TITLE

Let-7 restrains an oncogenic epigenetic circuit in alveolar type 2 cells to prevent ectopic formation of fibrogenic AT2 intermediates and pulmonary fibrosis.

AUTHORS

Matthew J. Seasock[1,2], Md Shafiquzzaman[2], Maria E. Ruiz-Echartea[3], Rupa S. Kanchi[3,4], Brandon T. Tran[5,6], Lukas M. Simon[7], Matthew D. Meyer[8], Phillip A. Erice[1,2], Shivani L. Lotlikar[2], Stephanie C. Wenlock[9], Scott A. Ochsner[3], Anton Enright[9], Alex F. Carisey[10,13], Freddy Romero[11,14], Ivan O. Rosas[11], Katherine Y. King[6], Neil J. McKenna[3], Cristian Coarfa[3,4], Antony Rodriguez[2,4,12,*]

[1]Immunology & Microbiology Graduate Program, Baylor College of Medicine, Houston, TX, 77030.

[2]Department of Medicine, Immunology & Allergy Rheumatology, Baylor College of Medicine Houston TX, 77030.

[3]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, 77030.

[4]Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine Houston, TX, 77030.

[5]Graduate Program in Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, TX, 77030.

[6]Department of Pediatrics, Division of Infectious Diseases, Texas Children's Hospital, Baylor College of Medicine, Houston, TX, 77030.

[7]Therapeutic Innovation Center, Baylor College of Medicine, Houston, TX, 77030.

[8]Shared Equipment Authority, Rice University, Houston, TX 77005.

[9]Department of Pathology, University of Cambridge, Cambridge, CB2 1TN, UK

[10]William T. Shearer Center for Immunobiology, Texas Children's Hospital, Houston, TX, 77030.

[11]Department of Medicine, Section of Pulmonary and Critical Care, Baylor College of Medicine. Houston, TX, 77030.

[12]Center for Translational Research on Inflammatory Diseases, Michael E. Debakey, Baylor College of Medicine, Houston, TX, 77030.

[13]Current Address: Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN

[14]Current Address: Vertex Pharmaceuticals, 3215 Merryfield Row, San Diego, CA, 92121.

*Author Correspondence: antonyr@bcm.edu

ABSTRACT

Analysis of lung alveolar type 2 (AT2) progenitor stem cells has highlighted fundamental mechanisms that direct their differentiation into alveolar type 1 cells (AT1s) in lung repair and disease. However, microRNA (miRNA) mediated post-transcriptional mechanisms which govern this nexus remain understudied. We show here that the *let-7* miRNA family serves a homeostatic role in governance of AT2 quiescence, specifically by preventing the uncontrolled accumulation of AT2 transitional cells and by promoting AT1 differentiation to safeguard the lung from spontaneous alveolar destruction and fibrosis. Using mice and organoid models with genetic ablation of *let-7a1/let-7f1/let-7d* cluster (*let-7afd*) in AT2 cells, we demonstrate prevents AT1 differentiation and results in aberrant accumulation of AT2 transitional cells in progressive pulmonary fibrosis. Integration of enhanced AGO2 UV-crosslinking and immunoprecipitation sequencing (AGO2-eCLIP) with RNA-sequencing from AT2 cells uncovered the induction of direct targets of *let-7* in an oncogene feed-forward regulatory network including BACH1/EZH2 which drives an aberrant fibrotic cascade. Additional analyses by CUT&RUN-sequencing revealed loss of *let-7afd* hampers AT1 differentiation by eliciting aberrant histone EZH2 methylation which prevents the exit of AT2 transitional cells into terminal AT1s. This study identifies *let-7* as a key gatekeeper of posttranscriptional and epigenetic chromatin signals to prevent AT2-driven pulmonary fibrosis.

INTRODUCTION

Lung respiratory diseases rank as major global health challenges and are among the leading causes of death worldwide. Lung regeneration after injury is facilitated by specialized epithelial progenitor stem cells (PSC)s found in airway and alveolar regions [1-3]. These cells are increasingly recognized as crucial factors in the onset of human lung diseases. Interstitial lung diseases (ILD)s, are respiratory disorders, characterized by excessive accumulation of

extracellular matrix (ECM) and fibrotic tissue in alveoli leading to respiratory failure [4]. This condition is linked to abnormalities in facultative epithelial PSCs, especially alveolar type 2 (AT2) cells [5,6]. AT2 cells are crucial for lung alveolar tissue repair, serving dual roles: as PSCs that can differentiate into alveolar type 1 (AT1) cells and as homeostatic cells (hAT2) essential for surfactant production [7-9]. Recent studies in some lung injury models found that hAT2 cells become primed (pAT2), self-renew and give rise to heterogeneous population of KRT8^{+} transitional cells (hereafter referred to as ADI, alveolar differentiation intermediates) which can become AT1s [10-13]. Studies have shown that ADIs are characterized by activation of signaling pathways (such as P53, TGFβ, EMT) and their pathological accumulation promotes lung injury and fibrosis [10-13]. In lungs of Idiopathic Pulmonary Fibrosis (IPF) patients (the most common endotype of ILDs), cells resembling ADIs, termed “aberrant basaloid” cells (AB), have been implicated in promotion of fibrosis [13,14].

Despite advances in understanding the molecular basis of AT2 dysfunction, the role of microRNAs (miRNA)s in AT2 to AT1 differentiation remain largely unexplored. Studies in human IPF tissues and animal models have shown reduced expression of *let-7* miRNA family members [15-18], as well as TGFβ/SMAD3 mediated repression of *let-7d* in epithelial cells [16]. Systemic inhibition of *let-7* function with an antagomiR was found to trigger thickening of alveolar septa and remodeling [16]. *Let-7* is recognized as a tumor suppressor which prevents the proliferation and metastasis of cancer stem cells via unrestrained induction of multiway hub oncogenes such as EZH2, MYC, and KRAS [19-22]. Also, *let-7* has been shown to act as an orchestrator of key pathways and processes involved in pulmonary fibrosis, such as PI3K/AKT/MTOR and epithelial mesenchymal transition (EMT) [22,23]. Moreover, we also recently reported a protective role of *let-7* family in CD8^{+} T cell-driven IL17 inflammation and lung tissue destruction in chronic obstructive pulmonary disease which points to a broader role in advanced lung disease(s) [24].

