_PLASMA_MEMBRANE	1.96	0	0.0206
12	INTRINSIC_TO_PLASMA_MEMBRANE	1.95	0	0.0218
13	MULTICELLULAR_ORGANISMAL_DEVELOPMENT	1.94	0	0.0224
14	PLASMA_MEMBRANE_PART	1.94	0	0.0210
15	AMINO_ACID_AND_DERIVATIVE_METABOLIC_PROCESS	1.93	0	0.0245
16	PLASMA_MEMBRANE	1.92	0	0.0281
17	RECEPTOR_ACTIVITY	1.91	0	0.0291
18	EXTRACELLULAR_REGION	1.90	0	0.0291
19	LIPID_RAFT	1.89	0	0.0326
20	LOCOMOTORY_BEHAVIOR	1.89	0	0.0324
21	INTRINSIC_TO_MEMBRANE	1.88	0	0.0326
22	EXTRACELLULAR_SPACE	1.88	0	0.0322
23	AMINE_CATABOLIC_PROCESS	1.88	0.00660066	0.0314
24	INTEGRAL_TO_MEMBRANE	1.88	0	0.0302
25	AMINE_METABOLIC_PROCESS	1.87	0	0.0341

D

RANK	NAME	NES	NOM p-val	FDR q-val
1	MRNA_PROCESSING_GO_0006397	-2.06	0	0.0702
2	STRUCTURAL_CONSTITUENT_OF_RIBOSOME	-2.02	0	0.0548
3	RIBONUCLEOPROTEIN_COMPLEX_BIOGENESIS_AND_ASSEMBLY	-1.98	0	0.0689
4	RIBONUCLEOPROTEIN_COMPLEX	-1.96	0	0.0651
5	RNA_PROCESSING	-1.96	0	0.0535
6	NUCLEAR_PORE	-1.95	0	0.0480
7	ENDONUCLEASE_ACTIVITY_GO_0016893	-1.92	0.0023	0.0548
8	NUCLEOLAR_PART	-1.91	0.0023	0.0573
9	RIBOSOME_BIOGENESIS_AND_ASSEMBLY	-1.89	0	0.0655
10	CYTOSKELETON_DEPENDENT_INTRACELLULAR_TRANSPORT	-1.88	0.0048	0.0627
11	NUCLEOLUS	-1.87	0	0.0608
12	RNA_SPLICING	-1.85	0	0.0725
13	RRNA_PROCESSING	-1.84	0.0023	0.0742
14	DNA_INTEGRITY_CHECKPOINT	-1.83	0	0.0774
15	RNA_SPLICINGVIA_TRANSESTERIFICATION_REACTIONS	-1.81	0.0076	0.0886
16	TRNA_METABOLIC_PROCESS	-1.79	0.0118	0.0981
17	CONDENSED_CHROMOSOME	-1.79	0.0025	0.0947
18	ACTIN_FILAMENT_BASED_MOVEMENT	-1.74	0.0069	0.1446
19	PORE_COMPLEX	-1.74	0.0078	0.1373
20	RIBONUCLEASE_ACTIVITY	-1.73	0.0117	0.1382
21	DNA_DAMAGE_CHECKPOINT	-1.71	0.0096	0.1551
22	CONDENSED_NUCLEAR_CHROMOSOME	-1.69	0.0112	0.1789
23	SMALL_NUCLEAR_RIBONUCLEOPROTEIN_COMPLEX	-1.68	0.0149	0.1783
24	SPLICEOSOME	-1.68	0.0102	0.1796
25	ENDORIBONUCLEASE_ACTIVITY	-1.67	0.0231	0.1837

Figure 2. WVs Represent a Subpopulation of CD61$^+$ LPs

(A) PCA of microarray data for distinct subsets of MECs shows WVs clustered more closely with LPs.

(B) qRT-PCR analysis confirmed the expression pattern of a number of well-known markers for different sorted MEC subsets (n = 4 mice). WVs, YFP-marked MECs from adult *Wap-Cre;R26YFP* virgin females; MaSC, mammary stem cell; LP, luminal progenitor; MLs, mature luminal cells; ALs, alveolar luminal cells (i.e., YFP-marked MECs from *Wap-Cre;R26Y* mice at mid-gestation). Data represent mean ± SEM.

(C and D) GSEA shows the top 25 enriched GO terms in WVs (C) and LPs (D), respectively. NES, normalized enrichment score; NOM p-val, nominal p value. False discovery rate (FDR) q value $\leq$ 0.25 was considered as significant.

See also Figure S2.

suggesting that the initial genetic marking is based on the specific activity of the *Wap* promoter, rather than non-specific labeling. To determine whether newly labeled WVs are similar to WVs in *Wap-Cre;R26YFP* virgin females at the molecular level, we measured expression levels of select markers in WVs (2 weeks post-injection) by qRT-PCR and found that they exhibited a similar LP gene expression pattern (Figure S3). By clonal analysis, we found that, 3 days after *Ad-Wap-Cre* injection, the majority of YFP$^+$ clones were found in mammary ducts (Figure 3C, ~75% in mammary ducts and ~25% in alveoli), and most of them were small clones (Figure 3D). Interestingly, after a long-term chase (4.5 months), we found that more YFP$^+$ clones were found in alveoli than in mammary ducts (Figure 3C, ~40% in mammary ducts and ~60% in alveoli), and most of the alveolar-localized clones were large clones (Figures 3D–3H). These data suggest that freshly labeled WVs contain long-lived progenitors committed to the alveolar lineage (i.e., alveolar progenitors [APs]).

Ad-Wap-Cre-Marked Nulliparous MECs Are Enriched with APs that Give Rise to Differentiating ALs during Pregnancy/Lactation and Are Self-Sustained over the Long Term

Although in virgin females we showed that freshly labeled WVs had a tendency to undergo spontaneous alveolar

