Supplementary figure legends

Supplementary Figure 1 - FACS strategies used for detection and sorting of $EGFP^+$ mammary epithelial cell populations in single cell suspensions from *Lgr6-CreERT2*$^{+/-}$ mammary glands at 5 weeks of age.

(a) FACS strategy applied to sort $EGFP^+$ and $EGFP^-$ mammary epithelial cells (quadrant I) from 5-6 week old mice for RNA sequencing analysis. **(b)** Heat map of RNA-Seq transcriptome analysis from two experiments (Ex.1/ Ex.2), showing all genes differentially expressed between $Lgr6^+$ and $Lgr6^-$ cells isolated from mammary glands of 5-6w-old Lgr6-$CreERT2^{+/-}$ mice. Red indicates upregulated genes; blue indicates downregulated genes. (n = 6 mice for each of the 2 biological replicates). **(c)** FACS strategy for determining the distribution of $EGFP^+$ cells over basal ($lin^- CD29^{hi}CD24^+$) and luminal ($lin^- CD29^{lo}CD24^+$) MECs. The gate for $EGFP^+$ cells was set using wildtype mice as fluorescence-minus-one (FMO) EGFP controls.

Supplementary Figure 2 - *Lgr6*-expressing cells in postnatal mammary gland development. **(a)** K8 immunostaining of *Lgr6-CreERT2*$^{+/-}$ mammary glands at postnatal day 14 (P14) and P30**.** EGFP^{+} cells in luminal (arrowheads) and basal cells (arrows). Scale bars: 25 µm. **(b)** Recombination efficiency 24 h after tamoxifen administration to 2w-old (n=5 mice) and 4w-old (n=4 mice) *Lgr6-CreERT2*$^{+/-}$*:tdTomato*$^{+/-}$ (LT) females. Mean ± s.e.m. **(c)** FACS plot showing the distribution of tdTomato^{+} cells over MECs 24 h after tamoxifen administration to 2w-old (left) and 4w-old (right) LT mice. **(d)** FACS analysis to exclude recombination of the *Rosa26-tdTomato* allele by *Lgr6-CreERT2* in the absence of Tamoxifen. Left/middle panel: MECs from 10w old wildtype (*WT*) or *LT* mice injected with sunflower seed oil at P12 and analysed 8w later. Right: Detection of tdTomato^{+} cells 8w after administration of 1mg tamoxifen at P12. **(e)** Mammary gland from *LT* mouse injected with sunflower seed oil at P12 and imaged 8w later. Scale bars: 1mm **(f)** F-actin staining of *LT* mammary glands 22w p.i. in pre-puberty (right panel) and 24w p.i. in puberty (left panel). Scale bars: 100 µm, 20µm (insets). **(g, h)** Distribution of basal and luminal multi-cellular tdTomato^{+} clones, and cell patches spanning both compartments after induction in pre-puberty (2w, **g**) and in puberty (4w, **h**). Numbers indicate % of multicellular areas comprising adjacent basal and luminal cells. (Analysed clones pooled from n = 3 mice per timepoint: n=1157, n=344, n=542, n=645, n=691 for 2w, and n=531, n=1476, n=1733, n=163, n=782, n=597 for 4w). Mean ± s.e.m. **(i)** K14/K8 immunostained mammary gland 16w p.i. (4w) showing a rare tdTomato^{+} “clone” spanning the basal (arrow) and luminal (arrowhead) compartments. Scale bar: 25 µm. **(j)** RNA *in situ* hybridization (ISH) analysis. Left: *Lgr6* (blue arrowheads) and *Axin2* (red

arrowheads). *Lgr6/Axin2* co-expression (arrows) can be observed in some cells. Right: Rare cells in the terminal end buds (TEBs) of pubertal mammary glands co-express *Lgr6* and *Lgr5* mRNA (arrows; inset). Scale bars: 10 µm (left), 20 µm (right). **(k)** Overview of transplantation assays of basal $EGFP^+$ and $EGFP^-$ cells from adult virgin LT females into emptied fat pads of NOD/SCID mice. See Supplementary Table 2 for source data for g, h.

Supplementary Fig. 3 – Basal and luminal *Lgr6*$^+$ progenitors contribute to the alveolar network during pregnancy

(a, b) Confocal images of K14/K8 immunostained whole mount mammary glands from LT female at 17.5 dpc in first pregnancy (17.5 dpc P1). tdTomato$^+$ clones dispersed over newly formed alveolar structures contain either basal K14$^+$ cells (a) or K8$^+$ alveolar cells (b). Scale bars: 40 µm (a), 20µm (b). **(c)** Confocal overview image of mammary gland on day 1 of the first lactation (1d L1). LN: lymph node. Scale bar: 2 mm. **(d)** 3D reconstruction of 1d L1 mammary gland showing an alveolar network of myoepithelial tdTomato$^+$ cells. Scale bar: 50 µm. **(e)** Quantification of alveoli containing basal and luminal tdTomato+ cells in the first lactation after tamoxifen administration to *LT* females at 4w of age. (1115 alveoli pooled from 3 mice were analysed). **(f)** Confocal z-stack image of tdTomato$^+$ cells lining ducts and alveolar remnants (arrows), and adjacent to apoptotic cells (arrowheads) at the end of involution (21d Inv1). Autofluorescent signal (arrowheads) comes from dying cells. Scale bar: 50 µm. **(g)** Confocal z-stack image of K14/K8 immunostained mammary duct at the end of involution. Scale bar: 25 µm. **(h)** 3D reconstruction of K8 stained mammary duct showing luminal clones remaining after involution. Scale bar: 50 µm.