This study delves into the molecular and cellular roles of the *let-7* family in AT2s, underscoring its vital role in lung tissue homeostasis, protection against injury, and fibrosis. We

revealed the crucial role of *let-7* as a molecular brake to uncontrolled expansion, activation of AT2 PSCs and the generation of excessive ADIs in the lung. Loss of *let-7* activity in AT2s also hampers cell fate where ADIs are prevented from transitioning into AT1s. Mechanistic studies identified a global miRNA-target interactome network controlled by *let-7* in AT2 cells which govern AT2 cell renewal, formation of fibrogenic ADIs, and blunted AT1 cell differentiation. Mechanistic experiments focus on the contribution of EZH2, a core component of the Polycomb Repressive Complex 2 (PRC2), which mediates tri-methylation of histone H3 on lysine 27 (H3K27me3), leading to chromatin compaction and gene silencing [25,26] . Our data found that *let-7afd*$^{-/-}$ AT2s exhibit upregulation of EZH2 methyltransferase activity which contributes to formation of ADIs and silencing of genes required for AT1 cell differentiation. Overall, our studies highlight a new dimension for *let-7* as a potent regulator in alveolar stem cell fate and epigenetic perturbations in pulmonary fibrosis.

RESULTS

Time-dependent downregulation of *let*-7 family expression during peak formation of ADIs after bleomycin-induced lung injury.

We examined *let-7* temporal expression dynamics after bleomycin-induced lung injury in mice (GSE195773)[27]. Our analysis of this published lung small RNA-seq dataset revealed transient down-regulation of *let-7* family members at 7- and 14-days relative to controls followed by return to baseline by 21-days post-bleomycin (Supplementary Fig. 1a). Total *let-7* activity decreased approximately 25% after 7-days post-bleomycin in lungs of mice compared to controls (Supplementary Fig. 1b). Approximately 75% of total activity was comprised from members: *let-7a*, *let-7b*, *let-7c*, *let-7d*, and *let-7f*. These *let-7* members are transcribed in part from conserved *let-7b/let-7c2* (*let-7bc2*) and *let-7a1/let-7f1/let-7d* (*let-7afd*) gene clusters [24,28] and contribute upwards of 20% and 55% respectively of total *let-7* activity in naive lung (Supplementary Fig. 1b,c). The temporal downregulation of *let-7* expression coincides with the

transient appearance and peak formation of injury associated KRT8^{+} intermediate AT2 cells which takes place one- and two-weeks following bleomycin injury [11,12]. The data also shows *let-7* expression returns to baseline coincident with disappearance of ADI cells. Based on these observations, we hypothesized that *let-7* family might govern ADI cell formation in alveolar regeneration.

Depletion of *let-7afd* in AT2 cells causes spontaneous lung injury acutely.

To understand the mechanism of *let-7* function in AT2 cell compartment, we utilized our *let-7bc2* and *let-7afd* conditional knockout (KO) mice[29]. We employed *let-7bc2*$^{f/f}$;*Sftpc-Cre*$^{ERT2/+}$, *let-7afd*$^{f/f}$;*Sftpc-Cre*$^{ERT2/+}$ mice with and without the *R26R-LSL-tdTomato* reporter (*Sftpc-tdT*) for phenotypic analysis – hereafter referred to as *let-7bc2*^{AT2} and *let-7afd*^{AT2} mice (Supplementary Fig. 1c). *Let-7bc2*$^{f/f}$, *let-7afd*$^{f/f}$ or wild-type (WT) *Sftpc-tdT* mice were used as controls and also treated intraperitoneally with tamoxifen (iTAM). Quantitative RT-PCR (qPCR) from flow sorted *Sftpc-tdT*$^{+}$ AT2 cells confirmed >90% excision of *let-7afd* cluster after iTAM (Supplementary Fig. 1d).

To assess the consequence of deletion of *let-7bc*2 or *let-7afd* clusters, lung histology and physiologic phenotypic analysis were carried out following 6-days of iTAM. Genetic loss of *let-7afd* in AT2 cells led to decreased arterial oxygen saturation (SpO2) following iTAM (Fig. 1a). Approximately 28% of *let-7afd*^{AT2} mice developed a diffuse pulmonary hemorrhage phenotype discerned by gross lung dissections and hematoxylin and eosin (H&E) stained sections (Fig. 1b,c). The histologic analysis of *let-7afd*^{AT2} mice revealed diffuse alveolar bleeding into the alveolar spaces and immune cell infiltration relative to controls (Fig 1c). The *let-7bc2*^{AT2} mice exhibited a milder temporal drop in SpO2 relative to *let-7afd*^{AT2} mice but without the hemorrhage phenotype (Fig. 1a).

Deletion of *let-7afd* in AT2 cells stimulates progenitor stem cell proliferation and

formation of ADI cells acutely in the lung.

We performed bulk RNA sequencing (RNA-seq) of whole-lung or flow cytometric sorting of *Sftpc-tdT*$^{+}$ cells respectively between *let-7afd*^{AT2} with a diffuse pulmonary hemorrhage phenotype and controls following 6-days of iTAM. Gene set enrichment analysis (GSEA) showed significant induction of cell cycle related genes (*e.g.*, *Mki67, Top2*), pAT2 (*e.g.*, *Foxp2*, *Hif1a*, *Il33*) and ADI (*e.g., Cldn4, Ctgf, Krt8)* identity genes upon deletion of *let-7afd* (Fig. 1d,e,f, Supplementary Fig. 2a,b,c). Loss of *let-7afd* led to downregulation of hAT2 (*e.g.*, *Sftpb, Sftpc*) and AT1 identity genes (Fig. 1e, Supplementary Fig. 2a,b,d). We additionally confirmed cycling AT2 lineage cells in lungs of *let-7afd*^{AT2} mice following 14-days of iTAM by immunofluorescence (IF) detection of nuclear KI67 in *Sftpc-tdT*$^{+}$ traced cells (Fig. 1g).