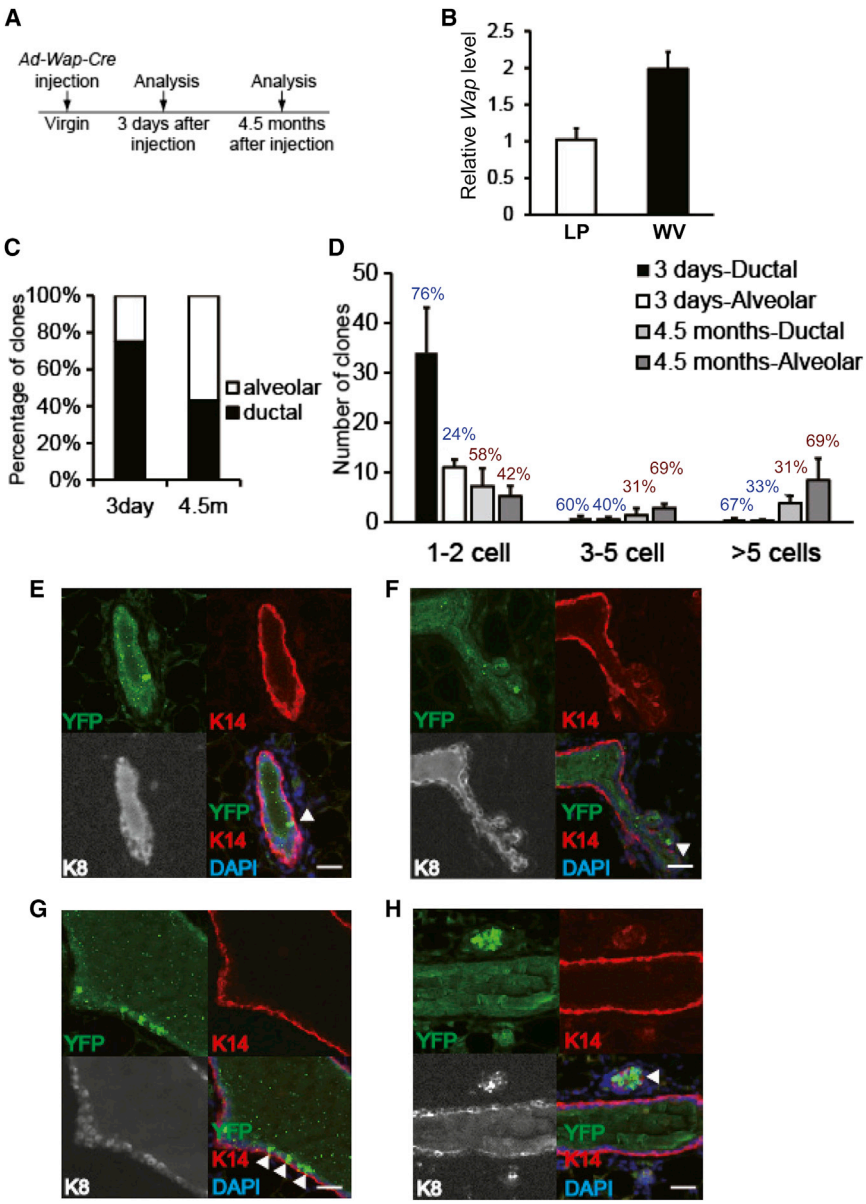


Figure 3. Freshly Labeled WVs Contain Long-Lived Progenitors that Exhibit Spontaneous Alveolar Differentiation in Nulliparous Females

(A) Timeline of pulse-chase lineage tracing by *Ad-Wap-Cre* intraductal injection into *R26YFP* nulliparous MGs. Animals were kept as virgins.

(B) qRT-PCR analysis shows higher expression of the endogenous *Wap* in freshly labeled WVs (WV) in relation to the bulk of $CD61^+$ LPs (LP) (pooled sorted cells from ten injected MGs from five mice). Expression was normalized to *Actb*. Data represent mean ± SEM (from three technical replicates).

(C and D) Relative percentages of YFP^+ ductal and alveolar clones (total = 100%) 3 days or 4.5 months (m) after injection (C) and distributions of small (1–2 cells), intermediate (3–5 cells), and large (>5 cells) YFP^+ clones in the indicated MG sections (D). Relative percentages of YFP^+ ductal and alveolar clones at each indicated chase time point (total = 100%) are shown. Data represent mean ± SEM; n = 4 (3 days) and 5 (4.5 months) injected MGs from three mice each.

(E–H) Representative IF pictures show ductal (E and G) or alveolar (F and H)-localized YFP^+ clones (arrowheads) 3 days (E and F) or 4.5 months (G and H) after injection.

See also Figure S3.

differentiation (Figures 3C–3H), we reasoned that their full potential would only be revealed upon pregnancy and lactation. We therefore injected *Ad-Wap-Cre* into mammary ducts of 2-month-old *R26YFP* virgin females and then allowed them to go through one or multiple rounds of pregnancy and lactation. We analyzed the YFP-marked lineage in them at different stages of MG development (Figure 4A). At both mid-gestation and lactation stages, by clonal analysis, we observed individual alveoli composed of almost entirely YFP^+ MECs (Figures 4B and 4C), and only ALs (but not alveolar myoepithelial cells) were marked by YFP (Figures S4A and S4B). These data suggest that at least a portion of freshly labeled WVs at the virgin stage are APs that clonally give rise to ALs during pregnancy and lactation.

Since many freshly marked WVs are localized in mammary ducts (Figures 3C–3H), it is uncertain whether such ductal-localized WVs are APs that produce ALs during pregnancy. To determine this, we generated a cohort of *R26YFP* females by injecting *Ad-Wap-Cre* and *Ad-K8-Cre* viruses to the contralateral sides of inguinal MGs in the same mouse, respectively, at the virgin stage. K8 (Keratin 8) is a pan-luminal marker and we showed previously that intraductal injection of *Ad-K8-Cre* led to genetic marking of the entire luminal lineage, including both ductal progenitors and APs (Tao et al., 2014). To account for the difference