Supplementary Figure 4 - Analysis of *Lgr6*-expressing cells in the adult mammary gland. **(a)** Flow cytometry dot plot showing distribution of EGFP expression in MECs of 8w-old *Lgr6-CreERT2*$^{+/-}$ females. **(b)** Bar graph demonstrating the recombination efficiency 24 hours after tamoxifen administration to 8w-old (n = 4 mice) *Lgr6-CreERT2*$^{+/-}$*:Rosa26-tdTomato*$^{+/-}$ (*LT*) females. Mean ± s.e.m. **(c)** Flow cytometry plot showing the distribution of tdTomato-labelled cells over MECs 24 hours p.i. **(d)** Flow cytometry dot plots of primary MECs 22w p.i., showing the distribution of tdTomato-labelled cells. **(e)** Representative flow cytometry plots showing very rare EGFP^{+} cells (left panel), present in both MEC compartments (right panel) of mammary glands from 30w-old *Lgr6-CreERT2*$^{+/-}$ females. **(f)** Confocal z-stack image (image 1) of K14/K8 immunostaining showing adjacent basal and luminal tdTomato^{+} clones (arrowheads) in 1d L1 alveoli of 8w-induced *LT* mice. Images 2-11 show the resolution of the 3D image into 10 single plane images. Insets show magnified regions containing adjacent basal and luminal tdTomato^{+} cells (arrowheads). Scale bar: 20 μm. **(g)** 3D reconstruction of duct at 14.5 dpc showing multi-cellular tdTomato^{+} cell patches containing K14^{+} (arrows) and K8^{+} cells (arrowheads). Scale bar: 50 μm.

Supplementary Figure 5 - Activation of oncogenic Ras combined with Fbxw7 deletion in *Lgr6*-positive cells results in breast cancer. **(a)** Description of the *FRL* mouse strain, derived from crossing three strains (*Fbxw7*$^{f/f}$; *LSL-KRas*^{G12D}*; Lgr6-CreERT2*). **(b)** Treatment scheme of 5-week-old *FRL* mice with tamoxifen and subsequent tumour analysis. **(c)** Immunohistochemistry of tumours from *FRL* mice 3 weeks post-injection with tamoxifen, stained with haematoxylin-eosin (HE) and probed for markers against basal (K5) and luminal (K8) cells. Scale bars: 100μm. **(d)** Immunohistochemistry of tumours from *FRL* mice 8 weeks post-tamoxifen injection stained with HE and probed for K5, K8 and ERα. Scale bars: 100 μm. **(e)** Summary of *BPL* and *FRL* tumour lesions.

Supplementary Figure 6 - *Lgr6*$^{+}$ cells in MMTV-PyMT-expressing mammary glands. **(a)** Mammary glands of *Lgr6-CreERT2*$^{+/-}$ (left) and *Lgr6-CreERT2*$^{+/-}$*:MMTV-PyMT*$^{+/-}$ females (right) at P14. PyMT-induced lesions are found at P14 (inset). Scale bars: 50 µm, 20 µm (inset). **(b)** Mammary glands of *Lgr6-CreERT2*$^{+/-}$ (left) and *Lgr6-CreERT2*$^{+/-}$*:MMTV-PyMT*$^{+/-}$ females (right) at P28. Middle T antigen-induced hyperplasia in puberty (right, inset). Scale bars: 100 µm, 20 µm (inset). **(c)** Middle T expression in early lesions of *Lgr6-CreERT2*$^{+/-}$*:MMTV-PyMT*$^{+/-}$ mammary gland at P14. Lesions stained positive for PyMT (arrows) whereas normally appearing ducts were negative (arrowhead). Arrowheads in inset: PyMT^{-} cells embedded within PyMT^{+} region. Scale bars: 100 µm, 20µm (inset). **(d)** Middle T and K14 expression in pre-malignant mammary gland tissue of 2w-old (left panel) and mammary tumour of 15w-old (right) *Lgr6-CreERT2*$^{+/-}$*:MMTV-PyMT*$^{+/-}$ females. Only rare Middle T oncogene expressing cells are positive for EGFP (arrows) and EGFP^{+} cells negative for PyMT oncogene are found (arrowheads). Scale bars: 50 µm (left), 200 µm (right panel), 50 µm (inset). **(e)** Adjacent EGFP^{+}/tdTomato^{+} (yellow arrows/arrowheads) and tdTomato^{+} (red arrows/arrowheads) clones 2d post P28-induction. Green arrows/arrowheads depict EGFP^{+} cells. Scale bar: 50 µm. **(f)** GFP/EdU co-stained *Lgr6-CreERT2*$^{+/-}$*:MMTV-PyMT*$^{+/-}$ mammary glands at 2w (left) and 15w of age (right). Left: EdU^{+} cells expressing *Lgr6* (arrows). Scale bars: 50 µm. Right: Tumour regions with EdU^{+}/*Lgr6*$^{+}$ cells (arrows and inset). Scale bar: 200 µm, 25 µm (inset). **(g)** GFP immunostaining in mammary tumour of 24w-old PLT female induced at P29. Scale bars: 100 µm, 10 µm (inset). **(h)** Tumours from *PLT* females injected with sunflower-seed oil at 2w (left) or 4w (right) did not

present with tdTomato$^+$ cells. Scale bars: 1 mm. **(i)** Tamoxifen administration and time-points to analyse the contribution of *Lgr6*$^+$ cells to tumour maintenance. **(j)** *PLT* mammary tumours 1d (inset) and 1w p.i. Scale bars: 100 µm (inset), 100 µm (inset). **(k)** Quantification of tdTomato$^+$ area in *PLT* tumours after induction at 14w. (n=6 tumours pooled from 4 mice). Line indicates mean. **(l)** *Lgr6 in situ* hybridization combined with immunostaining 1w after MPA (left) or MPA/DMBA (right) treatment demonstrating basal (arrows) and luminal (arrowheads) *Lgr6*$^+$ cells in mammary glands undergoing chemical carcinogenesis. Scale bars: 25µm.