To determine if transcriptomic changes arise in part due to intrinsic phenotypic changes in AT2 cells, we carried out colony forming efficiency (CFE) measurements in combination with transcriptome analysis by bulk RNA-seq on *let-7afd*^{AT2} and control AT2 cell-derived alveolospheres (organoids) grown for 14 days in AT2 maintenance medium (AMM) [30] (Fig. 1h). CFE and spheroid diameters were increased in *let-7afd*^{AT2} over controls (Fig. 1i, j, Supplementary Fig. 2e). GSEA and gene ontology (GO) biological process enrichment analysis respectively revealed a striking parallel between *let-7afd*^{AT2} organoids and *Sftpc-tdT* RNA-seq datasets in the induction of genes involved in cell cycle/growth and cell division (Fig. 1k, Supplementary Fig. 2f). The *let-7afd*^{AT2} organoids exhibited upregulation of pAT2 and ADI markers such as *Ctgf* and *Trp53* (Supplementary Fig. 2f). However, in contrast with sorted lineage traced *let-7afd*$^{-/-}$ AT2 cells where hAT2 markers appeared repressed, cultured *let-7afd*$^{-/-}$ AT2 organoids instead expressed higher levels of AT2 identity markers as compared to controls (Supplementary Fig. 2f,g). The *let-7afd*$^{-/-}$ organoids also showed repression of basal and club identity markers suggesting that it does not redirect towards airway cell fate (Supplementary Fig. 2g). Collectively, this data indicates that acute loss of *let-7afd* expression disrupts lung homeostasis and sets off aberrant lung remodeling by triggering inappropriate growth expansion

of AT2 cell progenitor stem cells and formation of ADI cells.

Integrated biochemical and transcriptomics analysis in the generation of a *let-7* mRNA targetome in AT2 cells.

To validate and identify functional targets of *let-7*, we performed enhanced Argonaute 2 UV-cross-linking immunoprecipitation-sequencing (AGO2-eCLIP) [31,32] in WT mice following 6-days of bleomycin injury (Fig. 2a). To focus on *let-7*-mRNA interactions the AGO2-eCLIP libraries were enriched with oligos corresponding to mature *let-7* family members (AGO2-eCLIP+let-7). The libraries yielded 13,027 chimeric *let-7*:mRNA called peaks in the 3'untranslated region (3'UTR) of genes (Fig. 2b). Motif analysis further showed highly significant and specific enrichment of *let-7* "seed" region binding motifs within chimeric *let-7* peaks (Fig. 2c). Collapsing of chimeric *let-7* peaks into individual transcription units yielded 2009 target genes in acute bleomycin-injured lungs (Fig. 2d).

To enrich for functional targets of *let-7* in AT2 cells, we integrated the AGO2-eCLIP+let-7 targetome with the sorted *Sftpc-tdT*$^{+}$ and/or organoid transcriptome datasets. Initially, we applied Sylamer[33] on *let-7afd*$^{-/-}$ transcriptome data to ascertain *let-7* post-transcriptionally represses mRNA targets. Sylamer showed prominent miR-mediated mRNA destabilization since the motifs corresponding to *let-7* were enriched within the 3'UTRs of induced genes in *let-7afd*$^{-/-}$ *Sftpc-tdT*$^{+}$ cells compared to controls (Fig. 2e). Next, we made use of Venn Diagram intersection analysis to select biochemically validated and significantly upregulated *let-7* target genes in sorted *Sftpc*-tdT^{+} and organoid transcriptome datasets. We refined the analysis by including experimentally published and validated direct *let-7* targets in the miRTarBase database[34]. The systematic approach revealed a *let-7* AT2 targetome comprised of 394 genes (Fig. 2f,g, Supplementary Tables 1,2).

Pathway discovery of *let-7* mRNA targetome identifies leading edge genes associated

with stem cell renewal, cell growth, and cell differentiation.

To identify the central pathways controlled by *let-7* we performed gene function annotation and functional enrichment analysis of the *let-7* targetome (Supplementary Tables 1,2,3). This analysis highlighted diverse networks of *let-7* targets associated with an array of cellular programs including stem cell renewal, growth, and differentiation. Interestingly, KEGG enrichment analysis showed the top pathways as miRNAs in cancer, focal adhesion, ECM-receptor interaction, and P53 signaling (Fig. 2h, Supplementary Table 3). We and others have shown miRNAs typically govern cell programs upstream of transcription factors (TF) and multiway hub genes in gene regulatory networks (GRN) to reinforce cell identity and cell fate decisions in development and disease[20-22,31,33]. Annotation of the *let-7* targetome revealed TFs which includes gene activators (*e.g. E2f1*, *Hif1a*, *Myc*), gene repressors (*E2f6*) and context dependent inducers/repressors of gene expression (*e.g.*, *Arid3a*, *Bach1*, *Ezh2, Foxp2*) (Supplementary Tables 1,2). Major hub genes, *Kras* and *Nras* were also identified as targets of *let-7* (Supplementary Tables 1,2). To ascertain whether hub TFs influence transcription-wide changes downstream of *let-7* in AT2 cells, we applied ChIP-Seq regulatory analysis on transcriptome datasets. This analysis showed significant enrichment for MYC, E2F1, E2F6, and BACH1 targets in genes upregulated after depletion of *let-7afd* (Fig. 2i). On the other hand, targets of EZH2 were highly overrepresented in downregulated genes upon loss of *let-7afd* (Fig. 2i). Intriguingly, the same is also true for the EZH2 partner proteins SUZ12/EED implying that absence of *let-7afd* augments EZH2 methyltransferase activity (Fig. 2i). Additionally, loss of *let-7afd* promotes the induction of KRAS pathway in AT2 cells (Fig. 2j). Protein interaction and functional enrichment analysis revealed strong connectivity of EZH2 with hub targets implicating the existence of a *let-7* dependent oncogenic GRN (OGRN) in AT2 cells (Fig. 2k). As we are interested in mechanisms of AT2 elicited lung injury and regeneration, we note the unique properties of OGRN genes as sufficient drivers of out-of-control stem cell growth and failed terminal differentiation in disease [25,35-38]. We annotated the location of the *let-7* chimeric peaks

and binding motifs for *let-7* within 3'UTRs (Fig. 2l,m, Supplementary Fig. 3a, Supplementary Tables 1,2). *Bach1*, *E2f1*, *Ezh2*, *Myc*, *Kras*, and *Nras* contain evolutionarily conserved sites and were previously identified as direct targets of *let-7* in non-AT2 cells [19-22,39]. *Hif1a* emerged as a direct target of *let-7*; however, our annotation of the binding site indicates it is missing in humans so we did not prioritize it (Supplementary Fig. 3a). We confirmed significantly increased mRNA expression of select *let-7* OGRN genes by qPCR detection in sorted *let-7afd*$^{-/-}$ *Sftpc-tdT*$^{+}$ cells following 14 days of iTAM (Fig. 2n).

