ARTICLE

doi:10.1038/nature12930

Alveolar progenitor and stem cells in lung development, renewal and cancer

Tushar J. Desai[1,2], Douglas G. Brownfield[1] & Mark A. Krasnow[1]

Alveoli are gas-exchange sacs lined by squamous alveolar type (AT) 1 cells and cuboidal, surfactant-secreting AT2 cells. Classical studies suggested that AT1 arise from AT2 cells, but recent studies propose other sources. Here we use molecular markers, lineage tracing and clonal analysis to map alveolar progenitors throughout the mouse lifespan. We show that, during development, AT1 and AT2 cells arise directly from a bipotent progenitor, whereas after birth new AT1 cells derive from rare, self-renewing, long-lived, mature AT2 cells that produce slowly expanding clonal foci of alveolar renewal. This stem-cell function is broadly activated by AT1 injury, and AT2 self-renewal is selectively induced by EGFR (epidermal growth factor receptor) ligands *in vitro* and oncogenic *Kras(G12D) in vivo*, efficiently generating multifocal, clonal adenomas. Thus, there is a switch after birth, when AT2 cells function as stem cells that contribute to alveolar renewal, repair and cancer. We propose that local signals regulate AT2 stem-cell activity: a signal transduced by EGFR-KRAS controls self-renewal and is hijacked during oncogenesis, whereas another signal controls reprogramming to AT1 fate.

Pulmonary gas exchange occurs in delicate alveolar sacs lined by two epithelial cell types[1] (Extended Data Fig. 1). Squamous alveolar type (AT) 1 cells mediate gas exchange, whereas cuboidal AT2 cells secrete surfactant that prevents alveolar collapse; AT2 cells are one of the medically most important cells in the neonate and, as described below, one of the most dangerous in adults. Serious diseases including respiratory distress syndrome and idiopathic pulmonary fibrosis involve a failure to establish or maintain AT1 and AT2 cells, and alveoli are a major site of lung cancer, the leading cause of cancer death[2].

Despite their importance, the identity of alveolar progenitor and stem cells is controversial and their activity throughout life uncharted[3,4]. Classical morphologic and autoradiographic studies in rodents suggested that progenitors mature into AT2 cells during development, some of which differentiate into AT1 cells[5]. Maintenance is difficult to study because of slow turnover[6,7], but lineage tag expression in isolated AT1 cells is observed following bulk labelling of AT2 cell populations[8,9]. To circumvent slow turnover, lung injury models have been used and provide evidence that AT2 cells can contribute to alveolar repair[8,10,11]. Currently, six different cell populations have been proposed as alveolar stem cells on the basis of their capacity for clonal propagation and multilineage differentiation in culture[9,12–14]. Transplantation assays have also been used, but because many cells are implanted they cannot assess whether individual cells self-renew and undergo multilineage differentiation[4,15]. When one of these putative stem cell populations was fate-mapped *in vivo*, it failed to demonstrate the multilineage differentiation achieved in culture[16]; similar disparity between *ex vivo* and *in vivo* behaviour of putative stem cells has been found for other organs[17]. Here we use a battery of alveolar markers, lineage tracing and clonal analysis in mice to identify alveolar progenitor and stem cells *in vivo* and map their locations and activity during lung development, maintenance and cancer.

AT1 and AT2 cells arise from a bipotent progenitor

Mature AT1 and AT2 cells appear about 1 day before birth, when distal tubules begin to dilate ('sacculation', Fig. 1a–c)[18,19]. We mapped progression of sacculation in three dimensions by analysing finely staged whole-mount lungs immunostained for E-cadherin (Cdh1) to visualize individual cells (Fig. 1a–c and Extended Data Fig. 1e–f). Dilation begins at the bronchoalveolar junction then progresses distally towards the airway tip (Fig. 1a–c).

The classical model proposing that progenitors in development are pre-AT2 cells is difficult to reconcile with the finding that some AT1 cell markers are expressed up to 5 days before sacculation[20]. To molecularly classify progenitors, we validated 15 extant AT1 and AT2 markers (Supplementary Table 1) then analysed the transition in labelling between distal (progenitors) and proximal (nascent AT1 and AT2 cells) positions in a sacculating airway (Fig. 1d) to infer dynamic expression changes during differentiation (Fig. 1e–p). Markers fell into six expression classes (Extended Data Table 1), distinguishing seven stages in alveolar development (Fig. 1e–p). However, instead of a progenitor to AT2 to AT1 progression, our data support a model in which bipotent progenitors (P) expressing a subset of AT1 (1) and AT2 (2) markers ($P1^E$, $P1^L$, $P2^E$ and $P2^L$) give rise to either AT1 or AT2 cells by shutting off inappropriate cell type markers early (E) or late (L) in differentiation, then turning on cell type-specific late (L) markers ($A1^L$, $A2^L$) as they complete maturation (Fig. 1q). Co-expression of AT1 and AT2 markers by progenitors indicates that these specialized cell types may have evolved from a primordial pneumocyte with features of both, similar to those in less derived vertebrates such as lungfishes[21].

Three additional lines of evidence support the bipotent progenitor model. First, clonal analysis of individual distal airway epithelial tip cells[22] labelled on embryonic day (E) 15 using an inducible Cre recombinase (linked with the oestrogen receptor (ER), Shh–Cre–ER) demonstrated localized alveolar lineage clusters with marked AT1 and AT2 cells (Extended Data Fig. 2a,b), confirming that individual cells are bipotent. Second, ultrastructural analysis of early sacculation revealed three classes of distal epithelial cells (Fig. 1r–u): cuboidal cells with glycogen vacuoles but no lamellar bodies (bipotent progenitors), cuboidal cells with vacuoles and lamellar bodies (early AT2 cells), and partially flattened cells with vacuoles (early AT1 cells). We never observed partially flattened cells with lamellar bodies, the presumed AT2-to-AT1 intermediate predicted by the classical model (Extended Data Fig. 1g). Third, lineage

[1]Department of Biochemistry and Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305-5307, USA. [2]Department of Internal Medicine, Division of Pulmonary and Critical Care, Stanford University School of Medicine, Stanford, California 94305-5307, USA.

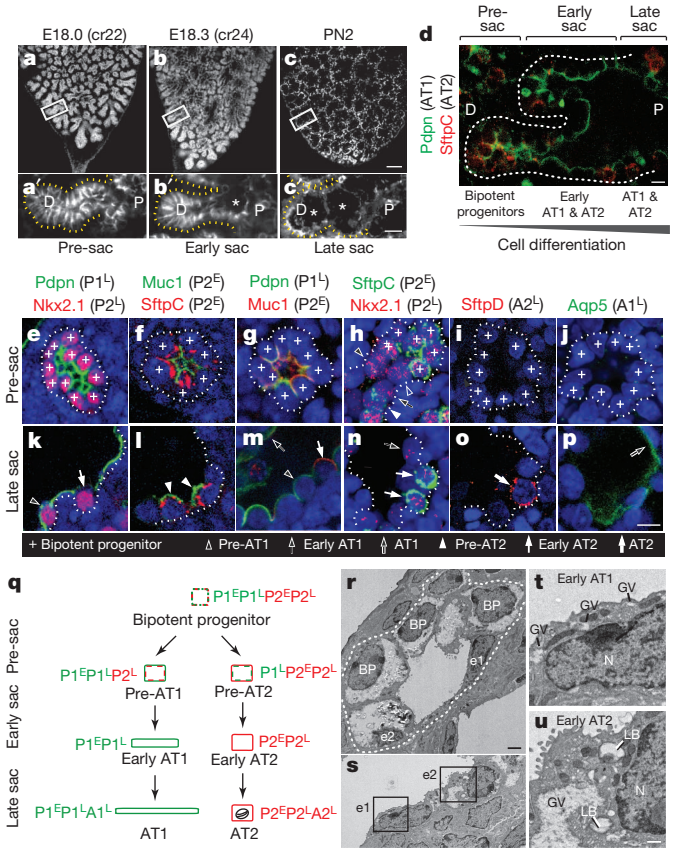


Figure 1 | Development of alveolar type 1 (AT1) and AT2 cells from bipotent progenitors. **a**–**c**, Mouse lung lobe tips stained for E-cadherin (E, embryonic day; cr, crown-rump length in mm; PN, postnatal day). **a'**–**c'**, Sacculation (sac; asterisks) proceeds proximally (P) to distally (D) along the airway. Scale bars, 100 μm (**a**–**c**), 20 μm (**a'**–**c'**). **d**, Staining for AT1 (Pdpn) and AT2 (SftpC) markers at E18.3 shows coexpression in pre-sacculation zone and restriction in late sacculation zone. Scale bar, 10 μm. **e**–**p**, Tips imaged in pre-sacculation (**e**–**j**) and late sacculation (**k**–**p**) zones identify six classes of marker expression profiles ($P1^E$,$P1^L$,$A1^L$,$P2^E$,$P2^L$,$A2^L$; Extended Data Table 1). Scale bar, 20 μm. **q**, Inferred differentiation pathways showing changes in AT1 (green) and AT2 (red) marker classes. Oval, lamellar body. **r**, **s**, Electron micrographs at E18.3 showing early (**r**) and late (**s**) sacculation zones. BP, bipotent progenitors; e1, early AT1 cells; e2, early AT2 cells. Scale bar, 2 μm. **t**, **u**, Boxed regions in S showing (e1) an early AT1 (squamous, glycogen vacuoles (GV) without lamellar bodies (LB)) and (e2) an early AT2 (cuboidal, GV, LB). N, nucleus; bar, 0.5 μm.

tracing of newly differentiated AT2 cells using a Cre recombinase knock-in (LysM–Cre) at the lysozyme M (*Lyz2*) locus, a late AT2 marker gene ($A2^L$), labelled many AT2 cells in the embryo but no marked AT1 cells were seen by 2 weeks postnatally (Extended Data Fig. 2c, e).

Little epithelial proliferation was detected during sacculation and for several weeks afterward (Extended Data Fig. 3), indicating that maturation of bipotent progenitors generates most or perhaps all AT1 and AT2 cells in development. Flow cytometry at early sacculation (E18.1) showed bipotent progenitors or pre-AT2 cells ($Muc1^+$ $Pdpn^+$) made up 8% of distal cells, whereas early or mature AT2 cells ($Muc1^+$ $Pdpn^-$) made up 5%. Bipotent progenitors seem to be fully exhausted by postnatal day 4, when sacculation has completed throughout the lung. In adult lungs few if any bipotent progenitors were found by flow cytometry (0.3% ± 0.2% $Muc1^+$ $Pdpn^+$ distal epithelial cells) or labelling with Shh–Cre–ER, suggesting a transition to other sources.

Alveolar renewal foci by rare activation of AT2 cells

Classical autoradiographic studies suggested alveolar epithelium is a slowly renewing population maintained by diffuse proliferation, presumably of AT2 cells[23]. We used *in vivo* lineage tracing to investigate the role of AT2 cells in maintenance, marking them in two complementary ways, with similar results. In one, a Cre–ERT2–rtTA knock-in at surfactant protein C (*SftpC*) locus (SftpC–Cre–ER) was used with a membrane-localized fluorescent reporter (mTmG) to pulse-label AT2 cells by tamoxifen induction at postnatal (PN) 18 days; mice were analysed 13 or 192 days later. Although SftpC is often used as a mature AT2 marker, the results above (Extended Data Table 1) show that as a class $P2^E$ marker it is also expressed by the bipotent progenitor. Hence, we also labelled AT2 cells using LysM–Cre described above, a class $A2^L$ marker that initiates expression only in mature AT2 cells (Extended Data Fig. 4a, b, d, m). Both lines robustly labelled AT2 cells throughout the lung, and at 2 months after marking (SftpC–Cre–ER) or postnatally (LysM–Cre) showed scattered AT1 cells labelled with the AT2 lineage tag (Extended Data Fig. 2d, f and Extended Data Fig. 5). AT2 cells did not significantly contribute to maintenance of bronchiolar lineages including Clara, ciliated and neuroendocrine cells (Extended Data Fig. 4c, e–m), and although rare cells near bronchoalveolar junctions that co-expressed SftpC and Clara cell marker CCSP/Scgb1a1 were also tagged ('bronchoalveolar stem cells' or BASCs[12]), they proliferated little if at all (Extended Data Fig. 4f–h, j–m).