Supplementary Figure 7 - Lgr6^{+} cells contribute to tumour maintenance in the MMTV-PyMT model of breast carcinoma. **(a)** Experimental setup to analyse chemotherapeutic resistance of Lgr6-expressing PyMT tumour cells *in vitro*. Tumour-bearing *PLT* mice were treated with tamoxifen and tumour organoids were cultured 24hrs later. The organoids were treated with 2μM doxorubicin for 48hrs and allowed to recover until analysis. **(b)** Combined light microscopy and fluorescence images of tumour organoids comparing tdTomato tracing between untreated *PLT* organoids (upper panels) and doxorubicin-treated organoids (lower panels). tdTomato+ cells survive chemotherapeutic treatment (lower middle panel) and are able to form completely traced, healthy organoids upon recovery (lower right panel). **(c)** Extent of tdTomato tracing within untreated and doxorubicin-treated organoids. (n=8 pooled tumours from n=4 mice, collected in 3 independent experiments). Mean ± s.e.m. $^{**}P < 0.01$, $^{***}P < 0.001$. (multiple unpaired t-tests). Scale bars: 200Scale bars: 200m, 100μm (insets). **(d)** Scheme of *PDL* tumour cell isolation, allele recombination and DT treatment in vitro before colony formation analysis. Images show examples of tumour spheres from untreated (left) and DT-treated (right) cells. **(e)** mRNA expression for *β-actin* (*Actb*), *Lgr6* and *DTR* (normalized over *GAPDH*) in tamoxifen- and DT-treated tumour cells from *PDL* mice. (n = 3 wells per group from n = 3 independent mice). Mean ± s.d. $^{***}P < 0.001$ (unpaired two-tailed t-test) **(f)** Analysis of colony formation of primary *PDL* tumour cells after treatment with tamoxifen alone or tamoxifen plus diphtheria toxin. (n = 12 wells per group from n = 3 independent mice). Mean ± s.d. $^{*}P < 0.011$ (Two-way ANOVA). **(g)** H&E staining of tumours formed by control and DT-treated *PDL* primary tumour

cells injected subcutaneously into nude mice as shown in Figure 7g. Scale bars: 500μm (1x), 50μm (40x).

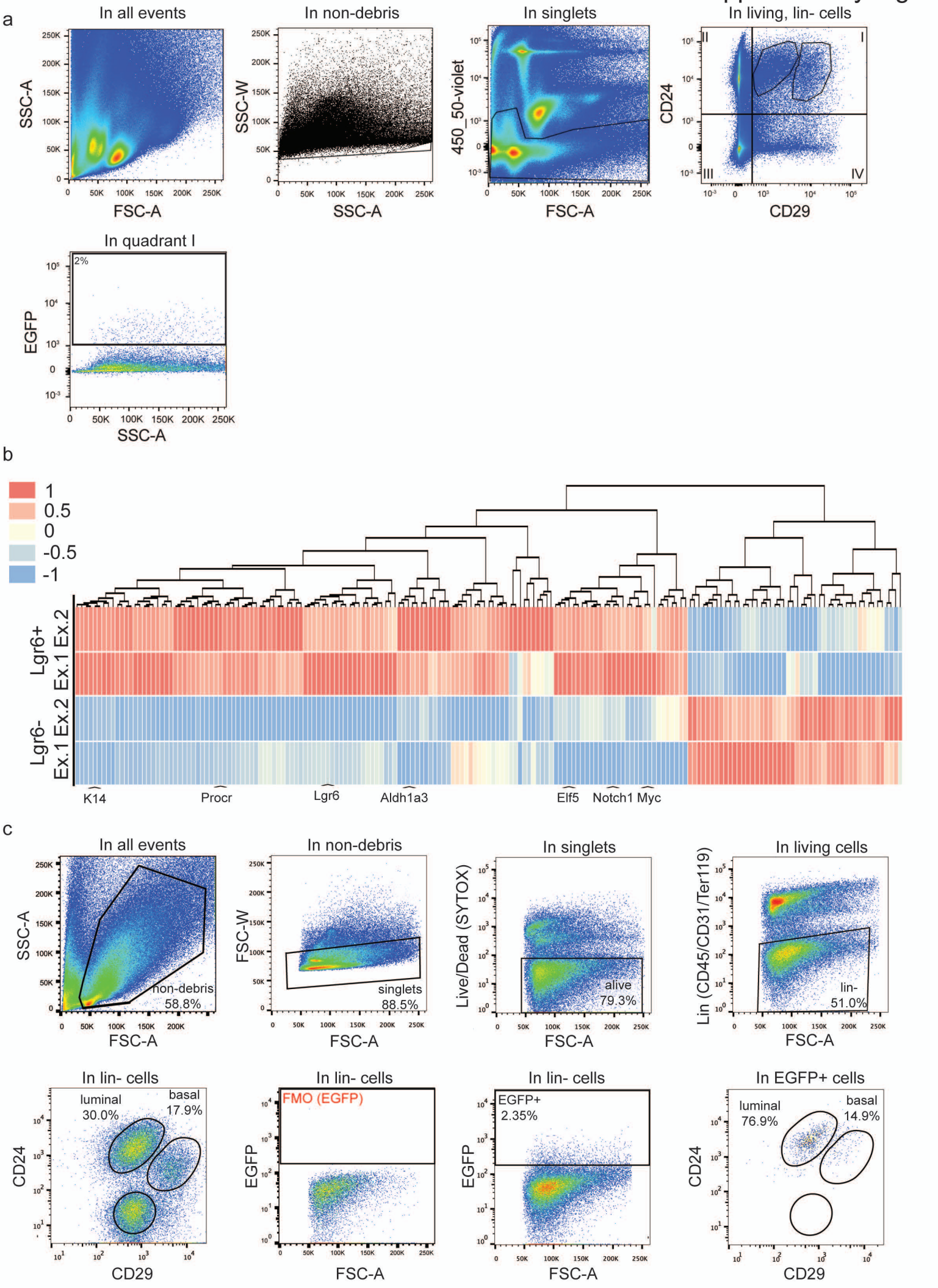
See Supplementary Table 2 for exact *p* values and source data for c, e, f.