Transcriptional convergence of alveolar and airway stem cells show enrichment of *let-7* targetome genes in ADI cells.

We evaluated the expression dynamics of the *let-7* targetome during AT2 cell "bridging" into ADIs following bleomycin injury from a published scRNA-seq dataset[12]. At the same time, we also incorporated expression data for MHC-II^{+} club airway cells because they were also associated with production of AT2 intermediate cells in bleomycin injury. Interestingly, *Bach1, Ezh2, Kras, Myc,* and Nras expression levels were increased along the differentiation trajectory from AT2 towards ADIs (Fig. 2o). Conversion of MHC-II^{+} club cells into KRT8^{+} ADI showed a more nuanced pattern for *Ezh2* and *Bach1* indicated by an increase followed by a reduction of expression (Fig. 2o). The majority of *let-7* targetome genes in our dataset exhibit a gradual increase and peak in expression in ADIs relative to both AT2 and MHC-II^{+} club cells (Fig. 2p). Jaccard cell identity overlap analysis further elaborates that the *let-7* targetome genes exhibit greater enrichment in AT1s and transitional AT2s than hAT2s (Fig. 2q). Together, these results support the notion that *let-7* serves as a braking mechanism in generation of pAT2/ADIs following lung injury and alveolar regeneration upstream of OGRN. These data prompted us to explore the phenotypic consequences of chronic absence of *let-7afd* as a catalyst for ILD.

Chronic deletion of *let-7afd* in AT2 cells induces progressive and spontaneous lung

fibrosis.

We examined the *let-7bc2*^{AT2} and *let-7afd*^{AT2} mice at 1- and 6-months following iTAM. At 1-month, both the *let-7bc2*^{AT2} and *let-7afd*^{AT2} mice (both sexes) showed spatially heterogenous disruption of the distal lung architecture marked by enlargement of alveolar spaces and increased presence of alveolar and peribronchial leukocytes and hemosiderin-laden macrophage infiltrates (Fig. 3a,b, Supplementary Fig. 4a,b). After 1-month of iTAM, approximately 30% of *let-7afd*^{AT2} mice furthermore exhibited heterogenous disorganization of alveoli denoted by septal wall thickening and the appearance of interstitial cells (Fig. 3a,b). The *let-7afd*^{AT2} mice additionally exhibited isolated areas of septal alveolar destruction, fibroblastic foci, and collagen positive staining with Masson's trichrome primarily at the periphery of the lung (Fig. 3a,b). These phenotypes worsened with aging to 6-months with thickening of interstitium distinguished by alveolar destruction, lung collagen deposition, and fibroblastic aggregates evident in >90% of mice (Fig. 3c,d). Alveolar septal destruction and collagen deposition in *let-7afd*^{AT2} mice were additionally supported by increased Ashcroft Scoring, Destructive Index (DI), and hydroxyproline levels respectively (Fig. 3e,f, Supplementary Fig. 4c). A parallel histologic evaluation of *let-7bc2*^{AT2} mice showed a milder phenotype without lung collagen deposition (Fig. 3c,d,f). Consistent with notable inflammatory phenotype detected by histology, the *let-7bc2*^{AT2} and *let-7afd*^{AT2} mice furthermore display increased total and differential immune cell counts in bronchoalveolar lavage fluid (BALF) (Fig. 3g). Pulmonary biomechanics and spirometry measurements further showed that the *let-7afd*^{AT2} mice exhibit progressive restrictive impairment in lung function beginning at 1-month of deletion (Supplementary Fig. 4d). To discern the transcriptome of the *let-7afd*^{AT2} mice, we carried out bulk RNA-seq on whole lungs of mice at 6-months of iTAM (Fig. 3h). Differential gene expression analysis and GSEA indicated marked induction of inflammation and fibrosis pathways in lungs of *let-7afd*^{AT2} mice compared to controls (Fig. 3i). Overall, these data establish that loss of *let-7* activity specifically in AT2 cells models progressive clinically relevant pulmonary fibrosis.

Ablation of the *let-7afd* in AT2 stimulates the persistence of ADI cells.

Based on this data, we hypothesized that AT2 loss of *let-7afd* could drive the proliferative expansion of pro-fibrogenic AT2 cell intermediates. We validated the presence of supernumerary AT2 lineage cells in lungs of *let-7afd*^{AT2} mice compared to controls after 1- and 5-months of deletion by IF labeling and quantification of *Sftpc*-tdT^{+} cells (Fig. 4a). Analysis of contiguous lung H&E sections for fibroblastic aggregates and IF detection of KRT8 revealed the spatial co-occurrence of ADIs in damaged areas with mesenchymal cell expansion and alveolar destruction in *let-7afd*^{AT2} mice after 1-month of iTAM (Fig. 4b). Further quantitative immunostaining analysis confirmed the persistence of KRT8^{+} ADIs in lung parenchyma of *let-7afd*^{AT2} mice compared to controls at 1- and 5-months of deletion (Fig. 4c,d). Cells expressing KRT8 exhibited distinct cell shape including elongated squamous morphology as recently described for ADIs (Fig. 4b,c)[11,12]. IF also revealed increased *Sftpc-tdT*$^{+}$ labelled cells co-expressing the ADI cell marker, CLDN4[10-12] in *let-7afd*^{AT2} mice compared to controls (Fig. 4e,f). Reinforcing our hypothesis, at 6-months of iTAM lungs of *let-7afd*^{AT2} mice furthermore exhibit enhanced and prolonged expression of ADI cell markers (Fig. 4g).

Deficiency of *let-7afd* stimulates abnormal changes in AT2 cell size and shape and upregulates the expression of fibrosis-inducing genes.