To investigate the frequency and spatial localization of AT1 renewal by AT2 cells, we labelled lungs as above using LysM–Cre and analysed them at 1, 2, 4, 8 and 16 months of age. Less than 1% of AT1 cells expressed the AT2 lineage tag at 1 month, 3.9% at 4 months and 7.5% at 16 months (Fig. 2a–c). AT1 cell replacement occurred preferentially in alveoli abutting arterioles and in the lung periphery (Fig. 3a, b). Wherever new AT1 cells arose, the alveolus was replaced essentially in half or its entirety, indicating alveoli include just one or two AT1 cells. With ageing, adjacent alveolar units were also renewed, evidenced by slowly enlarging clusters of AT1 cells marked with the AT2 lineage tag. These patches of AT1 replacement were indistinguishable from surrounding areas, except for AT1 labelling with the AT2 lineage tag; similar results were obtained with SftpC–Cre–ER (Extended Data Fig. 5). We conclude that AT1 replacement by AT2 cells occurs intermittently and in a spatially patchy distribution in 'renewal foci' that slowly enlarge over time, with perivascular and peripheral regions serving as relative 'hot spots.'

Renewal foci derive from single, differentiated AT2 cells

Each renewal focus could derive from a single AT2 cell, in which case the cells would be clonally related, or from multiple AT2 cells. To distinguish these possibilities, we used a Confetti Cre-dependent reporter that stochastically expresses one of four fluorescent proteins in each cell that undergoes recombination. We analysed the membrane CFP lineage marker and found that foci were similar in size and progressive enlargement to those observed using the single-colour reporter, indicating that renewal foci derive from a single 'founder' AT2 cell. Because of fortuitously inefficient recombination of the Confetti reporter using LysM–Cre (~2% of AT2 cells expressed the membrane-targeted CFP lineage tag at 16 months), we were able to distinguish marked AT2 cells within renewal foci. Most contained one or two AT2 cells and multiple AT1 cell progeny colonizing up to six contiguous alveoli (Fig. 2d–f). We also observed small clonal foci comprising only AT2 cells (Extended Data Fig. 6b, c), indicating they are capable of dividing without forming AT1 cells (self-duplication). Founder AT2 cells expressed mature AT2 markers including LAMP-1, a lysosome-associated protein, indicating they continue to produce surfactant (Extended Data Fig. 6a), and they did not express Clara, ciliated and neuroendocrine markers. Their only distinguishing feature was their nucleus, which was often slightly larger and stained more intensely for Nkx2.1 than nearby AT2 cells (see also Fig. 2d, e). We conclude that each renewal focus derives from a single, self-renewing AT2 cell that can generate multiple AT1 and/or AT2 cells, is long-lived, and remains closely associated with its progeny, as evidenced by its persistence in large foci at advanced ages (Fig. 2e). Because founder cells appear to be mature (LysM–Cre lineage positive), functional (SftpC–Cre–ER lineage and LAMP-1 positive) AT2 cells that share the essential properties of conventional stem cells (multipotent,

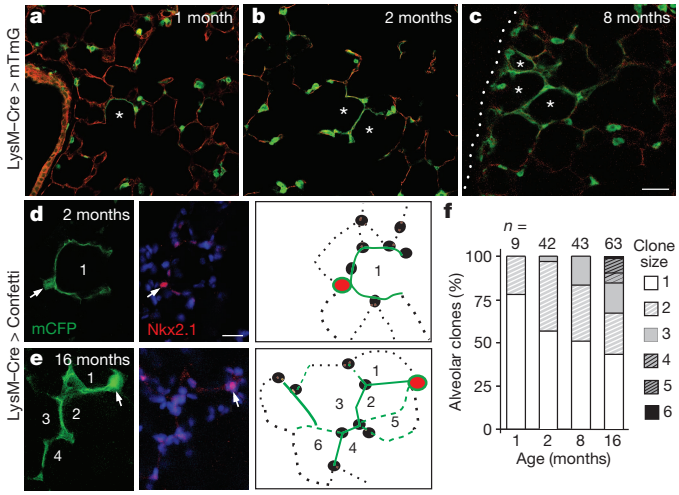


Figure 2 | Mature AT2 cells renew AT1 cells in clonal foci. a–c, AT2 lineage-labelled AT1 cell foci (green) enlarge with ageing, incorporating adjacent alveoli (asterisks). Dotted line, mesothelium. Scale bar, 150 µm. **d**, **e**, AT2 cells were sparsely marked with Confetti reporter, and mCFP-labelled foci (green) are shown co-stained for AT2 marker Nkx2.1 (red) at age 2 (**d**) and 16 (**e**) months. Note 'founder' AT2 cell (arrow) and its labelled AT1 progeny (green). Numbers, incorporated alveoli. Entire clone schematized in right panel. Red, founder AT2; green lines, AT1 daughters visible (solid) or outside (dashed) focal plane; black, unlabelled AT1 (dotted lines) and AT2 (ovals) cells. Scale bar, 20 µm (**d**, **e**). **f**, Clone size increases with ageing ($P = 0.03$, Kruskal–Wallis test). *n*, clones scored; clone size, number of incorporated alveoli.

self-renewing, persist for life), they are 'bi-functional' stem cells, executing both differentiated and regenerative functions.

AT2 cell activation by acute AT1 cell injury

Elevated oxygen tension is toxic to AT1 cells, but not to AT2 cells[24]. Exposure of 2 month old mice carrying the AT2 lineage tag to 88% oxygen for 120 h tripled the number of renewed AT1 cells (Fig. 3c–e). This finding shows that AT2 cells become activated following hyperoxic injury, and supports the idea that although normally only a rare subset of AT2 cells executes a stem cell function, others can be recruited to repair alveolar damage. Whether every AT2 cell can be activated this way could not be determined because more severe hyperoxia was lethal.

Oncogenic Kras selectively activates AT2 self-renewal

Adenocarcinoma, the major form of lung cancer, is associated with activating mutations in *Kras* or *Egfr*[25]. Typically located in peripheral lung regions, nearly all tumour cells express SftpC, leading to long-standing speculation that they originate from AT2 cells or their progenitors. However, the identity of the tumour-initiating cell(s) remains controversial[12,26–28]. To test the effect of oncogenic *Kras(G12D)* on mature AT2 cells, a conditional $Kras^{LSL\text{-}G12D}$ allele ($Kras^{LSL\text{-}G12D}$ is a knock-in at the *Kras* locus in which the wild-type *Kras* coding sequence is replaced by a lox-STOP-lox-*Kras(G12D)*) was activated using LysM-Cre along with the mTmG Cre-dependent reporter. Tumour nodules grew rapidly throughout the lungs (Fig. 4a–d), with dense replacement of virtually the entire alveolar region by 1 month after induction and death shortly thereafter (Fig. 4e). When lungs were examined in the first few days following induction, we found with a Rainbow multi-colour reporter that nearly every epithelial cell expressing the AT2 lineage tag proliferated, demonstrating highly efficient AT2 transformation by *Kras(G12D)* (Fig. 4f–i). The biggest tumours were found in peripheral and perivascular regions, sites where physiological AT1 renewal by AT2 cells was commonly observed (compare Figs 3a, b and 4o, p). Similar results were obtained using SftpC–Cre–ER to activate *Kras(G12D)* in adult mice (Extended Data Fig. 7a, c). By contrast, when we used CCSP–Cre–ER, most Clara cells were unaffected or divided minimally, whereas at bronchoalveolar junctions some formed small clonal adenomas (Extended Data Fig. 7b). We also used ubiquitously expressed ROSA–Cre–ER to activate the *Kras(G12D)* allele at random, resulting in many singlets and minimally affected cells throughout the lung, even 18 days after induction (Fig. 4j). The transformed AT2 cells comprising the adenomas continued to express AT2 markers (Nkx2.1, SftpD) and did not turn on a Clara marker (CCSP) or, with rare exceptions, AT1 markers (Pdpn, LEL) (Extended Data Fig. 8). Thus, oncogenic *Kras(G12D)* seems to selectively and permanently induce AT2 self-renewal, without deprogramming the cells to the bipotent progenitor or causing reprogramming to AT1 or Clara cell fates.

By examining lineage-tagged *Kras(G12D)* mutant lungs at progressive stages, we could infer the cellular mechanism of adenoma formation (Fig. 4k–m). Proliferation of the activated AT2 cell generates daughter cells that spread laterally yet maintain a monolayer ('lepidic' expansion), the first histological sign of the tumour (Fig. 4l). Later, cells heap up and form a nodule ('hilical' expansion) that obliterates the lumen and begins compressing and invading adjacent alveoli (Fig. 4m). Infrequent

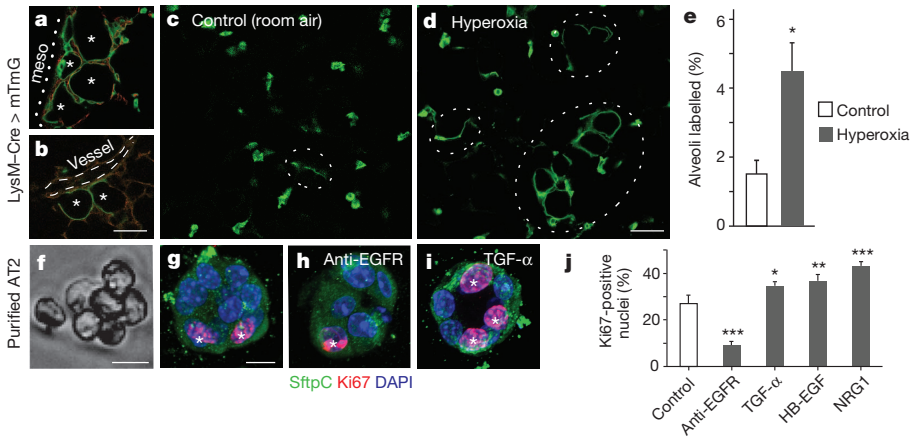


Figure 3 | Activation of AT2 stem cell function *in vivo* and proliferation *in vitro*. a, **b**, Renewal foci (AT2 lineage label, green; other cells, red) commonly involve alveoli (asterisks) in peripheral (**a**, mesothelium, dots) and perivascular (**b**, dashes) domains. Scale bar, 50 µm. **c**–**e**, Alveolar regions under room air (**c**) or after hyperoxia (88% O_2, 5 days) to injure AT1 cells (**d**). Note increased foci (dashed ovals) after injury. Scale bar, 50 µm. Quantification (**e**) shows increased alveolar surface ($n=2$; mean ± s.e.m.) from AT2-lineage labelled cells after hyperoxia. $P < 0.05$ (Mann–Whitney U test). **f**, **g**, Freshly isolated AT2 cells (**f**, phase contrast) cultured 4 days in Matrigel (**g**) proliferate, shown by Ki67 staining (red, asterisks), but maintain AT2 marker expression (SftpC, green). **h**–**j**, Images (**h**, **i**) and quantification (500 cells per biological replicate, $n = 4, 3, 3, 3, 4$) (**j**) of proliferation with EGFR blocking antibody (anti-EGFR, 2.5 µg ml^{-1}) and EGF ligands indicated (4 µM). Scale bar, 10 µm (**f**–**i**). Mean ± s.e.m.; $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ (Tukey's multiple comparisons test).