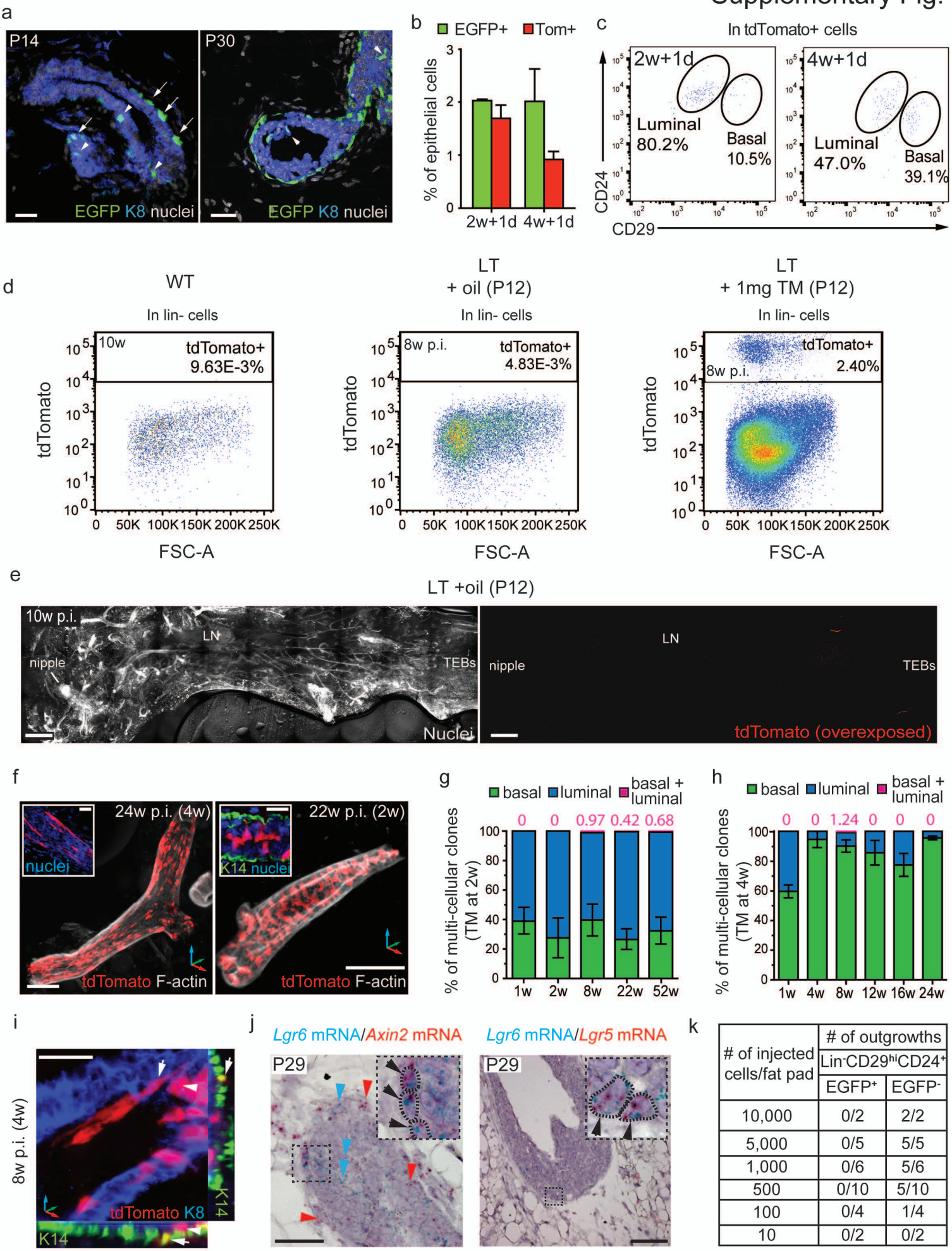
Supplementary Figure 8 - Scheme summarizing the contribution of *Lgr6*$^{+}$ cells to postnatal mammary gland development, pregnancy, and luminal mammary tumours. **(a)** During postnatal development, two populations of *Lgr6*$^{+}$ mammary epithelial progenitor cells (basal and luminal) contribute to ductal morphogenesis. In adult virgin females, however, *Lgr6*+ cells do not actively add to tissue homeostasis and their number decreases over time. Their clonal potential is reactivated by pregnancy (and stimulation with ovarian hormones) and they contribute to the alveolar network over multiple pregnancies. **(b)** *Lgr6*$^{+}$ progenitors can function as potent cells of origin of luminal mammary tumours. Transformation of *Lgr6*$^{+}$ progenitor cells with two different oncogenic combinations (loss of p53 and Brca1, or activation of mutant K-Ras^{G12D} and loss of Fbxw7, respectively) resulted in the formation of exclusively luminal mammary tumours, appearing with high penetrance. **(c)** *Lgr6*$^{+}$ cells contribute to luminal, but not basal mammary tumours. Lineage tracing showed that luminal *Lgr6*$^{+}$ progenitors contribute to MMTV-PyMT-induced tumours. In contrast, neither basal nor luminal Lgr6-expressing cells were traced in chemically induced, basaloid DMBA/MPA tumours.

Supplementary Table 1 - qRT-PCR Primer Sequences.

Supplementary Table 2 – Statistics source data. Source data for Fig. 1b, d, Fig. 2g – j, Fig. 4e, k, Fig. 6d, f, h, Fig. 7e, h, and Supplementary Fig. 7c, e, f.

Supplementary Fig. 1

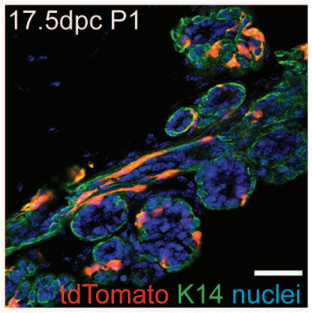




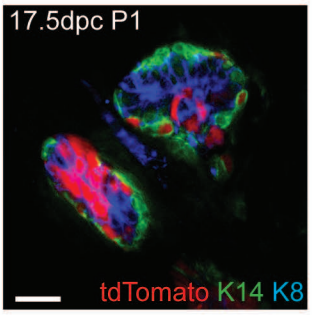
k

# of injected cells/fat pad	# of outgrowths Lin⁻CD29ʰⁱCD24⁺	
	EGFP⁺	EGFP⁻
10,000	0/2	2/2
5,000	0/5	5/5
1,000	0/6	5/6
500	0/10	5/10
100	0/4	1/4
10	0/2	0/2