To determine whether deletion of *let-7afd* promotes AT2 cell hypertrophy and/or other maladaptive ultrastructural cell changes associated with clinical ILD[5], we made use of transmission electron microscopy (TEM) in lungs of mice after 3-months of iTAM. TEM stereological analysis revealed the *let-7afd*^{AT2} mice exhibit enlarged or hypertrophic AT2 cells with increased area relative to controls (Fig. 5a,b). Notably, *let-7afd*$^{-/-}$ AT2 cells also displayed increased number and total cell area of lamellar bodies (LBs) with disorganized lamellae (Fig. 5a,b). The morphometric measurements of circularity furthermore show cell shape elongation of

295 *let-7afd*$^{-/-}$ AT2 cells compared to controls (Fig. 5a,b). Remarkably, *let-7afd*^{AT2} mice also exhibit
296 binucleated AT2 cells (11.1%, 5 of 45 cells) a unique feature not evidenced in control mice (0%,
297 0 of 44 cells) (Fig. 5a).

298 We hypothesized that *let-7* with its vast OGRN targets and impact on cell size and shape
299 might orchestrate fibrogenesis upstream PI3K/AKT/MTOR and ECM receptor/EMT pathways.
300 Correspondingly, analysis of our RNA-seq datasets revealed that loss of *let-7afd* activates gene
301 sets involved in the PI3K/AKT/MTOR and EMT in AT2 cells (Fig. 5c, Supplementary Table 3).
302 Consistent with transcriptome data, loss of *let-7afd* led to an increase of *Sftpc*-tdT^{+} cells labeled
303 with phospho-p70 S6 kinase (pS6K, a pan marker of MTORC pathway activation) compared to
304 controls at 1-month of deletion (Supplementary Fig. 5a). Integrative pathway annotation of the
305 *let-7* AT2 cell targetome revealed a functionally interconnected protein interactome network
306 comprised of over 50 gene targets of *let-7* directly or indirectly involved in PI3K/AKT/MTOR and
307 EMT pathways. Of note, the interactome was composed of many previously associated fibrosis
308 genes, including TFs (*Myc*) [40,41], receptors (*Itgb3*, *Itgb4*), ligands (*Col1a1*, *Col5a2*), enzymes
309 (*Nras*, *Kras*) [42,43] and cytoskeletal proteins (*Acta2*, *Tpm1*) (Fig. 5d,e, Supplementary Tables
310 1,2,3). The OGRN genes (*e.g.*, *Bach1*, *Ezh2*, *E2f1*) were included in the network given their
311 inductive roles in tissue fibrosis [44-46] and engagement of PI3K/AKT/MTOR [37,47] and EMT [23,48,49]
312 pathways respectively (Fig. 5d). To better understand the potential contribution of *Bach1* and
313 *Ezh2* to pulmonary fibrosis in *let-7afd*^{AT2} mice, we examined their expression by immunostaining
314 and qPCR. Consistent with our comprehensive phenotypic analysis, ablation of *let-7afd* led to a
315 significant increase in lung BACH1^{+} *Sftpc*-tdT^{+} and EZH2^{+} *Sftpc*-tdT^{+} cells following 5-months of
316 iTAM (Fig. 5f,g,h). Furthermore, transcript levels for *Bach1*, *Ezh2,* and other fibrogenic OGRN
317 genes (*E2f1*, *Myc, Kras*, *Nras*) remained induced in *let-7afd*$^{-/-}$ AT2 cells over controls after 2-
318 months of deletion (Fig. 5i). This suggests absence of *let-7afd* drives unchecked expression of
319 OGRN at the top of the hierarchy and ECM structural proteins downstream. Thus, we propose
320 the overall GRN endows *let-7afd*$^{-/-}$ AT2 cells with the maladaptive traits conducive to alveolar

destruction and remodeling.

Genetic loss of *let-7afd* makes AT2 cells vulnerable to DNA damage, senescence, and apoptosis and impedes AT2 to AT1 terminal differentiation.

Studies have shown that in different models of lung injury and fibrosis, AT2 cells can undergo cell senescence or apoptosis related to disruptions in quality control programs[5]. Moreover, some of the AT2 targets of *let-7* including *Casp3* and *Myc* [35] are associated with cell senescence and apoptosis (Supplementary Tables 1,2,3). We examined lungs of mice with γH2AX as a marker of DNA damage in conjunction with the β-galactosidase activity assay for cell senescence. At 1-month of deletion, lung immunostaining showed significant accumulation and frequency of γH2AX in SFTPC$^+$ AT2 cells of *let-7afd*^{AT2} mice compared to control mice (Fig. 6a,b). Correspondingly, *let-7afd*^{AT2} mice showed robust accumulation of β-galactosidase substrate X-gal in AT2s indicating induction of a cellular senescence program (Fig. 6d). In concordance with this data, *Trp53* transcript levels were induced in purified in *let-7afd*$^{-/-}$ *Sftpc*-tdT$^+$ cells relative to controls (Fig. 6e). We also assessed for apoptotic AT2 cells in lungs of mice by qPCR and IF detection of *Casp3* and active Casp-3 respectively. This analysis revealed increased *Casp3* expression and active Casp-3 *Sftpc*-tdT$^+$ labeled cells in *let-7afd*^{AT2} mice relative to controls (Fig 6a,c and Supplementary Fig. 6a). Congruent with these results, sorted *Sftpc*-tdT cells and whole lungs of *let-7afd*^{AT2} mice exhibit induction of DNA repair, P53 signaling, and apoptosis pathways (Supplementary Fig. 6b,c).

We posited that the persistence of ADIs in *let-7afd*^{AT2} mice might relate to a block in their ability to trans-differentiate into AT1s. Thus, we tested whether *let-7afd* is necessary for differentiation of AT2s into AT1s with lung immunostaining with AT1 cell marker (AGER). As expected, we did not observe *Sftpc-tdT*$^+$ cells co-expressing AGER in homeostatic control lungs (Fig. 6a,e). Surprisingly, lungs of *let-7afd*^{AT2} mice did not exhibit an appreciable increase in

frequency of AGER^{+}*Sftpc*-tdT^{+} cells despite the significant alveolar destruction after 1- and 5-months of deletion (Fig. 6a,f). Elaborating loss of AT1 cells, absence of *let-7afd* furthermore resulted in downregulation of AT1 identity genes in lungs of *let-7afd*^{AT2} mice (Supplementary Fig. 6d).