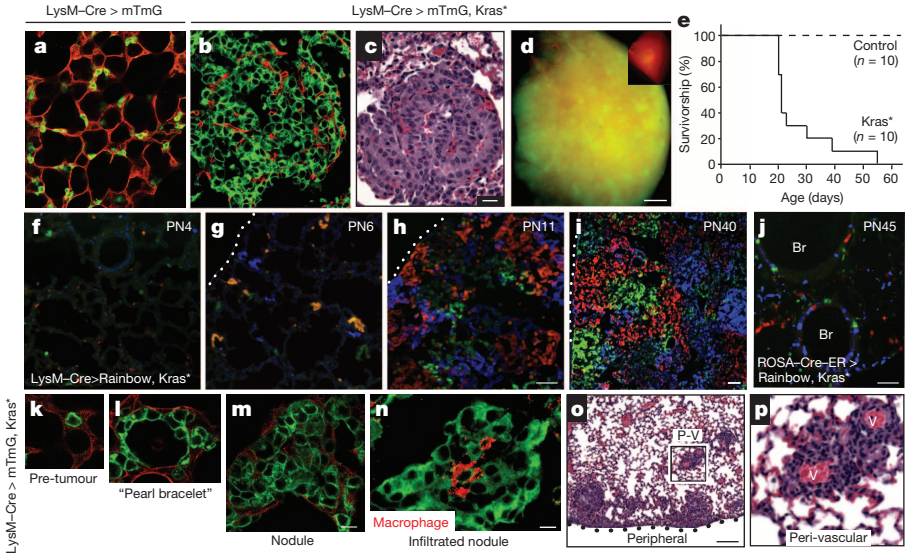


Figure 4 | Transformation of mature AT2 cells by activated Kras. **a**–**c**, LysM-Cre > mTmG control (**a**) and LysM-Cre > mTmG,Kras$^{LSL\text{-}G12D/+}$ (abbreviated Kras*) (**b**, **c**) lungs at age 7 weeks, showing proliferated AT2 cells (**b**, green cells) compressing surrounding cells (red) and forming adenomas (**c**, haematoxylin and eosin stain). Scale bar, 20 μm. **d**, Lung lobe as in **b** showing widespread infiltration by tumour (green). Inset, control lobe. Scale bar, 1 mm. **e**, Survival curves. **f**–**i**, LysM-Cre > Rainbow,Kras* lungs at indicated ages (PN, postnatal day). Note rapid clonal (single colour) expansion of labelled AT2 cells with minimal cell mixing. Scale bar, 50 μm (**f**–**h**), 100 μm (**i**). Dotted line, mesothelium. **j**, PN45 ROSA–Cre–ERT2> Rainbow,Kras* lung 18 days after induction. Note minimal expansion. Br, bronchi; scale bar, 50 μm. **k**–**n**, A recombined (green) AT2 cell (**k**) generates progeny that spread laterally, giving 'pearl bracelet' appearance (**l**). Cells later 'heap up' into nodules (**m**), some infiltrated by macrophages (**n**; red, anti-F4/80). Scale bar, 10 μm. **o**, **p**, Haematoxylin and eosin stain shows robust tumour growth peripherally and perivascularly (P-V). Boxed area (**o**, close-up in **p**) shows proliferated cells around blood vessels (v). Scale bar, 200 μm.

clones developed into large, well-formed papillary structures, although most formed simple adenomas[29]. At advanced stages, a notable pattern of large, closely packed, single-colour tumours was evident, showing minimal cellular exchange between neighbouring tumour foci (Fig. 4i). Some advanced tumours became infiltrated by macrophage 'nests' (Fig. 4n), a poor prognostic sign[30].

EGFR activity controls AT2 self-renewal in culture

We profiled the transcriptome of bipotent progenitors and AT2 cells. Both express *Egfr*, two other EGFR family members (*Erbb2* and *Erbb3*) (Supplementary Table 2), and receptors for many other signals (Supplementary Tables 2 and 3). Purified AT2 cells (Fig. 3f) retain a robust ability to both proliferate (>25% Ki67-positive, Fig. 3g, j) and to differentiate into AT1-like cells (>95% thin, flat morphology and aquaporin 5-positive, $n = 240$ cells) (Extended Data Fig. 6d). Addition of purified EGF ligands (transforming growth factor α (TGF-α), heparin-binding EGF (HB-EGF), NRG1) stimulated AT2 proliferation, whereas EGFR-blocking antibody inhibited proliferation (Fig. 3h–j). None of these treatments induced differentiation into AT1 cells. EGF signalling is therefore a critical and selective regulator of AT2 proliferation, at least under these conditions.

Discussion

Our characterization of alveolar progenitor and stem cells throughout the lifespan supports a model in which AT1 and AT2 cells arise independently during development from a bipotent progenitor (Fig. 5a). Several weeks after birth, when alveolar development is complete, there is a switch and mature AT2 cells become a renewable source of AT1 and AT2 cells. But these postnatal renewal events are rare and occur in monoclonal foci ('alveolar renewal focus', Fig. 5b) that slowly expand over months or years. Only a small fraction (~1%) of mature AT2 cells normally express this stem cell function and they divide intermittently (~40-day doubling time) and supply only local domains, giving an overall renewal rate of just 7% of alveoli per year. This stem cell function is more broadly induced by AT1 injury, indicating that other AT2 cells may have similar potential ('alveolar repair focus', Fig. 5b) that can be activated by dying AT1 cells. Classical pneumonectomy experiments support the idea that most AT2 cells possess latent regenerative capacity[31], and AT2 ablation induces self-duplication of surviving AT2 cells[9] ('AT2 replacement focus'). Our data do not exclude alternative sources of new alveolar cells, especially following severe injuries[14] that deplete AT2 cells regionally.

If many or all AT2 cells can serve as stem cells, why does only a minority execute this function for maintenance, producing large monoclonal foci? Perhaps alveolar turnover is coupled with a stem cell hierarchy whereby an initially activated AT2 cell suppresses nearby AT2 cells and becomes the dominant stem cell. It is also unclear why renewal foci slowly enlarge over time, because there is no obvious recurrent injury. Perhaps foci are programmed anatomical domains of alveolar renewal, with new cells moving out from a specialized niche akin to intestinal crypts, albeit with much slower turnover. Whatever the explanation, there are clearly regional influences as both renewal and tumour foci are preferentially located in perivascular and peripheral lung domains, possible sources of stem cell signals[32].

EGFR signalling selectively stimulates proliferation of AT2 cells *in vitro*, and oncogenic *Kras(G12D)* permanently and selectively activates proliferation *in vivo*, efficiently transforming AT2 cells into rapidly growing monoclonal tumours ('lung tumour focus', Fig. 5b). By virtue of their large numbers, class susceptibility and robust adenomatous response to *Kras(G12D)*, AT2 cells may constitute the major cell type responsible for human lung adenocarcinoma, making them among the most dangerous cells in the body.

We propose that the oncogenic potential of AT2 cells is a direct consequence of their stem cell function: AT2 cells are poised to function as alveolar stem cells and EGFR/KRAS signalling regulates the self-renewal part of the stem cell program (and the related process of self-duplication); another signalling pathway (Supplementary Table 2) must control AT2 reprogramming to AT1 fate (Fig. 5a). This model predicts that dying AT1 cells secrete an EGF that initiates self-renewal, plus another signal for fate reprogramming. Dying AT2 cells presumably produce only the former. It will be important to identify these signals and the events they control. This could suggest new strategies for early

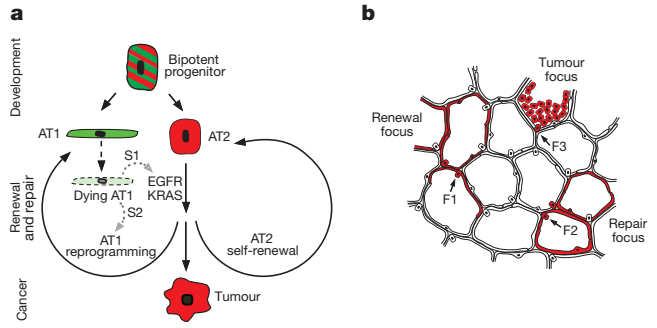


Figure 5 | Model of alveolar progenitors and stem cells in development, maintenance, and cancer. **a**, Bipotent progenitors expressing some AT1 (green) and AT2 (red) markers differentiate into AT1 or AT2 cells. Mature AT2 cells function as stem cells intermittently activated for alveolar renewal and repair. Dying AT1 cells are proposed to produce a signal (S1) transduced by EGFR-KRAS that activates division of a nearby AT2 cell (self-renewal); another signal (S2) reprograms a daughter into an AT1 cell. Activating mutations of *Egfr* or *Kras* in AT2 cells drive constitutive self-duplication, forming tumour of AT2-like cells. **b**, Rare AT2 cells function as stem cells, giving rise to clonal renewal foci (red, left) that slowly enlarge, with persistence of founder AT2 cell (F1). With injury, additional AT2 cells (F2) are recruited to generate repair foci (red, right). Activating *Kras* mutation in AT2 cells (F3) initiates tumour focus (red, top).

detection and treatment of tumours and for replenishing diseased alveoli.

METHODS SUMMARY

Cre recombinase was used to activate Rosa26-reporters[33] and a $Kras^{LSL\text{-}G12D}$ knock-in[34]. For immunostaining, lungs were agarose-inflated, then fixed in paraformaldehyde (PFA) and cryo-embedded or vibratome-sliced and fixed in methanol:dimethylsulphoxide or PFA. Secondary antibodies were Alexa Fluor-conjugated, or horseradish peroxidase-conjugated with tyramide amplification. Specimens were imaged by confocal (slices) or epifluorescence (cryosections) microscopy. For cell purification, lungs were dissociated then sorted by fluorescent markers or surface antigens using FACS or MACS. AT2 cells were cultured on Matrigel or on glass with serum. Expression profiling used Affymetrix platform.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 December 2012; accepted 3 December 2013.
Published online 5 February 2014.