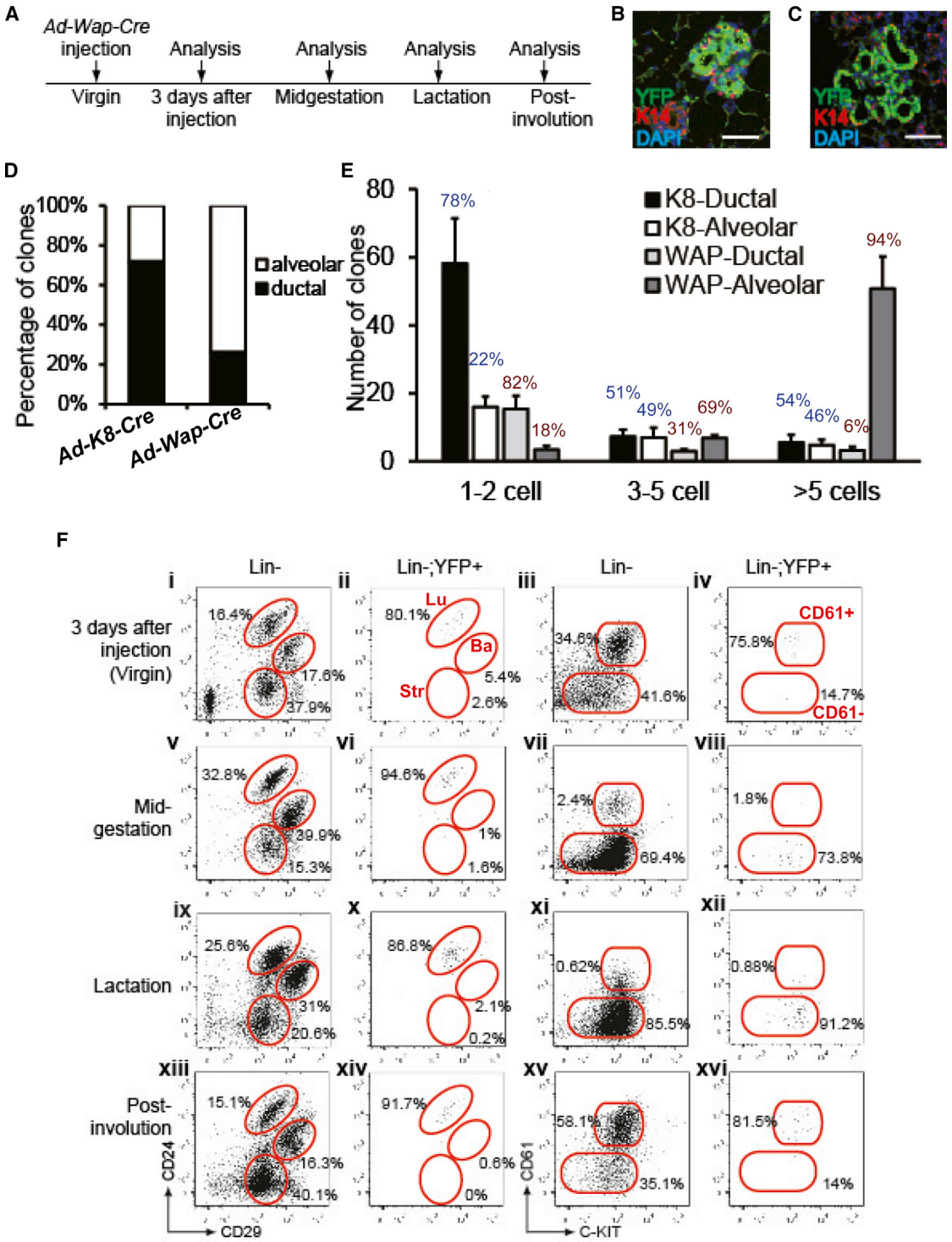


Figure 4. Freshly Labeled WVs Are Enriched with Self-Sustained, Clonogenic APs

(A) Timeline shows pulse-chase lineage tracing by *Ad-Wap-Cre* intraductal injection into *R26YFP* nulliparous MGs.

(B and C) IF staining of MG sections shows alveoli formed by YFP$^+$ cells (induced by *Ad-Wap-Cre* injection at the virgin stage) during mid-gestation (B) and lactation (C). Scale bars, 50 μm. Note only K8$^+$ ALs, but not K14$^+$ myoepithelial cells, were labeled by YFP (see also Figures S4A and S4B).

(legend continued on next page)

in the intrinsic labeling efficiency of different adenoviruses, we titrated *Ad-K8-Cre* so that the initial labeling (measured as the percentage of YFP$^+$ cells) was similar to that of *Ad-Wap-Cre*. Upon chasing to mid-gestation, we found that the majority of YFP$^+$ clones from *Ad-K8-Cre* injection were ductally localized, whereas most YFP$^+$ clones from *Ad-Wap-Cre* injection were found in alveoli (Figure 4D, overall quantifications; Figures S4C and S4D, representative IF pictures). Furthermore, we found that the majority of YFP$^+$ clones in *Ad-Wap-Cre*-marked MGs were large alveolar clones (i.e., more than five cells) (Figure 4E). We showed previously that, when injected using a higher viral titer, *Ad-K8-Cre*-marked luminal MECs at the virgin stage also could give rise to similar large alveolar clones at mid-gestation (Tao et al., 2014). The enrichment of alveolar clones in *Ad-Wap-Cre*-marked MGs at mid-gestation could be explained by two possibilities. One possibility is that a significant portion of the freshly labeled, ductal-localized WVs were indeed APs that gave rise to alveolar clones at mid-gestation. The second possibility is that YFP$^+$ alveolar clones at mid-gestation were mainly derived from freshly labeled, alveolar-localized WVs, whereas most of the freshly labeled ductal WVs were not maintained and disappeared at mid-gestation. Since the majority of freshly labeled WVs were ductal cells by location (Figure 3C), if they were not maintained, we would expect to observe a significant reduction in the total number of YFP$^+$ clones at mid-gestation compared to that from *Ad-K8-Cre*-marked MGs (as *Ad-K8-Cre* not only marks ductal WVs, but also other ductal MECs not targeted by *Ad-Wap-Cre*). However, we noted that, at mid-gestation, the total numbers of YFP$^+$ clones in *Ad-K8-Cre*-marked and *Ad-Wap-Cre*-marked MGs were not significantly different (*Ad-Wap-Cre*, 83 ± 7; *Ad-K8-Cre*, 99 ± 20; p value = 0.43; Figure 4E), which thus largely ruled out the second possibility. Overall, our lineage-tracing data demonstrate that *Ad-Wap-Cre*-marked nulliparous MECs (i.e., WVs) are enriched with clonogenic APs.

To further characterize the *Ad-Wap-Cre*-marked lineage, we performed FACS analysis. We found that, 3 days post-injection, almost all YFP$^+$ MECs were localized in the luminal gate and most of them were positive for the LP markers CD61 and C-KIT (Figure 4F, ii and iv). Upon chasing to mid-gestation and lactation, we found that most YFP$^+$ MECs became CD61$^-$ differentiating (and proliferating) ALs, although important, a small subset of them remained as CD61$^+$ cells (Figure 4F, viii and xii). Of note, a small population of CD61$^+$ YFP$^+$ MECs also was observed in *Wap-Cre;R26YFP* females at mid-gestation (Figures S2A and S2B). Following involution, the YFP$^+$ cells reverted to the nulliparous state and appeared as CD61$^+$ LPs again (Figure 4F, xiv and xvi). These data (Figures 4F and S4E) suggest that the *Ad-Wap-Cre*-marked WVs undergo full alveolar differentiation during pregnancy/lactation and may maintain themselves as long-lived APs after lactation/involution.

To further demonstrate the long-lived nature and differentiation property of WVs, we injected *Ad-Wap-Cre* to virgin females carrying the above-described conditional Cre-reporter *R26lz*. During lactation of the first pregnancy or after multiple rounds of pregnancy, we could observe large lobules composed of *lacZ*$^+$ cells (Figure S4F). Similarly, we also could detect YFP$^+$ luminal MECs in *R26YFP* females after multiple rounds of pregnancy, upon injection of *Ad-Wap-Cre* at the virgin stage (Figure S4G). Collectively, our lineage-tracing data suggest that WVs in the virgin MG are enriched with a pre-existing population of long-lived APs, in the physiological setting.