To ascertain whether loss of *let-7afd* blunts AT2 to AT1 cell differentiation in an intrinsic manner, we compared the formation of ADIs and AT1 differentiation by IF in control and *let-7afd*$^{-/-}$ AT2 cell organoids grown in AT1 cell differentiation medium (ADM) (Fig. 7a). On day 7 of ADM culture, CLDN4^{+} and KRT8^{+} expression levels and positive cells were enhanced significantly in *let-7afd*$^{-/-}$ organoids over controls which corroborates the increased ADIs in lungs of *let-7afd*^{AT2} mice (Fig. 7b,c, Supplementary Fig. 7a,b). In further support of impaired AT1 differentiation, *let-7afd*$^{-/-}$ organoids also exhibited reduced AGER^{+} and HOPX^{+} (another AT1 marker) cells and expression in contrast with controls (Fig. 7d,e,f Supplementary Fig. 7c). Of note, *let-7afd*$^{-/-}$ organoids also exhibit reduced expression and fewer LGALS3^{+} (a late ADI marker) cells compared to controls which highlight a late block in AT1 differentiation (Fig. 7g, Supplementary Fig. 7d). By comparison, SFTPC^{+} cells and levels were enhanced in *let-7afd*$^{-/-}$ organoids suggesting that absence of *let-7* promotes not only the accumulation of ADIs but also reinforces retention of AT2 identity (Fig. 7d,g, Supplementary Fig. 7e). To ascertain whether absence of *let-7afd* promotes DNA damage or death in differentiation impaired *let-7afd*$^{-/-}$ organoids, we examined γH2AX and active CASP3 by IF. The number of γH2AX^{+} and active CASP3^{+} cells were increased in *let-7afd*$^{-/-}$ AT2 cultures compared with controls (Supplementary Fig. 7f-h). Thus, *let-7afd* serves as a pivotal brake to the formation of lung injury inducing ADI cells which plays a critical role in AT1 cell differentiation. Furthermore, loss of *let-7afd* renders AT2 cells highly prone to senescence, DNA damage and apoptosis in a cell intrinsic manner which could aggravate pulmonary disease in the lung[50].

To explore the mechanism of impaired AT2 to AT1 differentiation of *let-7afd*$^{-/-}$ AT2 cultures, we examined BACH1 and EZH2 expression by immunostaining in ADM cultures.

Interestingly, *let-7afd*$^{-/-}$ cultured AT2 organoids exhibited higher levels and frequency of BACH1^{+} and EZH2^{+} cells as compared to controls (Fig. 7d,i,j, Supplementary Fig. 7i,j). Both BACH1 and EZH2 have been associated with impaired epithelial cell differentiation upon overexpression[46,51] but neither has been functionally studied in AT2 progenitor stem cells. Next, we examined their expression during ADI to AT1 cell transition after bleomycin-induced injury on a published scRNA-seq dataset [12]. This analysis showed a gradual decrease in their dynamic expression as ADIs move in the direction of AT1 cell differentiation (Fig. 7k). Other targets of *let-7* exhibited a similar expression profile (Fig. 7k, Supplementary Fig. 7k). Taken together these analyses lend support that loss of *let-7afd* hampers AT2 cell "bridging" into AT1 cells in part via ectopic expression of oncogenes that antagonize stem cell differentiation.

***Let-7afd* orchestrates AT2 to AT1 cell differentiation at the level of EZH2 histone methylation.**

In disease pathophysiology, unchecked expression of EZH2 and PRC2 activity promote histone trimethylation of H3K27, which ultimately leads to epigenetic silencing of genes associated with repression of cell growth and induction of cell differentiation [25,26]. Recent research has demonstrated that the genetic removal of *Ezh2* hampers the growth of lung epithelial cells and reduces the emergence of cycling epithelial cells, KRT17^{+} basal cells, and KRT8^{+} transitional epithelial cells in murine organoids [52] which is essentially the opposite phenotype of *let-7afd*$^{-/-}$ AT2s. Prompted by these findings, we investigated whether the unchecked expression of EZH2 in *let-7afd*$^{-/-}$ AT2s hinders their differentiation into AT1s. Our analysis began with lung RNA-seq data from *let-7afd*^{AT2} mice, six months post-deletion, focusing on the potential suppression of EZH2/PRC2 target genes. Subsequent ChIP-Seq regulatory network analysis showed significant enrichment of EZH2 and EED/SUZ12 targets in genes downregulated in *let-7afd*^{AT2} mice compared to controls (Fig. 8a). Then we examined histone H3K27me3 by immunostaining in lungs as well as cultured organoids derived from control and

let-7afd^{AT2} mice. Detection and quantification of H3K27me3 revealed a significant increase of H3K27me3 signal within nuclei of *Sftpc-tdT*$^{+}$ AT2s in *let-7afd*^{AT2} lungs compared to controls (Fig. 8b,c). By extension, organoid *let-7afd*$^{-/-}$ cultures exhibited higher nuclear expression levels for H3K27me3 compared to control cells (Fig. 8d,e).

Next, we performed cleavage under targets and release using nuclease (CUT&RUN)-DNA sequencing on chromatin obtained from flow sorted *let-7afd*$^{-/-}$ or control *Sftpc-tdT*$^{+}$ cells on day 6 after iTAM treatment. CUT&RUN was performed to monitor H3K27me3 and H3K27ac histone modifications (Fig. 8f). Differential peak analysis was carried out to identify increased and repressed histone peaks in *let-7afd*$^{-/-}$ compared to control WT AT2 cells. To determine which biological GO pathways are impacted by H3K27me3 in *let-7afd*$^{-/-}$ AT2 cells we applied the GREAT tool on differential peaks. The repressed H3K27me3 peaks in *let-7afd*$^{-/-}$ AT2 cells did not exhibit significant GO enrichment. In contrast, the increased H3K27me3 peaks in *let-7afd*$^{-/-}$ AT2 cells compared to control cells exhibited high enrichment in genes involved in epithelial cell differentiation and development (Fig. 8g, Supplementary Table 4). To identify potential TF regulatory networks, we performed motif analysis via HOMER. Three TF motifs were enriched including a motif closely matching SOX4, a master regulator of EMT and candidate regulator of AT2 to AT1 differentiation[12] (Fig. 8h, Supplementary Table 5). As expected, Cistrome analysis indicate high similarity of induced H3K27me3 peaks in *let-7afd*$^{-/-}$ AT2 cells to published SUZ12 and EZH2 ChipSeq datasets (Fig. 8i). To focus on repressed genes where H3K27me3 marks are likely to have a functional effect, we integrated the epigenomic analyses with our bulk transcriptomic dataset from flow sorted *let-7afd*$^{-/-}$ AT2 cells matched after 6-days of iTAM. We identified 267 differentially downregulated genes with increased H3K27me3 peaks in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells (Fig. 8j,k, Supplementary Table 6). Of note, this analysis revealed repression of genes that represent potential bottleneck defects in AT1 differentiation. This includes increased H3K27me3 deposition and repression of *Aqp5*, *Cav1,* and *Spock2* in *let-7afd*$^{-/-}$ cells compared to controls (Fig 8k,m, Supplementary Table 6). In further support of epigenomic data, AQP5 and