1. Bertalanffy, F. D. & Leblond, C. P. Structure of respiratory tissue. *Lancet* **266,** 1365–1368 (1955).
2. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2013. *CA Cancer J. Clin.* **63,** 11–30 (2013).
3. Rock, J. R. & Hogan, B. L. Epithelial progenitor cells in lung development, maintenance, repair, and disease. *Annu. Rev. Cell Dev. Biol.* **27,** 493–512 (2011).
4. Chapman, H. A. *et al.* Integrin α6β4 identifies an adult distal lung epithelial population with regenerative potential in mice. *J. Clin. Invest.* **121,** 2855–2862 (2011).
5. Adamson, I. Y. & Bowden, D. H. Derivation of type 1 epithelium from type 2 cells in the developing rat lung. *Lab. Invest.* **32,** 736–745 (1975).
6. Spencer, H. & Shorter, R. G. Cell turnover in pulmonary tissues. *Nature* **194,** 880 (1962).
7. Evans, M. J. & Bils, R. F. Identification of cells labeled with tritiated thymidine in the pulmonary alveolar walls of the mouse. *Am. Rev. Respir. Dis.* **100,** 372–378 (1969).
8. Rock, J. R. *et al.* Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proc. Natl Acad. Sci. USA* **108,** E1475–E1483 (2011).
9. Barkauskas, C. E. *et al.* Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123,** 3025–3036 (2013).
10. Evans, M. J., Cabral, L. J., Stephens, R. J. & Freeman, G. Renewal of alveolar epithelium in the rat following exposure to NO_2. *Am. J. Pathol.* **70,** 175–198 (1973).
11. Adamson, I. Y. & Bowden, D. H. The type 2 cell as progenitor of alveolar epithelial regeneration. A cytodynamic study in mice after exposure to oxygen. *Lab. Invest.* **30,** 35–42 (1974).
12. Kim, C. F. *et al.* Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* **121,** 823–835 (2005).
13. McQualter, J. L., Yuen, K., Williams, B. & Bertoncello, I. Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. *Proc. Natl Acad. Sci. USA* **107,** 1414–1419 (2010).
14. Kumar, P. A. *et al.* Distal airway stem cells yield alveoli *in vitro* and during lung regeneration following H1N1 influenza infection. *Cell* **147,** 525–538 (2011).
15. Kajstura, J. *et al.* Evidence for human lung stem cells. *N. Engl. J. Med.* **364,** 1795–1806 (2011).
16. Rawlins, E. L. *et al.* The role of $Scgb1a1^+$ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. *Cell Stem Cell* **4,** 525–534 (2009).
17. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479,** 189–193 (2011).
18. Burri, P. H. & Moschopulos, M. Structural analysis of fetal rat lung development. *Anat. Rec.* **234,** 399–418 (1992).
19. Buckingham, S., McNary, W. F., Jr & Sommers, S. C. Pulmonary alveolar cell inclusions: their development in the rat. *Science* **145,** 1192–1193 (1964).
20. Williams, M. C. & Dobbs, L. G. Expression of cell-specific markers for alveolar epithelium in fetal rat lung. *Am. J. Respir. Cell Mol. Biol.* **2,** 533–542 (1990).
21. Hughes, G. M. Ultrastructure of the lung of *Neoceratodus* and *Lepidosiren* in relation to the lung of other vertebrates. *Folia Morphol. (Praha)* **21,** 155–161 (1973).
22. Miller, L. A., Wert, S. E. & Whitsett, J. A. Immunolocalization of sonic hedgehog (Shh) in developing mouse lung. *J. Histochem. Cytochem.* **49,** 1593–1603 (2001).
23. Messier, B. & Leblond, C. P. Cell proliferation and migration as revealed by radioautography after injection of thymidine-H^3 into male rats and mice. *Am. J. Anat.* **106,** 247–285 (1960).
24. Bowden, D. H., Adamson, I. Y. & Wyatt, J. P. Reaction of the lung cells to a high concentration of oxygen. *Arch. Pathol.* **86,** 671–675 (1968).
25. Herbst, R. S., Heymach, J. V. & Lippman, S. M. Lung cancer. *N. Engl. J. Med.* **359,** 1367–1380 (2008).
26. Sutherland, K. D. & Berns, A. Cell of origin of lung cancer. *Mol. Oncol.* **4,** 397–403 (2010).
27. Xu, X. *et al.* Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. *Proc. Natl Acad. Sci. USA* **109,** 4910–4915 (2012).
28. Lin, C. *et al.* Alveolar type II cells possess the capability of initiating lung tumor development. *PLoS ONE* **7,** e53817 (2012).
29. Nikitin, A. Y. *et al.* Classification of proliferative pulmonary lesions of the mouse: recommendations of the mouse models of human cancers consortium. *Cancer Res.* **64,** 2307–2316 (2004).
30. Takanami, I., Takeuchi, K. & Kodaira, S. Tumor-associated macrophage infiltration in pulmonary adenocarcinoma: association with angiogenesis and poor prognosis. *Oncology* **57,** 138–142 (1999).
31. Brody, J. S., Burki, R. & Kaplan, N. Deoxyribonucleic acid synthesis in lung cells during compensatory lung growth after pneumonectomy. *Am. Rev. Respir. Dis.* **117,** 307–316 (1978).
32. Ding, B. S. *et al.* Endothelial-derived angiocrine signals induce and sustain regenerative lung alveolarization. *Cell* **147,** 539–553 (2011).
33. Kretzschmar, K. & Watt, F. M. Lineage tracing. *Cell* **148,** 33–45 (2012).
34. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15,** 3243–3248 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Andalon for technical assistance; H. Chapman (SftpC–Cre–ER-rtTA), B. Hogan (CCSP–Cre–ER), H. Ueno and I. Weissman (Rainbow), H. Clevers (Confetti), L. Luo (mTmG), and J. Sage ($Kras^{LSL\text{-}G12D}$) for strains; B. Stripp for goat anti-CCSP antibody; F. H. Espinoza for annotated gene lists; R. Metzger, H. Chapman, and members of the Krasnow laboratory for discussions and comments on the manuscript; and Maria Petersen for help preparing figures and the manuscript.

Author Contributions T.J.D. conducted the experiments except the gene expression profiling and AT2 cell cultures, which were done by D.G.B.; T.J.D., D.G.B. and M.A.K. conceived the experiments, analysed the data and wrote the manuscript. This work was supported by a Parker B. Francis Foundation Fellowship and NIH 5K08HL084095 Award (T.J.D.), NIH T32HD007249 (D.G.B.), and an NHLBI U01HL099995 Progenitor Cell Biology Consortium grant (M.A.K.). M.A.K. is an investigator of the Howard Hughes Medical Institute.

Author Information Microarray datasets were deposited at Gene Expression Omnibus (accession code GSE49346) and GEXC (https://gexc.stanford.edu/population/detail/998 and https://gexc.stanford.edu/population/detail/999). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.A.K. (krasnow@stanford.edu) or T.J.D. (tdesai@stanford.edu).

METHODS

Mouse strains. CD1 or B6 (Charles River Laboratories) were the wild type strains. The Cre recombinase system was used to manipulate gene and marker expression *in vivo* in specific mouse lung cell types and lineages[33] in mixed strain backgrounds, using a knock-in allele that expresses Cre or a tamoxifen-inducible Cre-oestrogen receptor (ER or ERT2) fusion crossed to a Cre-dependent target gene. For Cre-ER lines, tamoxifen (Sigma) was dissolved in corn oil by sonication, stored in 100 µl aliquots at −80 °C, and 1 mg (Rosa26–Cre–ERT2, CCSP–Cre–ER), 2 or 4 mg (Shh–Cre–ERT2) or 1, 2, or 3 mg (SftpC–Cre–ERT2-rtTA) injected intraperitoneally at specified ages to induce recombination in cells throughout the transgene expression domain. The Cre expression strains used were: Shh–Cre[35], which recombines throughout the endoderm-derived lung epithelium from the beginning of lung development; Shh–Cre–ERT2[35] expressed in early lung endoderm then restricts to distal tip epithelium during branching morphogenesis[22]; LysM–Cre[36] expressed in mature AT2 cells, macrophages, and neutrophils; SftpC–Cre–ERT2–rtTA[4] expressed in the bipotent alveolar progenitor and developing and mature AT2 cells, as well as in rare bronchiolar cells; CCSP–Cre–ER[16] expressed in Clara cells and rare SftpC-expressing alveolar cells, and Rosa26–Cre–ERT2[37] expressed throughout the lung and rest of the body. The Cre-responsive target genes used were: R26–EYFP[38], a ubiquitously expressed transgene that expresses enhanced yellow fluorescent protein (EYFP) following Cre recombination; R26–mTmG[39], which expresses membrane-targeted tdTomato before recombination and membrane-targeted enhanced green fluorescent protein (EGFP) after recombination; R26–Confetti[40], which expresses either red fluorescent protein (RFP), EYFP, nuclear GFP, or membrane cyan fluorescent protein (CFP) after recombination; R26–tdTomato[41], which expresses tdTomato after recombination; and R26–Rainbow[42], which expresses GFP before recombination and either CFP, mOrange, or mCherry after recombination. *Kras*$^{LSL\text{-}G12D}$ is a knock-in at the *Kras* locus in which the wild-type *Kras* coding sequence is replaced by a lox-STOP-lox-*Kras(G12D)* allele[34], which following Cre recombination results in endogenously regulated expression of a constitutively active KRAS protein. LysM–EGFP[43] is an EGFP knock-in at the *Lyz2* locus expressed by mature AT2 cells, macrophages, and neutrophils. Genotyping was performed by PCR of DNA extracted from ear clips using the Red-Xtract Kit (Sigma) and published primer sets. Mice were housed in filtered cages and all experiments were performed in accordance with approved IACUC (Institutional Animal Care and Use Committee) protocols.

Hyperoxic lung injury. 8-week-old female mice were housed in a sealed Plexiglas chamber with oxygen delivered by continuous flow to maintain levels at 88% for 5 days, after which the animals were returned to room air to recover for three weeks before analysing. Female siblings maintained under room air were used as controls. AT1 cell renewal was quantified as described below by blinded scoring of the number of alveoli in which AT1 cells expressed the AT2 cell lineage label in three randomly selected lung volumes (100–300-µm thick lung slices in randomly selected microscope fields using a 20× objective) from two animals per group.

Lung collection, fixation and processing. Embryos were staged by vaginal plugging of the mother, with noon on the day of appearance of the plug taken as embryonic day (E) 0.5. Individual embryos were also staged by fetal crown-rump length at time of euthanasia. Lungs were removed en bloc and a scalpel used to excise the tip of the accessory lobe, after which lungs and the tip were fixed in methanol: dimethylsulphoxide (DMSO) (4:1) with rocking overnight at 4 °C, bleached in methanol:DMSO:30% H_2O_2 (4:1:1) for 5 h at room temperature, then dehydrated and stored in methanol at −20 °C until staining.

Postnatal mice were euthanized by carbon dioxide inhalation, the abdominal aorta severed, and sternotomy performed. Phosphate buffered saline (PBS; Ca^{2+}- and Mg^{2+}-free, pH 7.4) was gently perfused into the right ventricle by manual pressure using a syringe with a 21 gauge needle to clear the pulmonary vasculature. For mice up to 7 days old, lungs were removed en bloc, the accessory lobe tip excised, lobes separated and trimmed to create a flat base, glued to a tissue disk with cyanoacrylate, then submerged in ice-cold PBS and sliced at a thickness of 50 to 300 µm using a vibrating microtome (Leica). Lung slices were fixed in methanol: DMSO as above. For direct imaging without immunostaining, slices were fixed in paraformaldehyde (PFA; 2% in PBS) overnight at 4 °C then stored in Vectashield (Vector) at 4 °C.

For mice 8 days or older, lungs were collected and processed as above except that following clearance of the pulmonary vasculature the ventral trachea was incised and cannulated with a blunt needle that was secured in place by tying a suture around the trachea. Lungs were then gently inflated to full capacity with molten low-melting-point agarose (Sigma, 2% in PBS) after which the cannula was withdrawn and the suture tightened to prevent leakage. Ice-cold PBS was dripped into the thorax to solidify the agarose, then the inflated lungs were removed en bloc and processed as above.

Lungs from mice carrying the Rainbow transgene were cleared and inflated as above, immersion fixed in PFA (4% in PBS) overnight at 4 °C, cryoprotected in 30% sucrose overnight at 4 °C, then submerged in OCT (Tissue Tek) in an embedding mould, frozen on dry ice, and stored at −80 °C. Sections of 10 and 30 µm obtained using a cryostat (Leica CM3050S) were collected on glass slides, mounted in Prolong Gold (Invitrogen) and stored at 4 °C after curing at room temperature overnight. Lungs for tumour histology were collected without perfusion, formalin-fixed, paraffin-embedded, then sections obtained using a microtome were collected on glass slides, stained with haematoxylin and eosin and mounted.

Lung fluorescent immunostaining and microscopy. Indirect immunohistochemistry was performed as described[44] with the following modifications. For signal amplification of whole-mount lung tips and pieces, a horseradish peroxidase-conjugated secondary antibody was used (Jackson ImmunoResearch) and detected by tyramide signal amplification (TSA, Plus Fluorescein System, PerkinElmer NEL741001KT). Primary antibodies raised in mouse were detected using the Mouse-on-Mouse Kit (Vector). For lung slices, secondary antibodies conjugated to an Alexa Fluor dye (A488, A555, or A633; Invitrogen) were used. DAPI (5 ng ml^{-1}) was included during incubation with secondary antibodies. Staining with biotin-conjugated lectins was performed as described for primary antibodies then detected with streptavidin conjugated to an Alexa Fluor dye (Invitrogen).