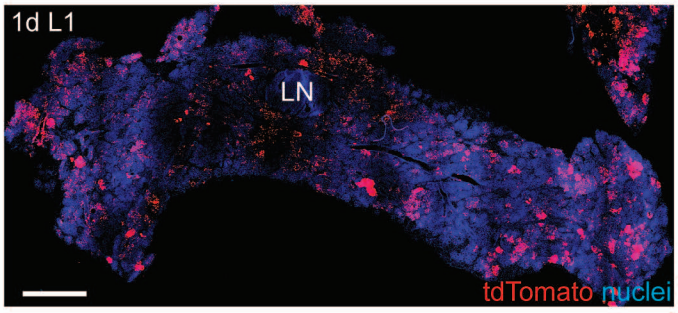
a


17.5dpc P1
tdTomato K14 nuclei

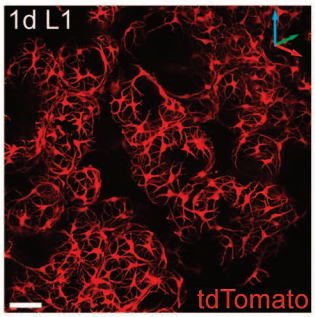
b


17.5dpc P1
tdTomato K14 K8

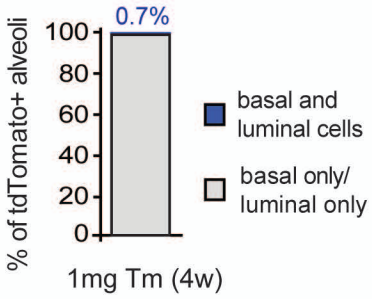
c


1d L1
LN
tdTomato nuclei

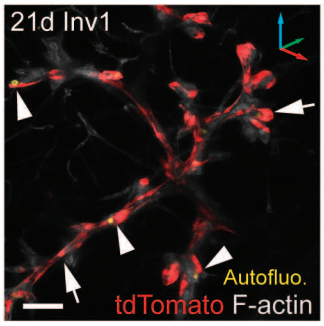
d


1d L1
tdTomato

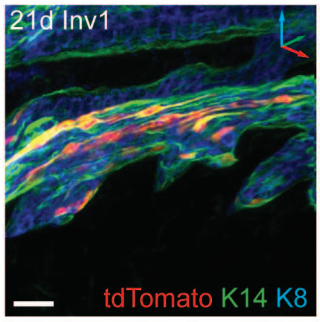
e


0.7%
100
80
60
40
20
0
% of tdTomato+ alveoli
basal and luminal cells
basal only/ luminal only
1mg Tm (4w)