CAV1 expression levels and positive cells were significantly repressed in *let-7afd*$^{-/-}$ organoids over controls (Supplementary Fig. 8a-e). Additionally, loss of *let-7afd* elicited the repression of important Notch, Hippo, or Wnt signaling pathway genes involved in specification and differentiation of lung epithelium (Fig. 8k, Supplementary Tables 6,7). Further, scrutiny of the expression dynamics of repressed genes during ADI cell "bridging" into AT1s indicate a gradual increase in their expression which reinforces fundamental roles in terminal differentiation (Fig. 8l). In conclusion, this collective data shows that temporal dysregulation of EZH2 methyltransferase activity is a crucial pathophysiologic mechanism which contributes to a block in AT2 to AT1 cell differentiation upon deletion of *let-7afd*.

Human Aberrant Basaloid (AB) cells from IPF patients exhibit enrichment of the *let-7* interactome.

We speculated that reduced *let-7* activity in lung epithelial cells of IPF patients might contribute to accumulation of AB cells. To investigate this hypothesis, we analyzed a published single-cell RNA sequencing dataset from human IPF lungs[14] and looked for differences in the gene expression profiles of *let-7* and the murine AT2 targetome. Few targets of *let-7* were upregulated in IPF AT2 relative to control AT2 cells (Fig. 9a,b, Supplementary Table 8). On the other hand, AB cells exhibited significant enrichment of the murine *let-7* interactome relative to AT1 and AT2 cells from IPF patients (Fig. 9a,b, Supplementary Tables 9,10). Furthermore, upregulated targets of *let-7* in AB cells compared to control AT2s figure prominently in fibrosis pathways including ECM and EMT (Fig. 9c). Relative to AT1 and AT2 cells, AB cells furthermore exhibit enrichment of EMT pathway genes (Fig. 9d). The *let-7* OGRN genes, *ARID3A*, *BACH1*, *EZH2*, *NRAS* were significantly upregulated in AB cells compared to control AT2 cells while FOXP2 and MYC were repressed (Fig. 9e and Supplementary Table 9). We also evaluated for consensus target signature for EZH2 in IPF. AB cells did not show reduced expression of EZH2 targets compared to control or IPF AT2 cells (Fig. 9e). However, PRC2

target genes were collectively reduced in IPF AT2 cells relative to control AT2 cells (Fig. 9e). Our analyses suggest the *let-7* OGRN axis might contribute to the formation of AB cells in IPF; however further studies will be required for more direct evidence.

DISCUSSION

Our study lends support to the importance of *let-7* miRNA in AT2s as a molecular gatekeeper in lung homeostasis, injury, and healing. We identified transient and coordinated downregulation of *let-7* family after bleomycin-induced lung injury, indicating a tightly regulated post-transcriptional response which paves the way for the emergence of repair-associated AT2s. This interpretation aligns with our studies indicating that AT2-cell specific genetic ablation of *let-7afd* removes a “brake” on AT2 activation resulting in uncontrolled cell growth, inflammation, and accumulation of ADIs in spontaneous pulmonary fibrosis in mice. Remarkably, this phenotype ensues without the requirement of exogenous injury. Additionally, we discovered that *let-7* not only restricts AT2 cells from transitioning into ADI intermediates but it is also crucial for exiting this state. Thus, we confirm that impaired *let-7* expression is a repressed marker of lung injury and pulmonary fibrosis, as well as a central hub for maintaining AT2 quiescence and directing differentiation towards AT1s.

To date experimental studies in epithelial PSC lung regeneration and disease primarily focused on protein-coding genes. The *let-7* miRNA family exerts broad cellular effects by silencing hundreds of target mRNAs in GRNs[22]. Here we demonstrate that *let-7* simultaneously regulates an extensive GRN in AT2 cells, including oncogenes such as BACH1, EZH2, E2F1, MYC, KRAS, and NRAS, previously linked to uncontrolled cell growth, inflammation, and profibrotic effects [37,38,40-46]. Our data favors that this OGRN is persistently upregulated in AT2 cells due to impaired *let-7* post-transcriptional regulation which would support progressive pulmonary fibrosis disease. Analysis of their expression by AT2 cell state trajectory modeling post-bleomycin injury further implicates these oncogenes in alveolar tissue regeneration and

repair. Impressively, OGRN *let-7* targets are enriched in cycling and pAT2 and ADIs but their expression decreases as cells progress toward AT1 identity. With permanent loss of *let-7afd*, our data indicates that AT2 cell identity is held within ADI state, where the *let-7* OGRN genes are most highly expressed.

The leading-edge targets of *let-7*, such as EZH2, BACH1, E2F1, MYC, KRAS, and NRAS, are associated with epithelial stem cell growth, transformation, and enhanced metastatic potential involving processes which include EMT [25,35,40,48]. Moreover, published studies on *let-7* have shown that its loss stimulates cell growth and EMT programs, while its overexpression limits these processes [22,23]. Intriguingly, BACH1 and EZH2 have also been associated with partial EMT in cancer [53,54], where epithelial cells linger as heterogeneous transitional cells which do not transition into mesenchymal cells but instead exhibit a continuum of epithelial and mesenchymal traits which potentiate migration, inflammation, and tumorigenesis [55]. Lineage tracing and scRNA-seq studies indicate that AT2 transitional cells do not transform into fibroblasts via EMT [10-13]; instead, they adopt a mixed gene expression program characterized by simultaneous expression of epithelial and mesenchymal genes, which influences lung remodeling and ECM deposition [13,56]. Recent studies showed that the accumulation of AT2 transitional cells can drive lung injury and ILD [13,56]. Given these observations, we propose a model where loss of *let-7* contributes to a continuum of partial EMT in transitional AT2s tied to deregulation of the OGRN. We note that the *let-7* OGRN likely functions as a powerful coherent or synergistic feedforward-loop because these genes converge on similar phenotypic disturbances in diseased epithelial cells. Also, our integrative analysis identified upregulated PI3K/AKT/MTOR and ECM genes as targets of *let-7*, which may further aggravate lung alveolar injury and fibrosis.