Primary antibodies against the following antigens (used at 1:200 dilution unless otherwise noted) were: pro-SftpC (rabbit, Chemicon AB3786), GFP (chicken, Abcam AB13970), F4/80 (rat, Serotec MCA497G; clone CI:A3-1), CGRP (calcitonin gene-related peptide) (guinea pig, Europroxima 2263BGP470-1), acetylated tubulin (mouse, Sigma T6793; clone 6-11B-1), lysozyme (rabbit, Thermo Scientific RB372A), RAGE (rat, R&D MAB1179), E-cadherin (rat, Life Technologies 131900; clone ECCD-2), CCSP (rabbit, EMD Millipore 07-623; goat, provided by B. Stripp), Pdpn (hamster, DSHB 8.1.1; 1:20), Aqp5 (rabbit, Calbiochem 178615), Muc1 (hamster, Thermo Scientific HM1630, clone MH1 (CT2); rabbit, Novus NBP1-02974), cathepsin H (goat, R&D AF1013), Ki67 (rat, DAKO M7249, clone TEC-3), Nkx2.1 (rabbit, Millipore 07-601; mouse, Thermo Scientific MS-699-P, clone 8G7G3/1), SftpB (rabbit, Chemicon AB3780), SftpD (rabbit, Chemicon AB3434), Abca3 (mouse, Seven Hills 13-H2-57), Lamp-1 (rat, DSHB 1D4B) and Lamp-2 (rat, DSHB GL2A7). Biotin-conjugated lectins used were: *Lycopersicon esculentum* (Vector B-1175) and *Ricinus communis* I (Vector B-1085).

After staining, lung slices were immersed in Vectashield (Vector) and either placed in a glass depression slide with a coverslip gently lowered on top and its edges secured with scotch tape, or into a chambered coverslip (Falcon). Images were acquired using a Leica Sp2 upright or Sp5 inverted laser scanning confocal microscope with LCS software. Cryosectioned Rainbow lungs were imaged using a Zeiss Axiophot upright fluorescence microscope with AxioVision 4.2 software. Intact lungs were visualized with a Leica MZ16FA fluorescent stereomicroscope and images captured with a Retiga 2000R (Q Imaging) camera using Image-Pro Plus (Media Cybernetics) software. Haematoxylin and eosin stained sections were imaged using a compound light microscope equipped with a digital camera. Images were processed (pseudocoloured, levels and contrast adjusted, overlaid) using Adobe Photoshop.

Electron microscopy and ultrastructural analysis. E18.3 accessory lobe tips were collected as described above then fixed in 3% glutaraldehyde/0.1 M cacodylate (pH 7.4) overnight at 4 °C, post-fixed in 1% OsO_4 for 1 h at 4 °C, washed, stained in 1% uranyl acetate overnight at 4 °C, dehydrated to 100% ethanol, infiltrated with Embed 812 resin (EMS), then oriented in casts and polymerized at 65 °C for 24 h. Blocks were sectioned using an ultramicrotome (Leica Ultracut) until fields containing partially sacculated airway tips were identified. Specimens were visualized with a transmission electron microscope (Jeol TEM1230) and images captured with a CCD camera then processed using Adobe Photoshop. Individual epithelial cells were scored for the presence or absence of glycogen vacuoles, lamellar bodies and cuboidal versus attenuated morphology.

Assigning alveolar cell marker classes. Each of the alveolar antigens and lectins was identified by systematically searching the published literature using PubMed. We then purchased all antibodies and lectins from our list that were commercially available and tested their expression in the adult lung against the canonical markers of AT1 (T1α/Pdpn) and AT2 (pro-SftpC) cells. In cases where we could not perform double-labelling with a canonical marker because the antisera were raised in the same species, we compared expression against other markers we had already validated that were raised in a different species. For AT1 markers that co-localized with Pdpn, we further attempted to confirm these were not also expressed by endothelial or AT2 cells by directly comparing their expression against anti-GFP in lung slices from Shh–Cre > mTmG mice, which provides pan-epithelial membrane marking. In total, we tested 77 antibodies raised against 48 antigens and seven lectins for robustness of signal, then for AT1 or AT2 selective marking. We then analysed the expression patterns of our panel of validated AT1 and AT2 cell markers in alveolar progenitors at the E17 accessory lobe tip, which we selected because its distinctive shape provides reliable landmarks for identifying the same airway position in different lung specimens. Only two expression profiles were observed, absent or ubiquitous. Of those expressed ubiquitously, each progenitor

cell would subsequently selectively downregulate either the AT1 or the AT2 subset, distinguishing pre- or early AT2 or AT1 cells, respectively. We designated these marker classes $P1^E$, $P2^E$, $P1^L$, and $P2^L$, based on the specificity and relative timing of their restriction (initial expression in Progenitors (P) with restriction to AT1 (1) or AT2 (2) cells Early (E) or Late (L) in the differentiation process). Next, the AT1 and AT2 markers that were absent in progenitors at E17 initiated selective expression in nascent AT1 or AT2 cells at E18, so we designated these two classes $A1^L$ and $A2^L$ (initial expression in AT1 (1) or AT2 (2) cells, Late (L) onset of expression). Classes were assigned based on analysis of five terminal airways (100 cells) for each developmental stage indicated.

Quantification of AT1 cell renewal. Alveolar microscope fields (using a 20× magnification objective) were randomly selected and scored for the presence and extent of AT1 cell labelling by GFP. An alveolar unit with at least half of the epithelial AT1 surface marked by the lineage tag was considered labelled, and contiguous extension of the tag into adjacent alveoli was considered part of the same focus. All scoring was performed with blinding to the experimental condition (for hyperoxia versus room air) and age (for clonal analysis). Percent of alveoli labelled was determined by counting the number of positive alveoli in a given volume of tissue then dividing by the total number of alveoli scored, with the latter estimated by assuming a cuboidal shape and average diameter of 75 μm per alveolus.

Bipotent progenitor and AT2 cell isolation. To quantify alveolar progenitor populations and isolate bipotent progenitors for expression profiling, E18 lungs from a single litter (~6 embryos) were removed en bloc without perfusion and a scalpel used to collect peripheral lobe edges enriched for alveolar cells. Lung tissue was rinsed briefly in PBS, minced with a razor blade into 1 mm^3 fragments, suspended in 5 ml of digestion buffer consisting of Elastase (3 U ml^{-1}; Worthington Biochemical Corporation) and DNase I (0.33 U ml^{-1}; Roche) in DMEM/F12, incubated with frequent agitation at 37 °C for 45 min, triturated briefly with a 5 ml pipette, then an equal volume of DMEM/F12 supplemented with 10% FBS and penicillin-streptomycin (1 U ml^{-1}, Thermo Scientific) was added and the cell suspension passed through a 100-μm mesh filter (Fisher) to remove residual tissue fragments, then centrifuged at 400*g* for 10 min. The pelleted cells were resuspended and incubated in red blood cell lysis buffer (BD Biosciences) for 2 min, passed through a 40-μm mesh filter (Fisher), centrifuged at 400*g* for 10 min, then resuspended in sorting buffer (PBS supplemented with 0.05% BSA and 2 mM EDTA). Antibodies against Muc1 (hamster, Thermo Scientific HM-1630-P1ABX; 1:100) conjugated using Alexa Fluor A488 Antibody Labelling Kit (Life Technologies A-20181), phycoerythrin-conjugated Pdpn (hamster, eBiosciences 12-5381-80, clone eBio8.1.1; 1:100), and Sytox Blue cell viability stain (Invitrogen) were added and the cells incubated for 15 min, after which cell populations were quantified by flow cytometry (Aria II, BD Biosciences) and viable $Muc1^+/Pdpn^+$ progenitors sorted directly into cell lysis buffer for RNA isolation.

To isolate LysM-lineage AT2 cells for expression profiling, adult (PN70) LysM-Cre > tdTomato mice were euthanized and pulmonary vasculature perfused clear (described above). The trachea was cannulated with a 20 gauge needle and lungs gently inflated with digestion buffer, then removed en bloc and lobes individually rinsed briefly in PBS, minced, suspended in digestion buffer, and a cell suspension was prepared (described above). A phycoerythrin-conjugated antibody against EpCAM (rat, eBiosciences 12-5791-81, clone G8.8; 1:100) was added and after incubation, $EpCAM^+/Tomato^+$ AT2 cells were sorted by FACS directly into cell lysis buffer for RNA isolation.

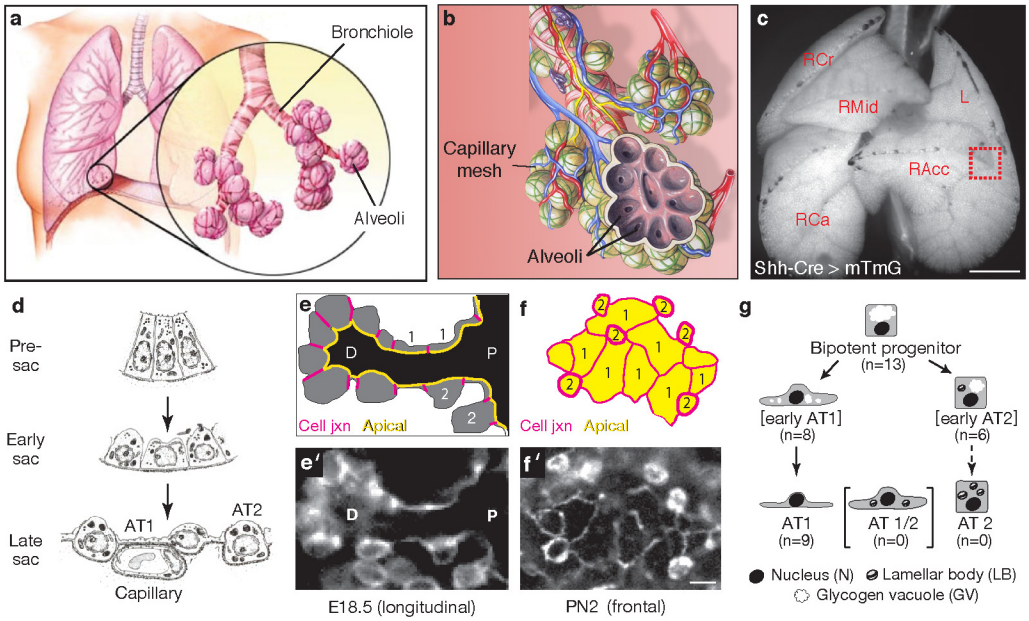
To isolate AT2 cells for culturing, a cell suspension (generated as above) from adult (PN50) B6 mouse lung was depleted of $CD45^+$ cells using a magnetic-activated cell sorter (MACS, Miltenyi Biotec) and CD45 microbeads (rat monoclonal, 30F11.1, Miltenyi Biotech 130-052-301) according to manufacturer instructions, incubated with a biotinylated antibody against EpCAM (rat, eBiosciences 13-5791-81; clone G8.8), then enriched for $EpCAM^+$ cells using streptavidin-coated microbeads (Miltenyi Biotech 130-048-102) as described[45,46]. Immunostaining for SftpC (rabbit, Chemicon AB3786; 1:100) was used to determine purity, with 90% of cells positive for this AT2 cell type marker.

To determine if Shh^+ embryonic progenitors were present in adult lung, we administered 4 mg tamoxifen to an adult (PN154) Shh–Cre–ER > mTmG mouse then collected and examined the lungs six days later for cells expressing the lineage mark, which were not observed. We also isolated distal epithelial cells from lungs of adult (PN70) B6 mice ($n = 3$) as above, performed FACS to select $EpCAM^+$ cells, then quantified the fraction of $Muc1^+/Pdpn^+$ cells.