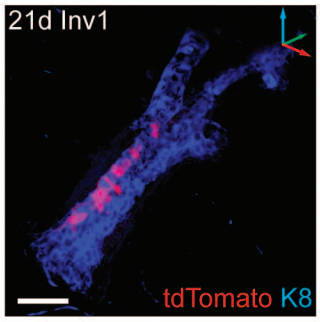
f

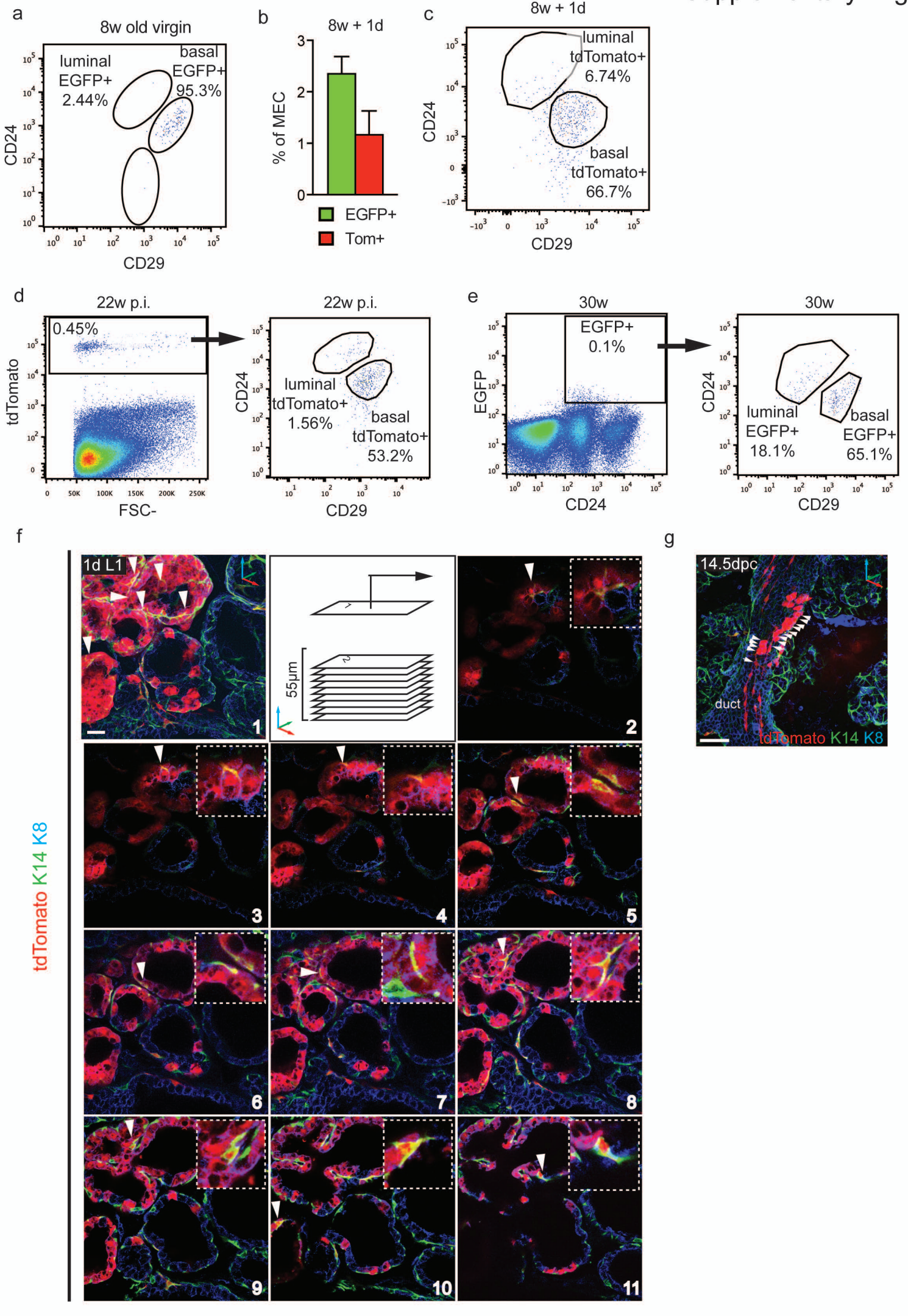

21d Inv1
Autofluo.
tdTomato F-actin

g

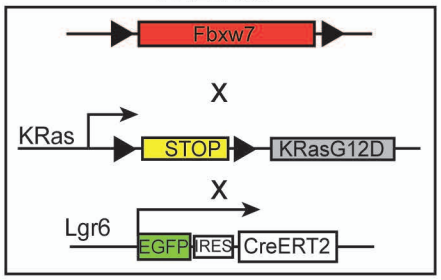

21d Inv1
tdTomato K14 K8

h

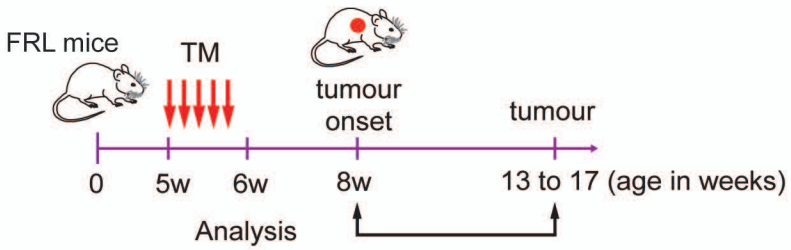

21d Inv1
tdTomato K8


a
8w old virgin
luminal
EGFP+
2.44%
basal
EGFP+
95.3%
CD24
CD29
b
8w + 1d
% of MEC
EGFP+
Tom+
c
8w + 1d
luminal
tdTomato+
6.74%
basal
tdTomato+
66.7%
CD24
CD29
d
22w p.i.
0.45%
tdTomato
FSC-
22w p.i.
luminal
tdTomato+
1.56%
basal
tdTomato+
53.2%
CD24
CD29
e
30w
EGFP+
0.1%
EGFP
CD24
30w
luminal
EGFP+
18.1%
basal
EGFP+
65.1%
CD24
CD29
f
1d L1
55µm
tdTomato K14 K8
g
14.5dpc
duct
tdTomato K14 K8