WVs Can Serve as Cells of Origin of Luminal Mammary Tumors

It was suggested previously that PI-MECs were the primary targets of neoplastic transformation in multiple *MMTV*-promoter-driven breast cancer models, including *MMTV-Neu* and *MMTV-PyMT* models (Henry et al., 2004; Jeselsohn et al., 2010), both of which develop luminal mammary tumors (Pfefferle et al., 2013). However, since PI-MECs were initially proposed as MECs induced by parity, it was uncertain whether WVs that are present in nulliparous MGs could serve as direct cells of origin in them. To address this, we bred these two models to *Wap-Cre* and *R26YFP* mice and generated *MMTV-*(*Neu* or *PyMT*)*;Wap-Cre;R26YFP* triple transgenic females. We found that most mammary tumors that developed in these triple-transgenic virgin females were indeed YFP$^+$ (Figures 5A and S5A). In both models, at the premalignant stage, we could already observe significant expansion of YFP-marked MECs, and all these YFP$^+$ premalignant

(D and E) Relative percentages of YFP$^+$ ductal and alveolar clones (total = 100%) in mid-gestation MGs (induced by *Ad-Wap-Cre* or *Ad-K8-Cre* [control] injection at the virgin stage) (D) and distributions of small (1–2 cells), intermediate (3–5 cells), and large (>5 cells) YFP$^+$ clones in the indicated mid-gestation MG sections (E). Relative percentages of YFP$^+$ ductal and alveolar clones from each injection (total = 100%) are shown. K8, *Ad-K8-Cre*; WAP, *Ad-Wap-Cre*. Data represent mean ± SEM; n = 4 (K8) and 5 (WAP) injected MGs from five mice (K8 and WAP injected to the contralateral sides of inguinal MGs in each mouse, respectively).

(F) FACS analysis of *R26YFP* virgin females injected with *Ad-Wap-Cre* chased to different developmental stages is shown. Lu, luminal gate; Ba, basal gate; Str, stromal cell gate. Detailed quantification of the percentage of each gate is shown in Figure S4E.

See also Figure S4.

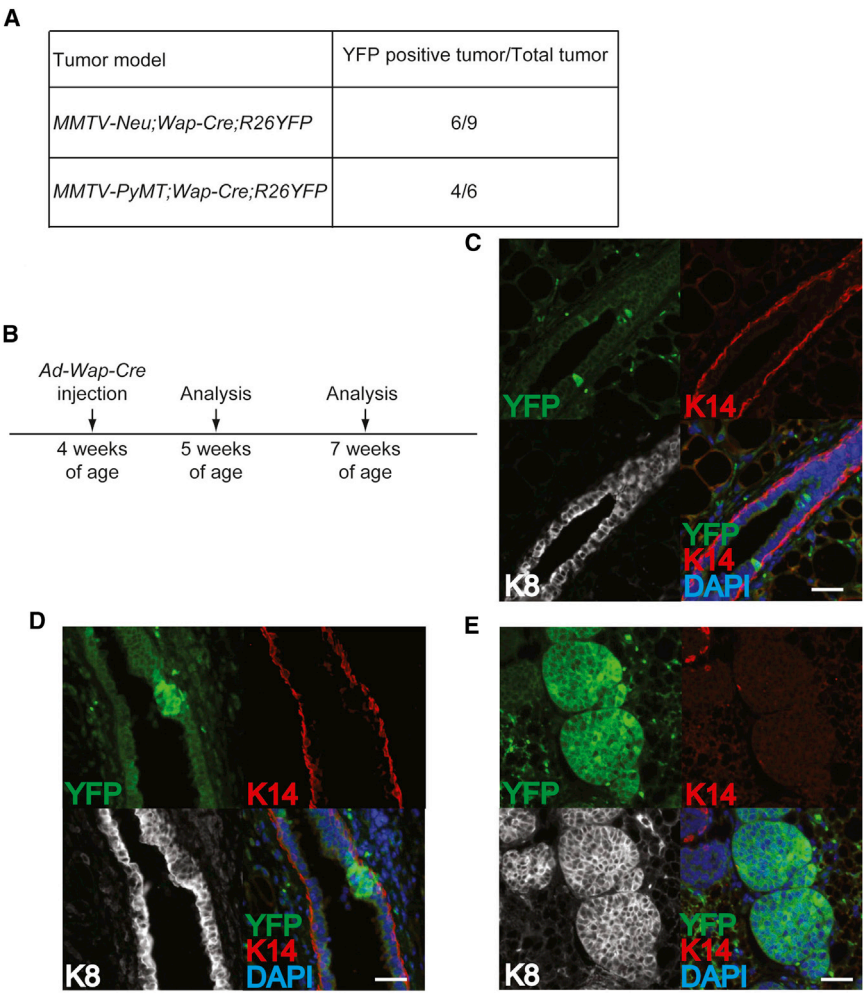


Figure 5. WVs Can Be Cells of Origin of Luminal Mammary Tumors

(A) Most tumors developed in *MMTV-Neu;Wap-Cre;R26YFP* and *MMTV-PyMT;Wap-Cre;R26YFP* virgin females were YFP$^+$ (i.e., >50% YFP$^+$Lin$^-$ cells in a tumor; see Figure S5A).

(B) Timeline to chase the cellular origin of *MMTV-PyMT* mammary tumor based on genetic marking by *Ad-Wap-Cre* intraductal injection is shown.

(C and D) IF staining of induced YFP$^+$ cells 1 week after *Ad-Wap-Cre* injection. While some YFP-marked cells remained comparatively normal (C), others began to form YFP$^+$ K8$^+$K14$^-$ in situ hyperplasia (D). Scale bars, 50 μm.

(E) IF staining of YFP$^+$ K8$^+$K14$^-$ adenocarcinomas induced by *Ad-Wap-Cre* injection. Note that adenocarcinomas were either completely YFP$^+$ or YFP$^-$. Scale bar, 50 μm. See also Figure S5.

MECs remained restricted to the luminal lineage (Figure S5B). These data supported that WVs could serve as cells of origin in these models when kept as virgin females.

To more definitively determine WVs as cells of origin, we focused on the *MMTV-PyMT* model. We injected *Ad-Wap-Cre* to *MMTV-PyMT;R26YFP* virgin females (to mark WVs by YFP, Figure 5B) at 4 weeks of age when their MGs still appeared normal (Figure S5C). At 1 week post-injection, in addition to individual YFP-marked luminal MECs (Figure 5C), we also observed clones of YFP$^+$ premalignant K8$^+$K14$^-$ luminal cells starting to emerge in mammary ducts (Figure 5D). At 3 weeks post-injection, we detected K8$^+$K14$^-$ luminal mammary tumors that were either entirely positive or negative for YFP (Figure 5E), suggesting that each tumor was clonally derived from its cell of origin, either marked by *Ad-Wap-Cre* or not. These data thus provide direct evidence to support that (*Ad-*)*Wap-Cre*-marked WVs are cells of origin of luminal mammary tumors in the *MMTV-PyMT* model.