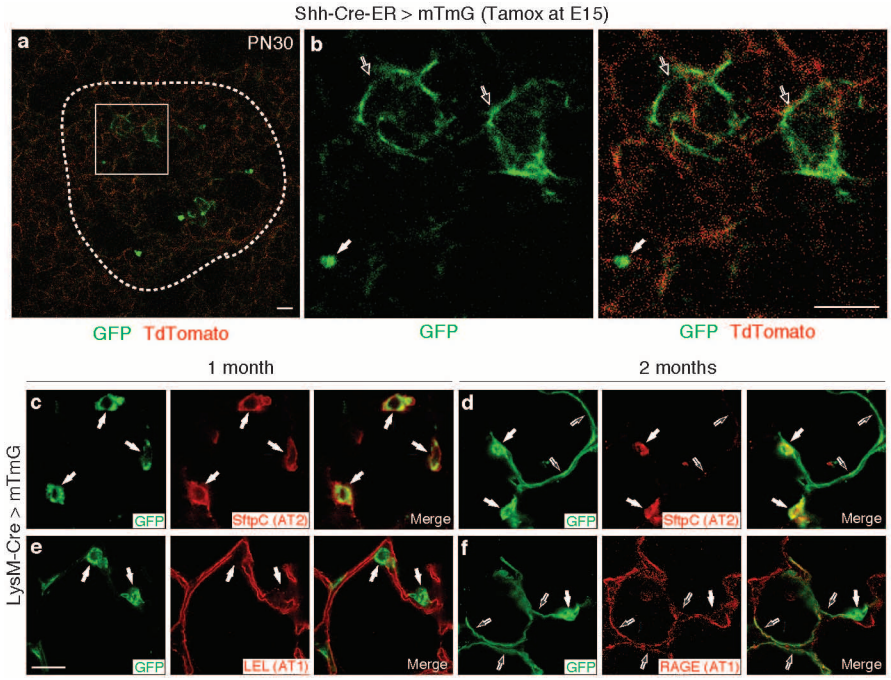
AT2 cell culture. Purified AT2 cells (described above) were seeded at a density of ~10,000 cells per well in #1 8-well coverglass chambers (Labtek) containing DMEM/F12 media supplemented with 10% fetal bovine serum. Cells were cultured at 37°C for 16 days (with media replacement every 2 days), during which they acquired an AT1-like cell phenotype[47]. Alternatively, AT2 cells were cultured on a layer of Matrigel (BD) in serum-free media[48] (DMEM/F12 supplemented with L-glutamine, non-essential amino acids, and penicillin-streptomycin) at 37 °C for 72 h, after which these conditions were either maintained or one of the EGF ligands TGFα (4 μM, R&D Systems 239-A-100), HB-EGF (4 μM, R&D Systems 259-HE-050), NRG1 (4 μM, R&D Systems 5898-NR-050) or a function blocking anti-EGFR antibody (rat, Abcam AB231, clone ICR10; 2.5 μg ml^{-1}) was added and the cells cultured for an additional 24 h, then fixed (Cytofix/Cytoperm, BD) and immunostained for SftpC (rabbit, Chemicon AB3786; 1:100) and Ki67 (rat, eBiosciences 41-5698-80, clone SolA15; 1:100) to identify proliferating AT2 cells. Stained samples were imaged using a Zeiss LSM 780 confocal microscope. Total and Ki67 positive nuclei were quantified (500 cells per well) for each condition in a blinded manner.

Alveolar progenitor and AT2 cell gene expression profiling. RNA was extracted from E18 bipotent progenitors or adult AT2 cells (isolated as described above, with yields of approximately 35,000 cells) using the RNeasy Micro Kit (Qiagen), cDNA synthesized using a 2 cycle cDNA synthesis kit (Affymetrix) and analysed on a Mouse Genome 430 2.0 Array (Affymetrix). RNA quality and quantity assessment, cDNA synthesis, probe preparation, labelling, hybridization and image scan were performed by the Protein and Nucleic Acid (PAN) facility at Stanford School of Medicine. Intensity values acquired from the Expression Console (Affymetrix) were analysed using Gene Expression Commons[49] (https://gexc.stanford.edu/).

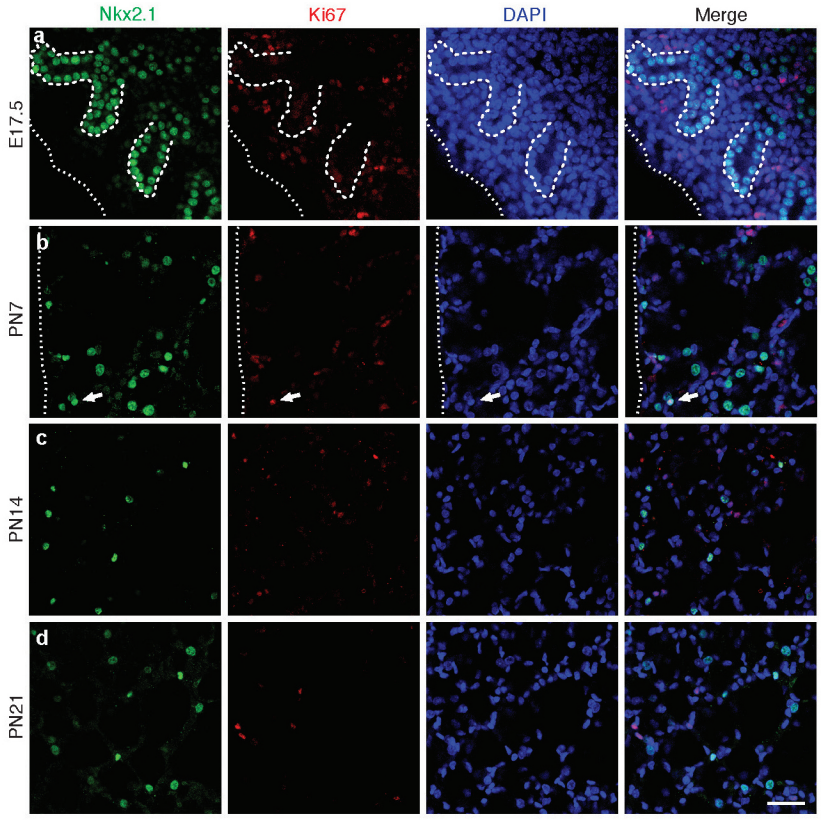
35. Harfe, B. D. *et al.* Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* **118,** 517–528 (2004).
36. Clausen, B. E., Burkhardt, C., Reith, W., Renkawitz, R. & Forster, I. Conditional gene targeting in macrophages and granulocytes using LysMcre mice. *Transgenic Res.* **8,** 265–277 (1999).
37. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445,** 661–665 (2007).
38. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of *EYFP* and *ECFP* into the *ROSA26* locus. *BMC Dev. Biol.* **1,** 4 (2001).
39. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45,** 593–605 (2007).
40. Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143,** 134–144 (2010).
41. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13,** 133–140 (2010).
42. Rinkevich, Y., Lindau, P., Ueno, H., Longaker, M. T. & Weissman, I. L. Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. *Nature* **476,** 409–413 (2011).
43. Faust, N., Varas, F., Kelly, L. M., Heck, S. & Graf, T. Insertion of enhanced green fluorescent protein into the lysozyme gene creates mice with green fluorescent granulocytes and macrophages. *Blood* **96,** 719–726 (2000).
44. Metzger, R. J., Klein, O. D., Martin, G. R. & Krasnow, M. A. The branching programme of mouse lung development. *Nature* **453,** 745–750 (2008).
45. Messier, E. M., Mason, R. J. & Kosmider, B. Efficient and rapid isolation and purification of mouse alveolar type II epithelial cells. *Exp. Lung Res.* **38,** 363–373 (2012).
46. Fujino, N. *et al.* A novel method for isolating individual cellular components from the adult human distal lung. *Am. J. Respir. Cell Mol. Biol.* **46,** 422–430 (2012).
47. Dobbs, L. G., Williams, M. C. & Brandt, A. E. Changes in biochemical characteristics and pattern of lectin binding of alveolar type II cells with time in culture. *Biochim. Biophys. Acta* **846,** 155–166 (1985).
48. Sugahara, K., Mason, R. J. & Shannon, J. M. Effects of soluble factors and extracellular matrix on DNA synthesis and surfactant gene expression in primary cultures of rat alveolar type II cells. *Cell Tissue Res.* **291,** 295–303 (1998).
49. Seita, J. *et al.* Gene Expression Commons: an open platform for absolute gene expression profiling. *PLoS ONE* **7,** e40321 (2012).
50. Giangreco, A. *et al.* Stem cells are dispensable for lung homeostasis but restore airways after injury. *Proc. Natl Acad. Sci. USA* **106,** 9286–9291 (2009).
51. Lynch, P. J. & Jaffe, C. C. Bronchial anatomy detail of alveoli and lung circulation. http://commons.wikimedia.org/wiki/File%3ABronchial_anatomy.jpg (2006).
52. Burri, P. H. in *Comprehensive Physiology* 1–46 http://dx.doi.org/10.1002/cphy.cp030101 (Wiley, 2011).



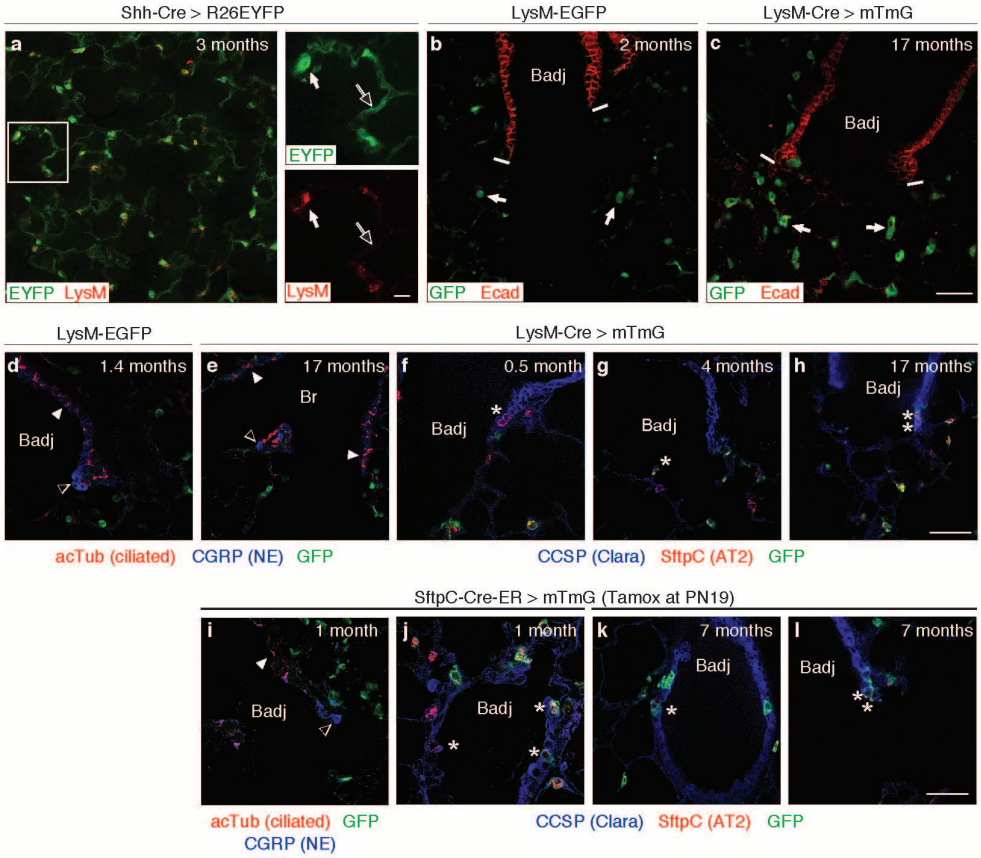
Extended Data Figure 1 | Mature and developing structure of the lung and alveoli. **a**, Mature lung showing close up (inset) of transition from bronchial tree to alveoli (from http://www.mayoclinic.com/, with permission). **b**, Alveoli are surrounded by a dense capillary network in which de-oxygenated blood (blue) from pulmonary arteries is oxygenated (red) and then returned to the heart through pulmonary veins (modified from ref. 51). **c**, E16.5 Shh–Cre > mTmG mouse lung in which GFP is expressed throughout the epithelium. Lobes are labelled (RAcc, right accessory; RCr, right cranial; RMid, right middle; RCa, right caudal; L, left) and boxed region shows the tip of the accessory lobe that was used for developmental analyses. By E16.5, the bronchial tree has formed and 1 day later sacculation begins as flat (squamous) AT1 and cuboidal AT2 cells mature to generate a functional gas exchange interface. Saccules subsequently undergo subdivision ('secondary septation') into mature alveoli that provide an extensive gas-exchange surface. Scale bar, 1 mm; **d**, Schematic of cell morphogenesis during sacculation (sac) (from ref. 52). Progenitors form flat AT1 cells adjacent to capillaries and AT2 cells specialized to secrete surfactant. **e**, **f**, **e′**, **f′**, Schematics (**e**, **f**) and images (**e′**, **f′**) of E-cadherin-stained accessory lobe tips at E18.5 and postnatal day 2 (PN2) showing longitudinal (**e**) and frontal (**f**) views of maturing AT1 (1) and AT2 (2) cells. P, proximal; D, distal; red, cell junctions (jxn); yellow, apical surfaces. Note lack of AT1 cells distally in sacculating airway. Scale bar, 10 μm (**e′**, **f′**). **g**, Quantification of ultrastructural classification of cell types in sacculating airways in E18.3 lungs (see Fig. 1r–u). Values shown are the numbers of each progenitor and cell type observed with the indicated ultrastructural features. No cells ($n = 36$) had features of an AT2> AT1 intermediate (AT1/2) or mature AT2 cell. E, embryonic day; PN, postnatal day.