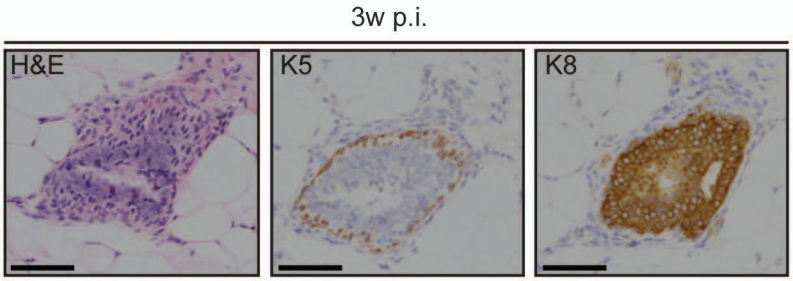
a



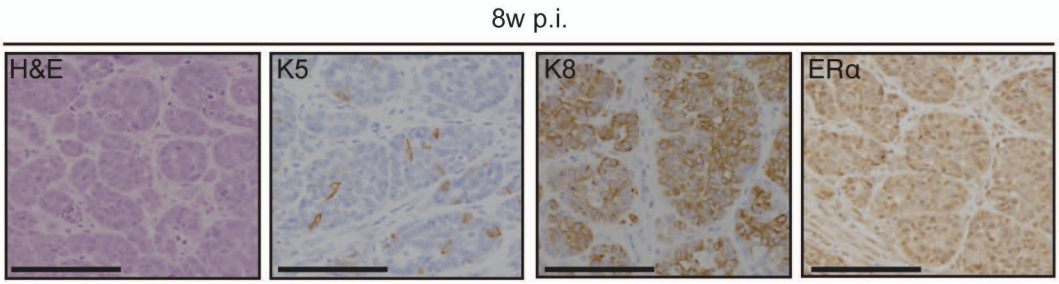
b



c

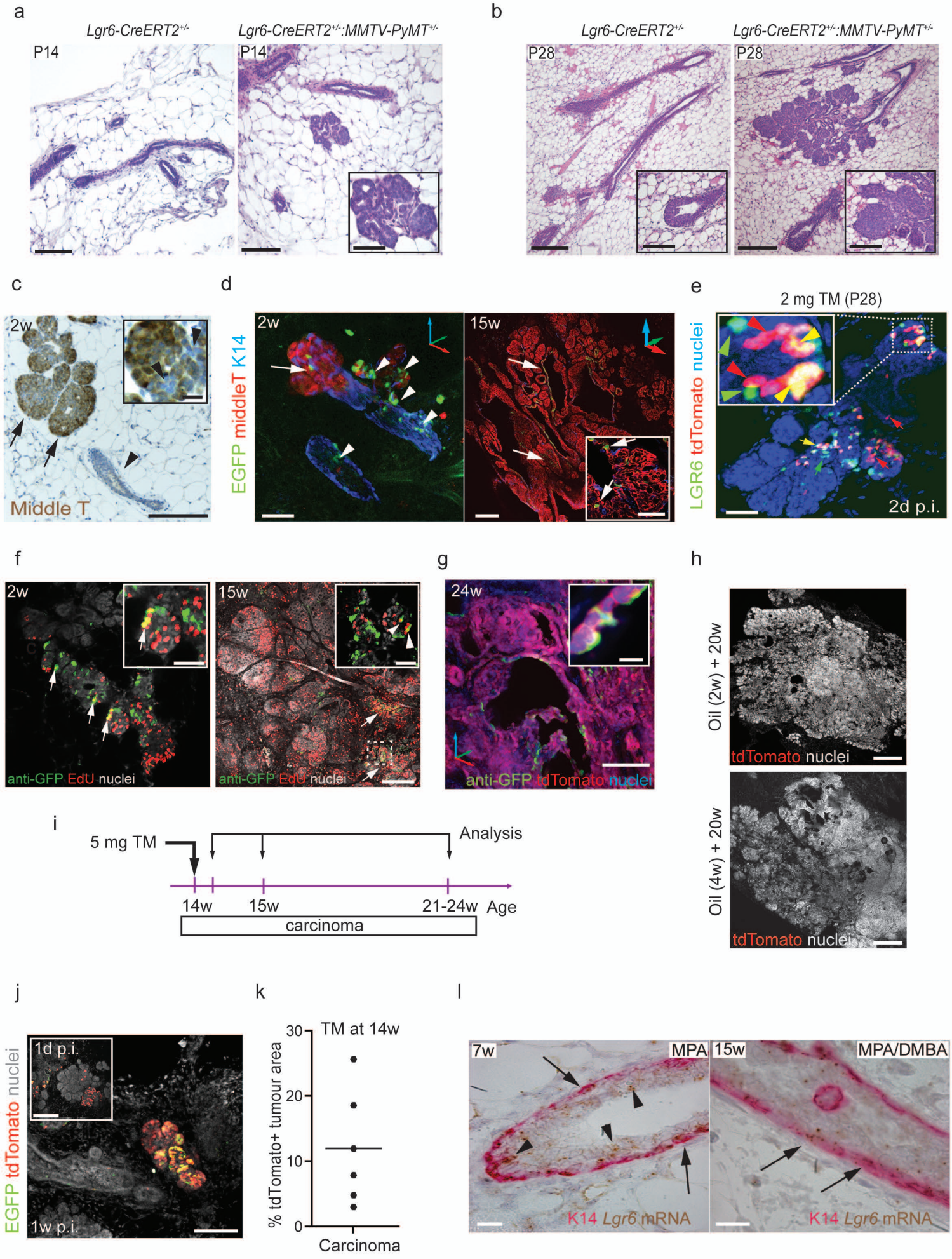


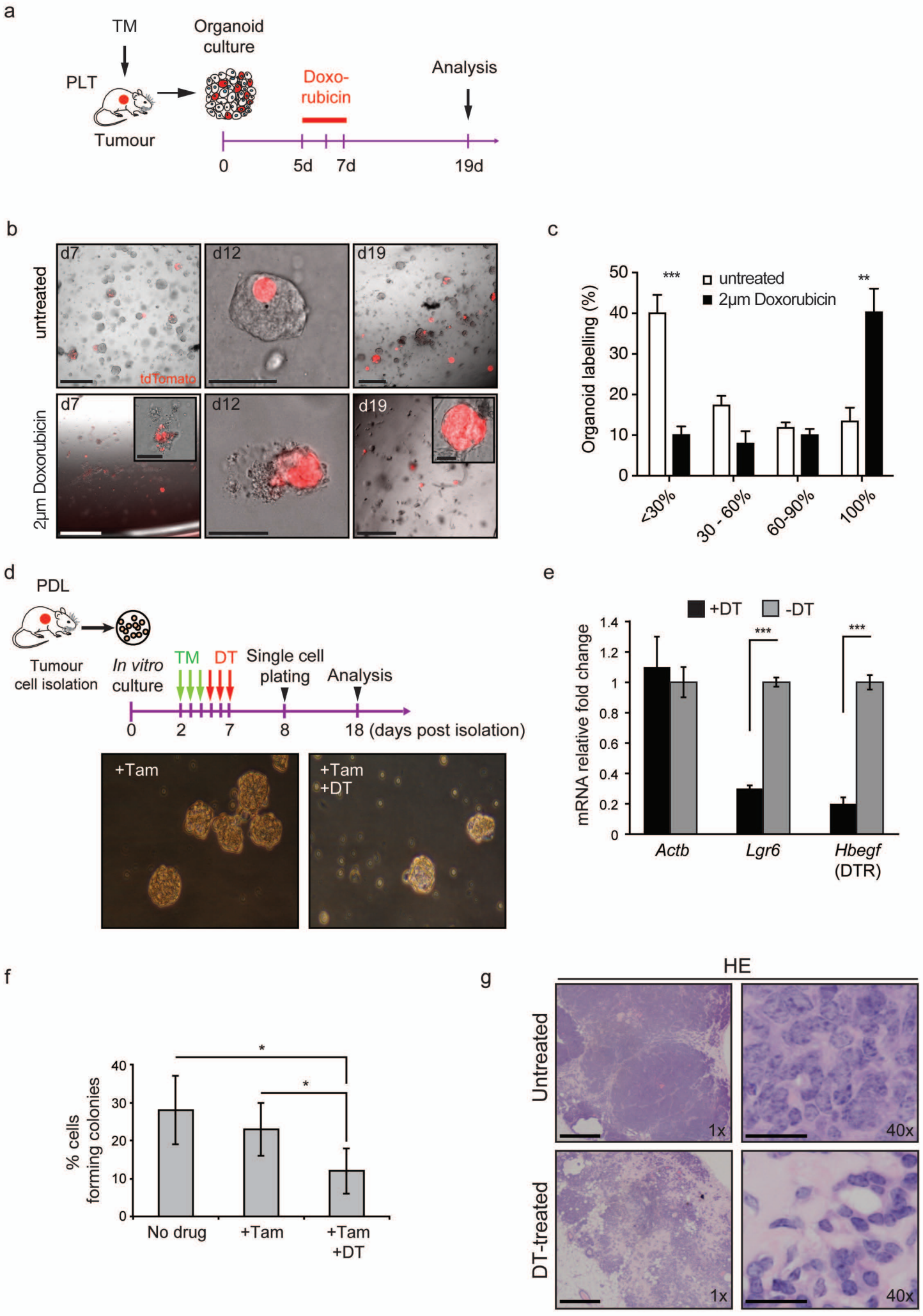
d

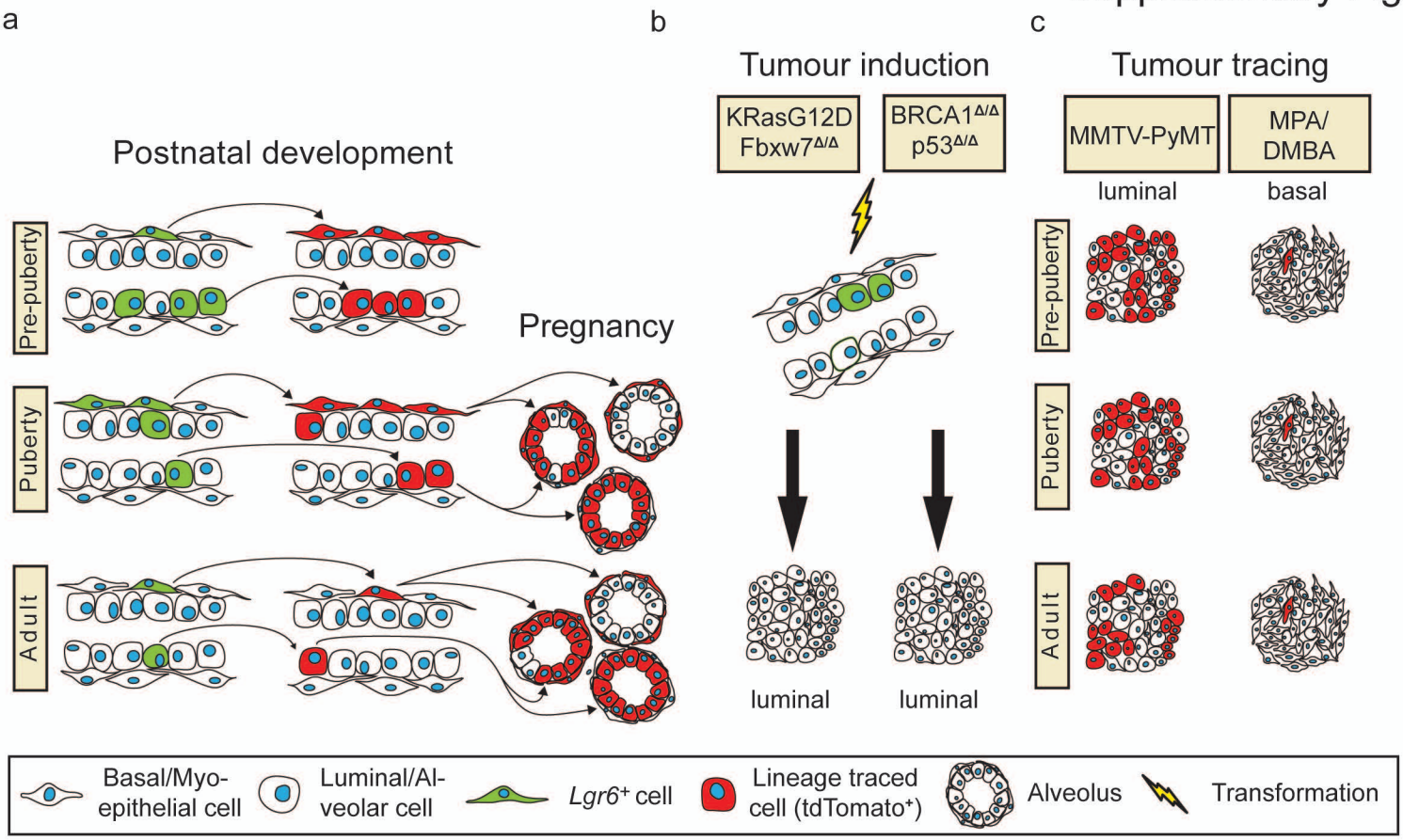


e

	FRL	BPL
Total number of animals	7	10
Penetrance (number of animals harbouring lesions when sacrificed)	100% (7 out of 7)	60% (6 out of 10)
Total number of K8+ tumour lesions	100% (54 over 54 counted)	100% (32 over 32 counted)
Total number of K5+ tumour lesions	0% (0 out of 54)	0% (0 out of 32)


a
Lgr6-CreERT2+/-
Lgr6-CreERT2+/-:MMTV-PyMT+/-
P14
P14
b
Lgr6-CreERT2+/-
Lgr6-CreERT2+/-:MMTV-PyMT+/-
P28
P28
c
2w
Middle T
d
2w
15w
EGFP middleT K14
e
2 mg TM (P28)
LGR6 tdTomato nuclei
2d p.i.
f
2w
15w
anti-GFP EdU nuclei
anti-GFP EdU nuclei
g
24w
anti-GFP tdTomato nuclei
h
Oil (2w) + 20w
tdTomato nuclei
Oil (4w) + 20w
tdTomato nuclei
i
5 mg TM
Analysis
14w
15w
21-24w
Age
carcinoma
j
EGFP tdTomato nuclei