WVs Can Serve as Cells of Origin of Mammary Tumors with Basal Differentiation

We showed previously that *Wap-Cre*-mediated activation of the conditional *Etv6-NTRK3* (*EN*) knockin allele in virgin females led to the development of mammary tumors with basal differentiation. These tumors were either composed of K8$^+$K14$^{-/low}$ luminal tumor cells surrounded by (and/or mixed with) K5$^+$K14$^+$ basal tumor cells (type I tumors) or composed of K8$^+$K14$^+$K5$^-$ tumor cells (type II tumors) (Li et al., 2007). Since we showed that WVs are K8$^+$K14$^-$ CD61$^+$ luminal cells (Figures 1 and 2), we asked whether activation of *EN* led to abnormal luminal-to-basal differentiation directly from WV luminal cells. To determine this, we used two genetic-marking approaches. In the first approach, we generated *Wap-Cre;EN;R26YFP* virgin females and checked their MGs at the premalignant stage. We observed a significant portion of YFP$^+$ premalignant MECs emerged in the basal gate, compared to premalignant MECs in *MMTV-*(*Neu* or *PyMT*)*;Wap-Cre;R26YFP* virgin females, which were almost entirely restricted to the

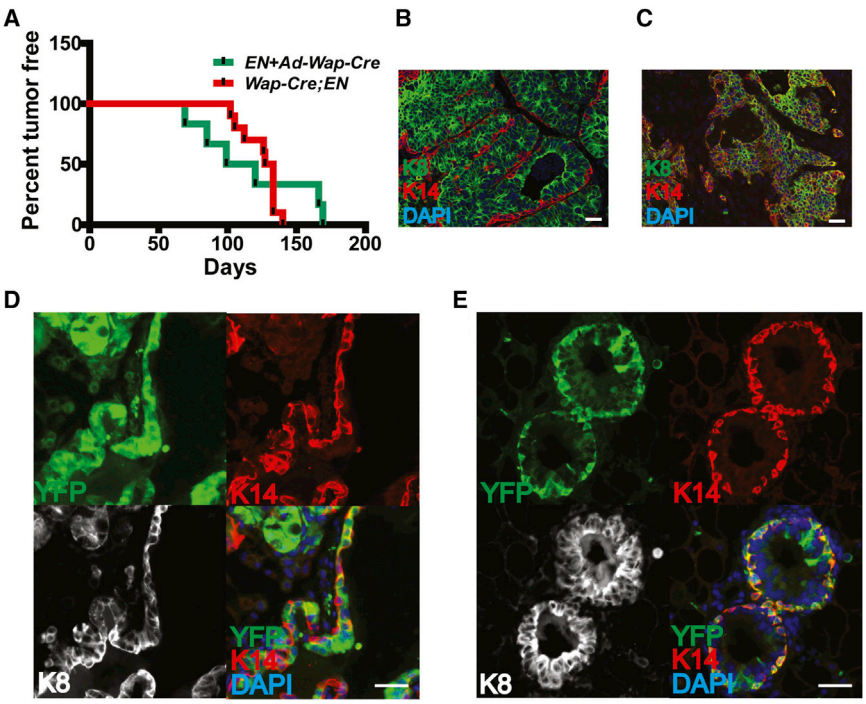


Figure 6. WVs Can Be Cells of Origin of Mammary Tumors with Basal Differentiation

(A) *EN* mice induced by *Ad-Wap-Cre* (n = 6) showed similar tumor latency as *Wap-Cre;EN* mice (n = 10); both cohorts had the same FVB background.

(B and C) IF staining confirmed that the two tumor types (type I, B; type II, C) originally identified in *Wap-Cre;EN* mice were recapitulated in *EN* mice induced by *Ad-Wap-Cre*. Scale bars, 50 μm.

(D and E) IF staining of *EN;R26YFP* MGs after *Ad-Wap-Cre* intraductal injection showed abnormal basal differentiation during the preneoplastic stage. (D) Most YFP^+ preneoplastic clones contained $K8^+K14^+$ cells. (E) Some YFP^+ clones contained largely $K8^-K14^+$ basal cells mixed with a few $K8^+K14^-$ luminal cells. Scale bars, 50 μm.

See also Figure S6.

luminal gate (Figure S5B). To further confirm this, we sorted YFP^+ MECs from *Wap-Cre;EN;R26YFP* premalignant MGs, and, by IF staining, we indeed detected MECs with basal differentiation (i.e., $K14^+K8^-$ and $K14^+K8^+$ cells), in addition to $K14^-K8^+$ luminal cells (Figure S5D).

To definitively prove WV luminal cells as cells of origin of *EN* mammary tumors with basal differentiation, we injected *Ad-Wap-Cre* to mammary ducts of virgin females carrying the same *EN* conditional allele and found that all injected females developed mammary tumors with a latency similar to that of *Wap-Cre;EN* females (both under FVB background) (Figure 6A). By IF staining, we found that the two types of tumors (i.e., types I and II) developed in our original *Wap-Cre;EN* mouse model (Li et al., 2007) were both recapitulated (Figures 6B, 6C, and S6). These new data thus provide direct evidence to support WVs as the cellular origin in this tumor model. Next we analyzed the *Ad-Wap-Cre*-injected *EN;R26YFP* MGs with premalignant lesions. By IF staining, we detected YFP^+ lesions composed of $K14^+K8^+$ MECs, as well as YFP^+ clones with $K14^+K8^-$ basal cells mixed with $K8^+K14^-$ luminal cells (Figures 6D and 6E). These data suggest that, upon turning on *EN* in WVs, the EN oncoprotein could drive abnormal basal differentiation (e.g., upregulation of K14) directly from WV luminal cells, eventually leading to the development of mammary tumors with basal differentiation.

Overall, these data demonstrate that, under the influence of different oncogenic insults, WV luminal cells can undergo abnormal differentiation, including basal differentiation. It was shown recently that mammary tumors that developed in *Wap-Cre;EN*, *MMTV-PyMT*, and *MMTV-Neu* mice resemble human HER2-enriched, luminal B, and luminal A breast cancers at the molecular level, respectively, and all three models highly express a gene signature composed of many genes involved in alveolar function (Pfefferle et al., 2013). Thus, it is likely that these three models may share the same cellular origin, and our data provide support for a model in which a common luminal subpopulation enriched with long-lived APs may give rise to distinct breast cancer subtypes, when transformed by different oncogenic events.