Extended Data Figure 2 | Clonal analysis of alveolar progenitor cells and lineage marking and tracing alveolar type 2 (AT2) cells with LysM-Cre. **a**, **b**, Shh–Cre–ER > mTmG embryos were induced in utero with a limiting dose (2 mg) of tamoxifen (tamox) at E15 to pulse-label epithelial cells at the distal lung tips (alveolar progenitors) with GFP (0.2 labelled cells per embryonic lung lobe) shortly before the onset of differentiation then examined 34 days later at PN30. **a**, An isolated clone (dashed circle) expressing the GFP lineage tag (green) in a PN30 lung. **b**, Close up of boxed region showing several flat AT1 cells (open arrows) and a cuboidal AT2 cell (filled arrows) within the alveolar clone, indicating that the tagged progenitor was bipotent. Tagged cells are interspersed with unrecombined cells (tdTomato, red). E, embryonic day; PN, postnatal day; Scale bar, 50 μm. **c**–**f**, Close-ups of alveoli of 1 (**c**, **e**) and 2 month old (**d**, **f**) LysM-Cre > mTmG lungs stained for the AT2 lineage tag (GFP, green) and the AT2 (**c**, **d**) or AT1 (**e**, **f**) markers indicated. Note that at 1 month lineage marked cells (green) express the AT2 (**c**) but not the AT1 marker (**e**). At 2 months (**d**, **f**), the lineage mark is also observed in some flat AT1 cells. Filled arrows, AT2 cells; open arrows, AT1 cells; E, embryonic day; PN, postnatal day. Scale bar, 20 μm (**c**–**f**).



Extended Data Figure 3 | Proliferation analysis of bipotent progenitors and alveolar epithelial cells. **a**–**d**, Late gestational (E17.5, **a**) and early postnatal (PN7, 14, 21; **b**–**d**) lungs stained for Nkx2.1 (green) for epithelial and Ki67 (red) for actively cycling cells. Note essentially exclusive labelling, indicating minimal proliferation of bipotent progenitors (**a**) or AT1 and AT2 cells (**b**–**d**). Arrow, a rare proliferating AT2 cell. Dashes outline distal epithelial tips; dotted line indicates mesothelium. E, embryonic day; PN, postnatal day; Scale bar, 35 μm.

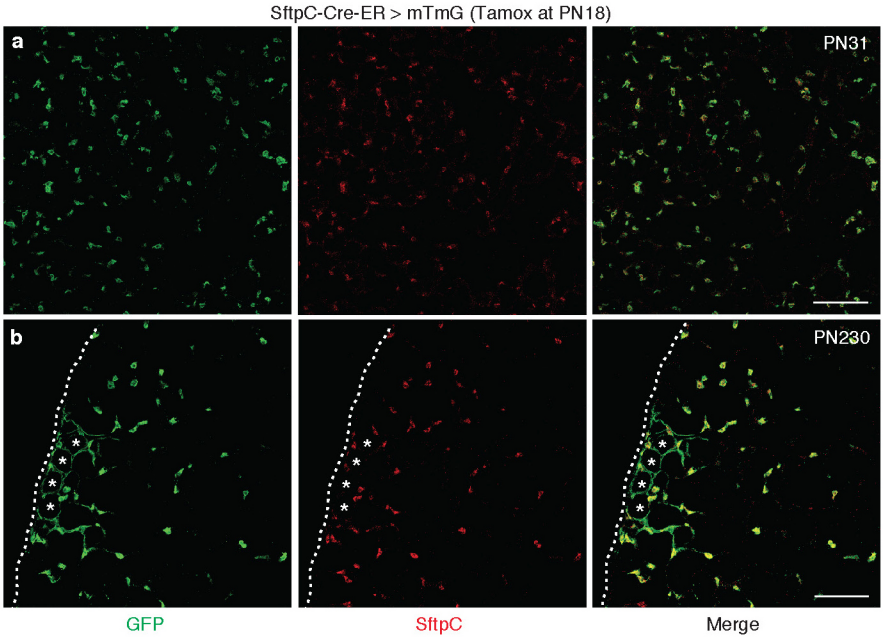


m

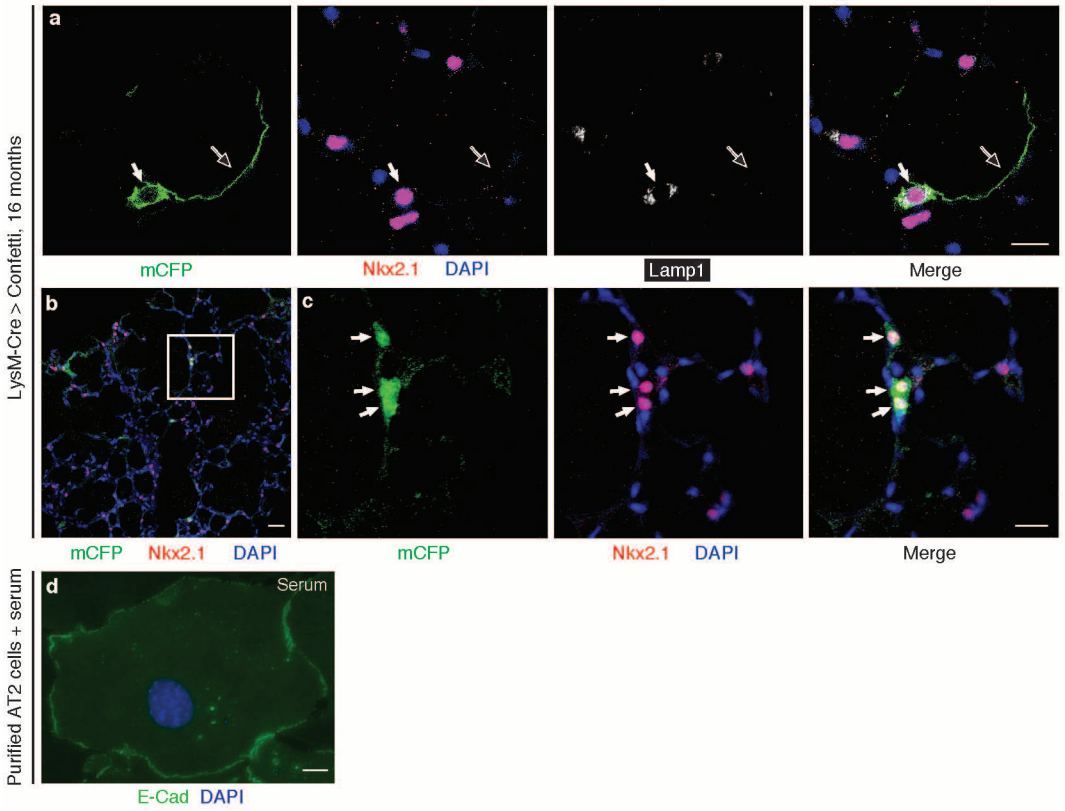
Genotype, Age	Cell type (cell marker)					
	Ciliated (acTub)	NE (CGRP)	AT2 (SftpC)	Clara (CCSP)	Double positive (CCSP/SftpC)	No. of Badj analyzed
LysM-EGFP, 9d	0/850	0/29		0/400	0/4	8
LysM-EGFP, 1.4 mo	0/350	0/35		0/500	1/7 (14%)	14
LysM-EGFP, 9.5 mo	0/550	0/86		0/400	0/3	9
LysM-Cre>mTmG, 0.5 mo	0/650	1/42 (2.3%)	150/650 (23%)	0/700	0/7	19
LysM-Cre>mTmG, 4.2 mo	0/500	0/24	363/455 (80%)	0/500	*5/10 (50%)	11
LysM-Cre>mTmG, 17 mo	0/450	0/31	156/236 (66%)	*40/900 (4.4%)	*17/43 (40%)	13
SftpC-CreER>mTmG, 1 mo (1 mg Tamox + 13d)	0/550	0/51	554/588 (94%)	0/500	*4/7 (57%)	11
SftpC-CreER>mTmG, 7 mo (1 mg Tamox + 192d)	0/900	1/47 (2.1%)	499/514 (97%)	*8/900 (0.9%)	*9/11 (82%)	22

* >90% of the marked cells were solitary and almost all others were doublets. The largest observed cell clusters were 3 ($CCSP^+$/$SftpC^+$) and 5 ($CCSP^+$) cells.

