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

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.
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 Li1,2,*

1Division of Genetics, Brigham and Women’s Hospital (BWH), Boston, MA 02115, USA
2Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
3Co-first author
*Correspondence: zli4@rics.bwh.harvard.edu
http://dx.doi.org/10.1016/j.stemcr.2015.05.014
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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−CD24hiCD29med 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 (Shen 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+ cells (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.
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

<table>
<thead>
<tr>
<th>RANK</th>
<th>NAME</th>
<th>NES</th>
<th>NOM p-val</th>
<th>FDR q-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RESPONSE_TO_WOUNDING</td>
<td>2.12</td>
<td>0</td>
<td>0.0086</td>
</tr>
<tr>
<td>2</td>
<td>RECEPTOR_ACTIVITY</td>
<td>2.12</td>
<td>0</td>
<td>0.0049</td>
</tr>
<tr>
<td>3</td>
<td>IMMUNE_RESPONSE</td>
<td>2.10</td>
<td>0</td>
<td>0.0057</td>
</tr>
<tr>
<td>4</td>
<td>RESPONSE_TO_EXTERNAL_STIMULUS</td>
<td>2.10</td>
<td>0</td>
<td>0.0043</td>
</tr>
<tr>
<td>5</td>
<td>INFLAMMATORY_RESPONSE</td>
<td>2.05</td>
<td>0</td>
<td>0.0022</td>
</tr>
<tr>
<td>6</td>
<td>DEFENSE_RESPONSE</td>
<td>2.03</td>
<td>0</td>
<td>0.0028</td>
</tr>
<tr>
<td>7</td>
<td>AMINO_ACID_METABOLIC_PROCESS</td>
<td>2.00</td>
<td>0</td>
<td>0.0154</td>
</tr>
<tr>
<td>8</td>
<td>NITROGEN_COMPOUND_CATABOLIC_PROCESS</td>
<td>1.99</td>
<td>0</td>
<td>0.0147</td>
</tr>
<tr>
<td>9</td>
<td>OXYGEN_BINDING</td>
<td>1.97</td>
<td>0</td>
<td>0.0096</td>
</tr>
<tr>
<td>10</td>
<td>INTEGRAL_TO_PLASMA_MEMBRANE</td>
<td>1.94</td>
<td>0</td>
<td>0.0020</td>
</tr>
<tr>
<td>11</td>
<td>TRANSCRIPTION</td>
<td>1.93</td>
<td>0</td>
<td>0.0039</td>
</tr>
<tr>
<td>12</td>
<td>PLASMA_MEMBRANE</td>
<td>1.92</td>
<td>0</td>
<td>0.0081</td>
</tr>
<tr>
<td>13</td>
<td>RECEPTOR_ACTIVITY</td>
<td>1.91</td>
<td>0</td>
<td>0.0029</td>
</tr>
<tr>
<td>14</td>
<td>INTEGRAL_TO_EPIDERM</td>
<td>1.90</td>
<td>0</td>
<td>0.0021</td>
</tr>
<tr>
<td>15</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0022</td>
</tr>
<tr>
<td>16</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0023</td>
</tr>
<tr>
<td>17</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0024</td>
</tr>
<tr>
<td>18</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0025</td>
</tr>
<tr>
<td>19</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0026</td>
</tr>
<tr>
<td>20</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0027</td>
</tr>
<tr>
<td>21</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0028</td>
</tr>
<tr>
<td>22</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0029</td>
</tr>
<tr>
<td>23</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0030</td>
</tr>
<tr>
<td>24</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0031</td>
</tr>
<tr>
<td>25</td>
<td>INTEGRAL_TO_MEMBRANE</td>
<td>1.90</td>
<td>0</td>
<td>0.0032</td>
</tr>
</tbody>
</table>

C

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 ≤ 0.25 was considered as significant.

See also Figure S2.
differentiation (Figures 3C–3H), we reasoned that their full potential would only be revealed upon pregnancy and lactation. We therefore injected Ad-Wap-Cre intraductal injection into R26YFP nulliparous MGs. Animals were kept as virgins.

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 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 4E, 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
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+ K8+K14+/C0 premalignant luminal cells starting to emerge in mammary ducts (Figure 5D). At 3 weeks post-injection, we detected K8+K14+/C0 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 tumor cells surrounded by (and/or mixed with) K5+K14+ basal tumor cells (Figure 5E), compared to premalignant tumor cells in MMTV-(Neu or PyMT);Wap-Cre;R26YFP virgin females, which were almost entirely restricted to the luminal lineage.
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+/C0 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 WV, 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., WV) 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 ALs that work together with the two novel MEC subpopulations revealed by Notch2-based genetic marking (i.e., S [small] and L [large] cells) (Sale et al., 2013), for the formation of alveolar clusters in response to alveogenic 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 ALs, 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 7 A). These ALs 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 can 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 ALs.

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 7 B). Overall, our study provides direct support for a luminal subpopulation enriched with ALPs 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 ALPs, 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
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).

**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.
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 FACS Aria sorter (BD Biosciences).

Whole-Mount Staining, IHC and IF Staining, and Clonal Analysis
Whole mounts of MGs of sham and ovariectomized mouse 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 with 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 manufacturer’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 Foundation 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 24, 2015

REFERENCES