DISCUSSION

Whether the luminal MEC lineage in adults is sustained by lineage-restricted LuSCs (Van Keymeulen et al., 2011) or by bipotent basal stem cells (Rios et al., 2014; Wang et al., 2015) in the physiological setting has become a topic of intense debate. Nevertheless, one common observation that started to emerge was evidence for long-lived LuSCs/LPs that could sustain the luminal lineage in vivo, yet their identity and role in breast cancer remained largely elusive. Our study uncovered that (*Ad-*)*Wap-Cre*-marked MECs in virgin females (i.e., WVs) are luminal MECs enriched with a pre-existing population of long-lived LPs committed to the alveolar luminal sub-lineage. In contrast to the previous finding in parous females for PI-MECs, we found

that, during pregnancy and lactation, these WVs only produced ALs, but not myoepithelial cells, in their native habitat. This observation is consistent with a recent clonal analysis study based on *K5-rtTA/TetO-Cre* and a multicolor Cre-reporter, which demonstrated different cellular origins of luminal and myoepithelial cells within individual alveoli (Rios et al., 2014).

WVs appear to be the target of recently revealed paracrine basal-to-luminal cell signaling controlled by p63 (Forster et al., 2014), and they overlap with LPs defined by either *Elf5*-based (Rios et al., 2014) or *Notch1*-based lineage tracing (Rodilla et al., 2015). WVs may represent the classic APs that work together with the two novel MEC subpopulations revealed by *Notch2*-based genetic marking (i.e., S [small] and L [large] cells) (Šale et al., 2013), for the formation of alveolar clusters in response to alveologenic signals. Furthermore, due to the self-sustained, long-lived nature of WVs, our study provides an alternative model to explain the developmental origin of PI-MECs. Rather than derived from de-differentiation of secretory alveolar cells that survived involution, PI-MECs detected via *Wap-Cre* might represent a combination of *equipotent* APs, either marked by *Wap-Cre* at the virgin stage (i.e., WVs) or freshly marked by *Wap-Cre* during late gestation (i.e., the second wave of genetic marking when the *Wap* promoter activity is dramatically upregulated; Figure 7A). These APs may be intrinsically resistant to involution. As there should be a lot more of them freshly labeled by *Wap-Cre* during late gestation (due to much higher *Wap* activity at this stage) than those pre-labeled at the virgin stage, it would appear as if a larger population of *Wap-Cre*-marked long-lived MECs emerged only after parity (Wagner et al., 2002). This model is supported by a recent study characterizing parous PI-MECs using *Wap-Cre* mice (Chang et al., 2014). Lastly, the following also should be pointed out: (1) our model does not exclude the possibility that PI-MECs also can be derived from de-differentiation of mature secretory ALs that survive involution; (2) our model does not suggest that nulliparous WVs and parous PI-MECs are identical, and, in fact, parity is expected to have a profound effect on this progenitor cell lineage (in support of this, it was shown recently that, depending on the timing of acquiring premalignant lesions and of pregnancy, the *Wap*-marked alveolar cell population was the target cell mediating the dual roles of pregnancy in increasing or reducing breast cancer risks [Haricharan et al., 2013, 2014]); and (3) our model does not rule out a possibility that WVs may include LuSCs that can give rise to both ductal luminal cells and APs.

Although MaSCs (i.e., bipotent basal stem cells) were initially hypothesized as cells of origin of breast cancer, recent studies suggested that most breast cancers, including basal-like breast cancer (BLBC), might originate from luminal cells (Skibinski and Kuperwasser, 2015). In particular, human LPs have been suggested as the cellular origin of BLBC with *BRCA1* mutation (Lim et al., 2009; Proia et al., 2011), an observation supported by mouse modeling (Molyneux et al., 2010). Our in vivo data provided evidence to support that WVs, a subpopulation of $CD61^+$ LPs, also could serve as cells of origin of breast cancer. Several properties of these cells, such as resistance to involution and self-renewal capacity, may contribute to their increased risk to serve as cells of origin of breast cancer.

Recent lineage-tracing studies revealed that bipotent basal stem cells could contribute to the luminal lineage in vivo (Rios et al., 2014; Wang et al., 2014). Due to their long-lived nature, it is possible that these bipotent stem cells are the cell type in which an initiating mutation is acquired (i.e., cells of mutation). However, it is also possible that breast cancer may not initiate directly from such mutation-bearing bipotent basal stem cells; instead, breast cancer may initiate from more committed progenitors (i.e., cells of origin), when they inherit the initiating mutation from their parental bipotent stem cells. This concept was described previously in the setting of glioma (Liu et al., 2011). Furthermore, since our study suggests that WVs are long-lived cells, one could argue that there would be a long time window for them to acquire and accumulate mutations as well, making it possible for them to serve as both cells of mutation and cells of origin of breast cancer. Of note, our study does not suggest that WVs are the only cells of origin of breast cancer; ductal progenitors also may serve as cells of origin of breast cancer (Jeselsohn et al., 2010).

Finally, although the normal fate of WVs is to give rise to ALs under the influence of ovarian and pregnancy hormones, upon transformation by different oncogenes, they can undergo aberrant luminal-to-luminal or luminal-to-basal differentiation, eventually leading to the development of heterogeneous mammary tumors (Figure 7B). Overall, our study provides direct support for a luminal subpopulation enriched with APs as the cellular origin of multiple subtypes of breast cancers, and it may have important implications for understanding the cell of origin and heterogeneity of breast cancer. As reproductive history of an individual has a profound impact on her lifetime breast cancer risk, identification of a long-lived MEC subpopulation enriched with APs, which are excellent candidate target cells for mediating pregnancy-induced risk changes, may lead to novel strategies to prevent and treat breast cancer.

EXPERIMENTAL PROCEDURES

Mouse Strains, Ovariectomy, and Mammary Gland Intraductal Injection

Wap-Cre mouse was obtained from the Mouse Models of Human Cancers Consortium (MMHCC) repository. *R26YFP*, *R26lz*, *MMTV-Neu*, and *MMTV-PyMT* mice were purchased from the