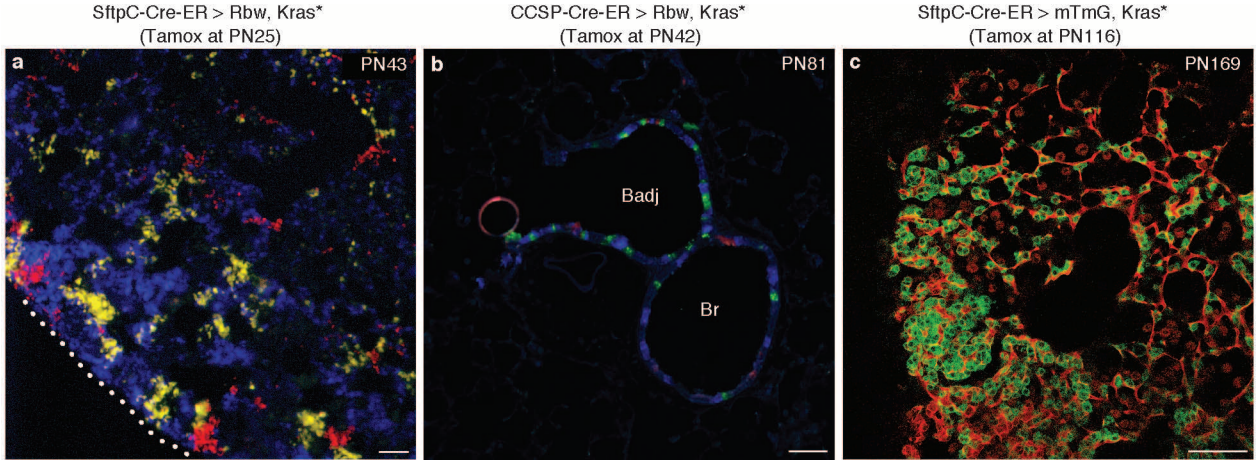
Extended Data Figure 4 | Quantification of cell type labelling and long-term lineage contribution of LysM-Cre and SftpC–Cre–ER marked cells. **a**, Alveolar region of a PN 3 month (mo) Shh–Cre> R26EYFP mouse lung co-stained for GFP (green, epithelial cytoplasm) and LysM (red). Inset shows close-up of boxed region. LysM is detected in cytoplasm of many AT2 cells (filled arrow) but not AT1 cells (open arrow). **b**, Bronchoalveolar lung region of a PN 2 mo LysM reporter mouse expressing GFP from the endogenous locus stained for E-cadherin (red) to mark airway epithelium and GFP (green) to mark LysM-expressing cells. Note AT2 (filled arrows) but not bronchiolar cells (dashes mark bronchoalveolar junction (Badj)) express the LysM reporter. **c**, PN 17 mo LysM-Cre > mTmG lung stained for E-cadherin (red) and the AT2 lineage marker (green). Note many marked AT2 cells (filled arrows) but absence of lineage-marked cells in the terminal bronchiole (dashes, Badj). **d**, **e**, Lungs from LysM-EGFP (**d**) and LysM-Cre > mTmG (**e**) mice of the indicated ages stained for ciliated (acetylated tubulin, acTub, red) and neuroendocrine (NE) cell (CGRP, blue) markers and GFP (green) show no co-expressing cells, indicating ciliated (filled arrowhead) and NE (open arrowhead) cell types do not express LysM and do not derive from AT2 cells. Br, bronchus. **f**–**h**, Lungs from LysM-Cre > mTmG mice of the indicated ages stained for the AT2 lineage tag (green), the AT2 cell marker SftpC (red), and the Clara cell marker CCSP (blue) show $CCSP^+$/$SftpC^+$ (double-positive) cells (*) at the Badj, some of which are tagged. Marked double-positive cells are solitary (**g**) or in doublets (**h**). **i**, Lung from PN 1 mo SftpC–Cre–ER > mTmG mouse (administered 1 mg tamoxifen at PN19) analysed 13 days after pulse-labelling and stained as in **d**. Note no co-expressing cells, indicating that ciliated (filled arrowhead) and NE (open arrowhead) cell types are not tagged. **j**–**l**, Lungs from mice labelled as in panel **i** analysed 13 days (**j**) and 192 days (**k**, **l**) later, stained as in **f**. Note double-positive $CCSP^+$/$SftpC^+$ cells (*) at the Badj, some of which are pulse-labelled (**j**). After 192 day chase, marked double-positive cells are solitary (**k**) or in doublets (**l**). **m**, Quantification of lung cell types marked under different labelling and lineage trace conditions. The number of marked and total cells of each type scored (and Badj analysed for $CCSP^+$/$SftpC^+$ cells) is shown for each genotype and age analysed. For the SftpC–Cre–ER line, the dosage of tamoxifen (tamox) and interval time until analysis is also indicated. Because bronchial maintenance involves proliferation without significant cell dispersion[50] (as we find for alveolar maintenance), the presence of marked $CCSP^+$/$SftpC^+$ cells primarily in isolation (>90%) at advanced ages suggests they did not contribute significantly to physiological bronchiolar renewal. d, days; PN, postnatal. Scale bar, 50 μm (**a**–**l**), 10 μm (insets in **a**).



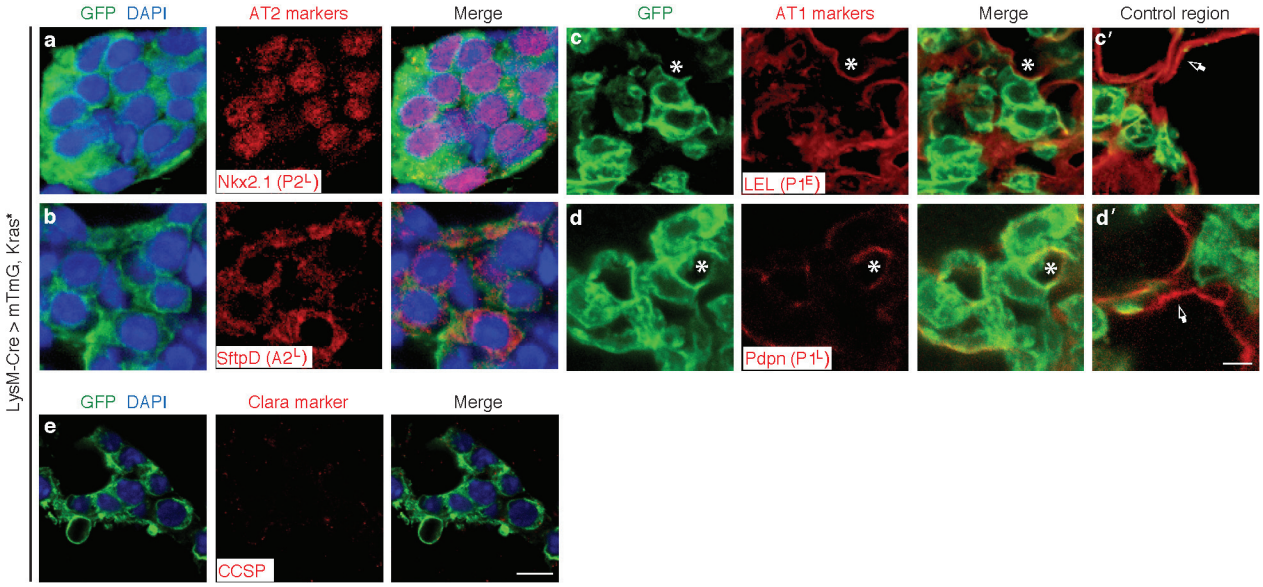
Extended Data Figure 5 | Lineage tracing alveolar type 2 (AT2) cells using SftpC–Cre–ER. SftpC–Cre–ER > mTmG mice were administered 1 mg tamoxifen (Tamox) at PN18 then analysed later by staining for the lineage label (GFP, green) and AT2 cells (SftpC, red). **a**, **b**, 13 days later only AT2 cells are marked (**a**), whereas after 212 days flat AT1 cells also express the AT2 lineage tag (**b**). A peripheral renewal focus (mesothelium indicated by dotted line) involving multiple alveoli (asterisks) is shown, similar to results using LysM–Cre. Quantification revealed that 94% of AT2 cells were marked 13 days after tamoxifen induction and 97% after 192 days (see Extended Data Fig. 5m), indicating that the $SftpC^{+}$ population is maintained by self-duplication, and that new AT2 cells do not derive from another cell population during physiological ageing. PN, postnatal day; bar, 100 μm.



Extended Data Figure 6 | Alveolar type 2 (AT2) founder cell functional marker expression and self-duplication *in vivo*, and reprogramming into AT1 cells *in vitro*. **a**–**c**, 16 month LysM-Cre > Confetti mouse lung stained for mCFP lineage tag (green), AT2 cell marker (Nkx2.1, red), and DAPI (blue) to identify clonal renewal foci. **a**, AT2 founder cell (filled arrow) associated with a daughter AT1 cell (open arrow) shown to express the $A2^{L}$ marker LAMP-1 (white), a protein associated with surfactant-containing lysosomes in mature, secretory AT2 cells. **b**, **c**, A clonal focus (**b**, boxed region) in which an AT2 cell has generated two additional AT2 cells (filled arrows) but no AT1 cells (**c**, close up of boxed region), demonstrating isolated self-duplication without AT1 cell reprogramming *in vivo*. Scale bars, 25 μm (**a**–**c**). **d**, Freshly isolated AT2 cells (Fig. 3f) cultured 4 days on glass with 10% serum adopt flat AT1 morphology (E-cadherin, green) and initiate AT1 marker expression (Aqp5, not shown). Scale bar, 10 μm.



Extended Data Figure 7 | Clonogenic response to activated Kras using SftpC–Cre–ER and CCSP–Cre–ER. **a**–**c**, Lungs of SftpC–Cre–ER (**a**) and CCSP–Cre–ER (**b**) mice carrying oncogenic $Kras^{LSL\text{-}G12D/+}$ (Kras*) and Rainbow (Rbw), injected with 3 mg (**a**) or 1 mg (**b**) of tamoxifen (tamox) at indicated ages to induce Kras* expression and clonal lineage marking of StfpC- and CCSP-positive cells, respectively. Lungs were analysed after 17 days at PN43 (**a**) or after 39 days at PN81 (**b**). Note that multifocal tumours result when SftpC–Cre–ER mice are induced at 25 (**a**) days of age, whereas induction of Kras* in CCSP–Cre–ER mice at 42 days (**b**) shows many unresponsive cells and doublets throughout the bronchi (Br) as well as small clonal tumours located at bronchoalveolar duct junctions (Badj). **c**, An Sftpc–Cre–ER mouse carrying Kras* and mTmG alleles induced at PN116 by injection of 2 mg tamox and analysed after 53 days also demonstrates adenomas. Dotted line indicates mesothelium; PN, postnatal day; scale bar, 100 μm.



Extended Data Figure 8 | Marker expression analysis of alveolar type 2 (AT2) cell derived adenomas. LysM–Cre > mTmG, $Kras^{LSL\text{-}G12D/+}$ (abbreviated Kras*) lungs stained for the AT2-lineage marker (GFP, green), the nuclear stain DAPI (blue), and the indicated AT2, AT1, and Clara cell markers (red). **a**–**e**, Note tumour cells (green) maintain expression of AT2 markers (**a**, **b**) and do not turn on Clara (**e**) or AT1 markers (**c**, **d**), except possibly rare cells (*). **c′** and **d′** show control (non-tumour) regions with normal AT1 staining (arrows). Scale bars, 10 μm.

Extended Data Table 1 | Developmental expression patterns of alveolar epithelial cell markers

Marker class*	Antigen or lectin (synonym)	Subcellular localization†	Expression pattern‡ E17.5	E18.0	E18.1	E18.5
ATI markers						
$P1^E$	Ager (RAGE)	bm				
	LEL	m				
	RCA I	m				
$P1^L$	Pdpn (T1α)	am				
$A1^L$	Aqp5	am				
AT2 markers						
$P2^E$	SftpC	c				
	Muc1	am, c§				
	Ctsh	c				
$P2^L$	Nkx2.1 (Ttf-1)‖	n				
$A2^L$	SftpB	c				
	SftpD	c				
	Lyz2 (LysM)	c				
	Abca3	c				
	Lamp-1	c				
	Lamp-2	c				

* First position of marker class name denotes whether alveolar lineage marker is first expressed by the bipotent progenitor (P) or a mature alveolar (A) cell type, second position denotes the mature cell type labelled by the marker (AT1 (1) or AT2 (2) cell), and third position denotes whether specificity is achieved early (E) or late (L) in the differentiation program.
† bm, basal membrane; m, membrane; am, apical membrane; c, cytoplasmic; n, nuclear.
‡ A pair of bipotent progenitors (squares) at a fixed representative airway position at embryonic day (E) 17.5 is shown developing into an AT1 cell (thin rectangle) and an AT2 cell (square with oval lamellar body) at E18.5. Green, cell expressing AT1 marker; red, cell expressing AT2 marker; *, cell in which marker subsequently turns off; dashed line, cell in which marker expression has recently ceased or reduced.
§ Changes localization from apical in progenitor to apical and cytoplasmic in AT2 cell.
‖ Expression is reduced but persists in AT1 cells.
LEL, *Lycopersicon esculentum* lectin; RCA I, *Ricinus communis* agglutinin I lectin.