1d p.i.
1w p.i.
k
TM at 14w
% tdTomato+ tumour area
30
20
10
0
Carcinoma
l
7w
MPA
K14 Lgr6 mRNA
15w
MPA/DMBA
K14 Lgr6 mRNA


a
TM
Organoid culture
PLT
Tumour
Doxo-rubicin
Analysis
0
5d
7d
19d
b
untreated
2µm Doxorubicin
d7
d12
d19
tdTomato
c
Organoid labelling (%)
untreated
2µm Doxorubicin
***
**
<30%
30 - 60%
60-90%
100%
d
PDL
Tumour cell isolation
In vitro culture
TM
DT
Single cell plating
Analysis
0
2
7
8
18 (days post isolation)
+Tam
+Tam +DT
e
mRNA relative fold change
+DT
-DT
***
***
Actb
Lgr6
Hbegf (DTR)
f
% cells forming colonies
*
*
No drug
+Tam
+Tam +DT
g
HE
Untreated
DT-treated
1x
40x
1x
40x


a
Postnatal development
Pre-puberty
Puberty
Adult
Pregnancy
b
Tumour induction
KRasG12D
Fbxw7Δ/Δ
BRCA1Δ/Δ
p53Δ/Δ
luminal
luminal
c
Tumour tracing
MMTV-PyMT
MPA/
DMBA
luminal
basal
Pre-puberty
Puberty
Adult
Basal/Myo-epithelial cell
Luminal/Al-veolar cell
Lgr6+ cell
Lineage traced cell (tdTomato+)
Alveolus
Transformation