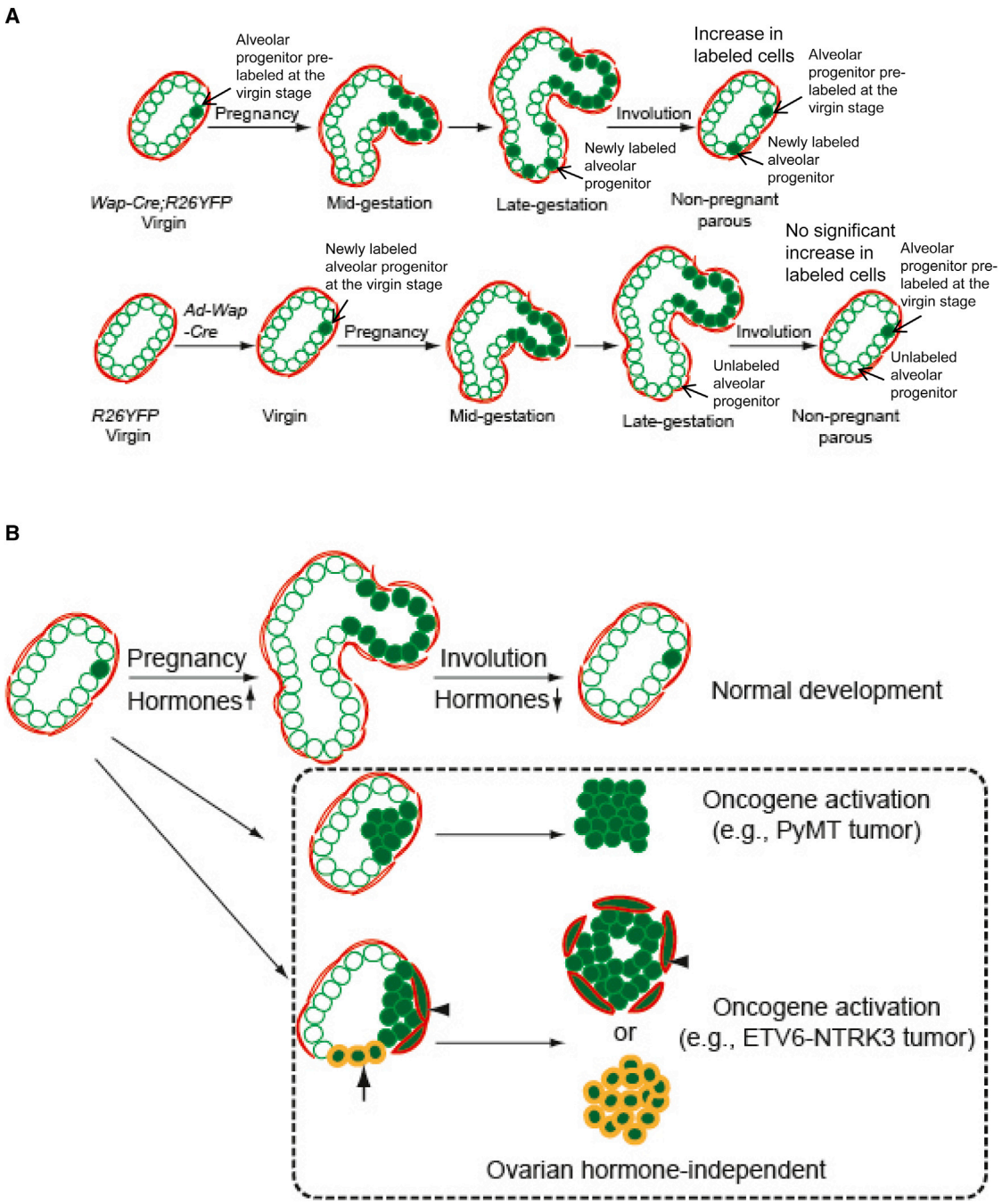


Figure 7. Proposed Models for Normal and Abnormal Fates of WVs

(A) Schematic diagrams show AP cell fate mapping using constitutive *Wap-Cre;R26YFP* mice (top) and the *Ad-Wap-Cre*-based lineage-tracing approach (bottom).

(B) Schematic diagram shows abnormal (during cancer development) fates of WVs. Arrow and arrowhead indicate abnormal differentiation of WVs to $K14^{+}K8^{+}$ and $K14^{+}K8^{-}$ basal-like cells, respectively.

Jackson Laboratory. *EN* conditional knockin mouse was generated previously (Li et al., 2007). All mice were under a pure FVB background unless otherwise indicated. Ovariectomy was performed on 10-week-old *Wap-Cre;R26YFP* virgin mice and analyzed 8 weeks after surgery. Intraductal injection of *Ad-Cre* was performed as described previously (Tao et al., 2014). All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC).

Mammary Gland Cell Preparation, FACS Analysis, and Cell Sorting

Thoracic and inguinal MGs were dissected and cell suspensions were prepared as previously described (Shackleton et al., 2006). Flow cytometric analysis was performed with an Accuri C6 analyzer (BD Biosciences) and analyzed with CFlow (BD Biosciences), or it was performed with a DXP11 flow cytometer and analyzed with FlowJo. FACS sorting was performed with a FACSAria sorter (BD Biosciences).

Whole-Mount Staining, IHC and IF Staining, and Clonal Analysis

Whole mounts of MGs of sham and ovariectomized mice were fixed and processed as previously described (Jones et al., 1996). IHC and IF staining of MG tissue was performed on inguinal MGs that were fixed in 10% formalin (Fisher Scientific) and embedded in paraffin. FACS-sorted cells used for IF were fixed in 4% paraformaldehyde and permeabilized by 0.5% Triton X-100. For determination of the clone size of YFP-marked cells in the MG of *Wap-Cre*;*R26YFP* adult virgin females, at least 100 YFP^+ clones (single cell, two cells, or more than two cells) were counted in four different mice. In vivo clonal analysis was performed as described previously (Tao et al., 2014).

Microarray Analysis

Total RNA from sorted MECs was prepared by the RNeasy kit (QIAGEN) and amplified with the Ovation RNA Amplification System V2 (NuGEN). Mouse Genome 430 2.0 Array (Affymetrix) was used to generate the expression profiles. All arrays were normalized by dCHIP, and PCA was performed by using the PCA module in GenePattern (http://www.broadinstitute.org/cancer/software/genepattern/).

qRT-PCR

For qRT-PCR, total RNA from sorted MEC subsets was prepared by the RNeasy kit and cDNA was generated with iScript (Bio-Rad) according to the manufacture's protocol. PCR was performed using FastStart SYBR Green Master (Roche).

Ad-Wap-Cre Adenovirus Vector Construction and Adenovirus Production

Mouse *Wap* promoter (Wagner et al., 1997) was amplified from mouse genomic DNA. The resultant PCR fragment was inserted into adenovirus shuttle vector pAd5 (Gene Transfer Core of the University of Iowa) to drive expression of Cre recombinase with a nuclear localization sequence (nlsCre, Addgene). *Ad-Wap-Cre* adenovirus was produced as described previously (Tao et al., 2014).

Statistics

Student's t test was used to calculate the p values. Data were reported as mean ± SEM.

ACCESSION NUMBERS

The accession number for the microarray data reported in this paper is GEO: GSE47376.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.05.014.

AUTHOR CONTRIBUTIONS

L.T. performed all *Ad-Wap-Cre*-based lineage-tracing experiments. M.P.A.v.B. performed all *Wap-Cre* mice-based experiments. All authors contributed to the design, data analysis, and writing of this paper.

ACKNOWLEDGMENTS

We are grateful to G. Losyev and Y. Qiu for expert technical assistance with FACS sorting and Y. Shao and E. Fox for microarrays. This research was supported by a Pathway to Independence K99/R00 grant from the National Cancer Institute (CA126980), a seed grant and a Cancer Program Pilot grant from Harvard Stem Cell Institute, a Milton Fund award from Harvard University, the Hearst Foundation Young Investigator award from BWH, and the Start-up Fund from BWH to Z.L.

Received: March 25, 2014
Revised: May 26, 2015
Accepted: May 26, 2015
Published: June 25, 2015

REFERENCES

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.

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.

Booth, B.W., Boulanger, C.A., and Smith, G.H. (2007). Alveolar progenitor cells develop in mouse mammary glands independent of pregnancy and lactation. J. Cell. Physiol. *212*, 729–736.

Boulanger, C.A., Wagner, K.U., and Smith, G.H. (2005). Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression. Oncogene *24*, 552–560.

Chang, T.H., Kunasegaran, K., Tarulli, G.A., De Silva, D., Voorhoeve, P.M., and Pietersen, A.M. (2014). New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. Breast Cancer Res. *16*, R1.

Choi, N., Zhang, B., Zhang, L., Ittmann, M., and Xin, L. (2012). Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. Cancer Cell *21*, 253–265.

Forster, N., Saladi, S.V., van Bragt, M., Sfondouris, M.E., Jones, F.E., Li, Z., and Ellisen, L.W. (2014). Basal cell signaling by p63 controls

luminal progenitor function and lactation via NRG1. Dev. Cell *28*, 147–160.

Haricharan, S., Dong, J., Hein, S., Reddy, J.P., Du, Z., Toneff, M., Holloway, K., Hilsenbeck, S.G., Huang, S., Atkinson, R., et al. (2013). Mechanism and preclinical prevention of increased breast cancer risk caused by pregnancy. eLife *2*, e00996.

Haricharan, S., Hein, S.M., Dong, J., Toneff, M.J., Aina, O.H., Rao, P.H., Cardiff, R.D., and Li, Y. (2014). Contribution of an alveolar cell of origin to the high-grade malignant phenotype of pregnancy-associated breast cancer. Oncogene *33*, 5729–5739.

Henry, M.D., Triplett, A.A., Oh, K.B., Smith, G.H., and Wagner, K.U. (2004). Parity-induced mammary epithelial cells facilitate tumorigenesis in MMTV-neu transgenic mice. Oncogene *23*, 6980–6985.

Jeselsohn, R., Brown, N.E., Arendt, L., Klebba, I., Hu, M.G., Kuperwasser, C., and Hinds, P.W. (2010). Cyclin D1 kinase activity is required for the self-renewal of mammary stem and progenitor cells that are targets of MMTV-ErbB2 tumorigenesis. Cancer Cell *17*, 65–76.

Jones, F.E., Jerry, D.J., Guarino, B.C., Andrews, G.C., and Stern, D.F. (1996). Heregulin induces in vivo proliferation and differentiation of mammary epithelium into secretory lobuloalveoli. Cell Growth Differ. *7*, 1031–1038.

Kordon, E.C., and Smith, G.H. (1998). An entire functional mammary gland may comprise the progeny from a single cell. Development *125*, 1921–1930.

Li, Z., Tognon, C.E., Godinho, F.J., Yasaitis, L., Hock, H., Herschkowitz, J.I., Lannon, C.L., Cho, E., Kim, S.J., Bronson, R.T., et al. (2007). ETV6-NTRK3 fusion oncogene initiates breast cancer from committed mammary progenitors via activation of AP1 complex. Cancer Cell *12*, 542–558.

Li, W., Ferguson, B.J., Khaled, W.T., Tevendale, M., Stingl, J., Poli, V., Rich, T., Salomoni, P., and Watson, C.J. (2009). PML depletion disrupts normal mammary gland development and skews the composition of the mammary luminal cell progenitor pool. Proc. Natl. Acad. Sci. USA *106*, 4725–4730.

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.; kConFab (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913.

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.

Liu, C., Sage, J.C., Miller, M.R., Verhaak, R.G., Hippenmeyer, S., Vogel, H., Foreman, O., Bronson, R.T., Nishiyama, A., Luo, L., and Zong, H. (2011). Mosaic analysis with double markers reveals tumor cell of origin in glioma. Cell *146*, 209–221.

Matulka, L.A., Triplett, A.A., and Wagner, K.U. (2007). Parity-induced mammary epithelial cells are multipotent and express cell surface markers associated with stem cells. Dev. Biol. *303*, 29–44.

Molyneux, G., Geyer, F.C., Magnay, F.A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A., et al. (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell *7*, 403–417.

Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. Nature *406*, 747–752.

Pfefferle, A.D., Herschkowitz, J.I., Usary, J., Harrell, J.C., Spike, B.T., Adams, J.R., Torres-Arzayus, M.I., Brown, M., Egan, S.E., Wahl, G.M., et al. (2013). Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. Genome Biol. *14*, R125.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. *12*, R68.

Prater, M.D., Petit, V., Alasdair Russell, I., Giraddi, R.R., Shehata, M., Menon, S., Schulte, R., Kalajzic, I., Rath, N., Olson, M.F., et al. (2014). Mammary stem cells have myoepithelial cell properties. Nat. Cell Biol. *16*, 942–950.

Proia, T.A., Keller, P.J., Gupta, P.B., Klebba, I., Jones, A.D., Sedic, M., Gilmore, H., Tung, N., Naber, S.P., Schnitt, S., et al. (2011). Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. Cell Stem Cell *8*, 149–163.

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.

Robinson, G.W., McKnight, R.A., Smith, G.H., and Hennighausen, L. (1995). Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. Development *121*, 2079–2090.

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.

Šale, S., Lafkas, D., and Artavanis-Tsakonas, S. (2013). Notch2 genetic fate mapping reveals two previously unrecognized mammary epithelial lineages. Nat. Cell Biol. *15*, 451–460.

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.

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.

Skibinski, A., and Kuperwasser, C. (2015). The origin of breast tumor heterogeneity. Oncogene.

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.

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.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl. Acad. Sci. USA *100*, 8418–8423.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993–997.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Tao, L., van Bragt, M.P.A., Laudadio, E., and Li, Z. (2014). Lineage tracing of mammary epithelial cells using cell-type-specific cre-expressing adenoviruses. Stem Cell Reports *2*, 770–779.

van Amerongen, R., Bowman, A.N., and Nusse, R. (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.

Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189–193.

Visvader, J.E. (2011). Cells of origin in cancer. Nature *469*, 314–322.

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.

Wagner, K.U., Boulanger, C.A., Henry, M.D., Sgagias, M., Hennighausen, L., and Smith, G.H. (2002). An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. Development *129*, 1377–1386.

Wang, Z.A., Mitrofanova, A., Bergren, S.K., Abate-Shen, C., Cardiff, R.D., Califano, A., and Shen, M.M. (2013). Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. Nat. Cell Biol. *15*, 274–283.

Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.O., Yang, L., and Zeng, Y.A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. Nature *517*, 81–84.