We specifically investigated the contribution of chromatin remodeler EZH2 which restricts the exit of AT2 transitional cells into AT1s upon *let-7afd* loss. Our mechanistic findings showed that depletion of *let-7afd* in AT2 cells promotes ectopic EZH2 expression and aberrant

H3K27me3 modifications in nuclei of cells. Prior studies have shown the key importance of heritable H3K27me3 chromatin compaction mediated by the EZH2/PRC2 complex in preserving cellular fate identity of stem cells during cell division [25,26]. H3K27me3 marks are associated with gene repression to exert control over cell identity and differentiation. As our epigenomic results showed, *let-7afd*$^{-/-}$ AT2s exhibit increased H3K27me3 deposition in repressed genes associated with epithelial cell differentiation, cell identity, and cell-cell communication. We highlight *Aqp5*, *Cav1* and *Spock2* as important AT1 cell identity markers [10-12] which were repressed upon *let-7afd* loss. *Spock2* is a known important growth suppressor target of EZH2 in lung adenocarcinomas [57]; and *Cav1,* an essential component of caveoli, which promotes alveolar septa defects and pulmonary fibrosis upon deletion in mice [58]. Thus, we propose that a critical mechanism of the *let-7*/EZH2 axis is aberrant epigenetic reprogramming which prevents AT2 transitional cells from transitioning into AT1 cells.

In future work it will be important to ascertain the molecular function and sufficiency of EZH2 and other *let-7* targets identified in induction of alveolar destruction and fibrosis downstream of *let-7*. These studies will necessitate a toolbox of oncogene gain-of-function animal models to recapitulate features of deficiency of *let-7afd*. Furthermore, although we demonstrated dynamic changes in AT2 cells upon loss of *let-7afd* by integrated approaches in mice, we did not achieve transcriptomic and epigenomic depth at a single cell resolution which would have allowed more insight on the role for *let-7* on AT2 cell transition states and alveolar niche cells. Furthermore, this study did not focus on how loss of *let-7afd* leads to activation of AT2 cells and contributes to DNA damage, senescence, and apoptosis. More studies will be needed to resolve the contribution of these processes in the sequelae of lung fibrosis upon deletion of *let-7afd*.

In conclusion, our study provides insights into *let-7* as a potent braking mechanism to oncogene dysregulation in AT2 cell-driven pulmonary fibrosis. Also, our observation that AB cells from IPF patients express higher levels of the *let-7* targetome offers the enticing possibility

that targeted delivery of *let-7* could be used to slow the progression of fibrosis. We point that several reports indicate the effectiveness of targeted *let-7* delivery in mitigating fibrosis [59,60], and pharmacologic targeting of EZH2 is employed against both cancer and fibrosis [61,62]. Furthermore, since AT2 cells are significant source of lung adenocarcinomas and IPF patients are at high risk of cancer; we propose that existing anti-cancer drugs might offer new approaches to target AB cells and IPF.

MATERIALS AND METHODS

Mice and tamoxifen administration.

Mice of both sexes were used for experiments. We recently described the creation and PCR genotyping strategy of the C57BL/6 isogenic *let-7bc2*$^{fl/fl}$ or *let-7afd*$^{fl/fl}$ mice [24]. Mice were crossed to *Sftpc*$^{tm1(cre/ERT)Blh}$ (*Sftpc-CreER*^{T2}), *Rosa26R-CAG-LSL-tdTomato (R26tdT)* purchased from JAX and genotyped by PCR. For acute phenotypic studies at 6-day or 14-days of iTAM induced deletion, mice aged 8-12 weeks old mice were given 4 doses of 160 mg/kg tamoxifen (Sigma, T5649) dissolved in corn oil (Sigma, C8267) by intraperitoneal injections (i.p.) every other day. Chronic studies over 1 month of duration gave mice 4 doses of 80 mg/kg iTAM every other day once a month. Mice were genotyped by PCR from ear clippings with published primers [24]. All studies with mice were approved by the Baylor College of Medicine Institutional Animal Care and Use Committee and followed the National Research Council Guide for the Care and Use of Laboratory Animals.

Physiologic measurements.

Arterial oxygen saturation (SpO2) measurements were collected using the STARR Life Sciences MouseOx Plus (Oakmont, Pennsylvania) on awake mice with a neck cuff sensor to obtain five readings over a 30 min interval. The mice were shaved on the neck and allowed 45 minutes to acclimate without restrains with a neck cuff sensor prior to good quality recordings.

The lung respiratory biomechanics values were obtained with the flexiVent (Scireq) [63] system. Mice undergoing flexiVent procedure were subjected to anesthesia, followed by tracheostomy and cannulation with an 18G metal cannula. The mice were placed under mechanical ventilation at a respiratory rate of 150 breaths/min, a tidal volume of 10 mL/kg and a PEEP set at 3cmH2O in enclosed chamber for pulmonary function measurements following manufacturer guidelines. The average of three measurements with a coefficient of determination ≥ 0.95 were calculated for each mouse.

Lung histology and histomorphometry analysis.

Mouse lungs were perfused with sterile PBS via right ventricle cardiac perfusion. For formalin-fixed paraffin embedding tissues, lungs were inflated with instillation of 4% paraformaldehyde (PFA) via a tracheal cannula at 25-cm H_2O pressure for 15mins, and then tied by nylon suture. The lungs were then transferred into 50 mL polypropylene tube containing PFA for overnight fixation at 4°C. Then lung lobes were separated, transferred into tissue cassettes, and then immersed in 70% ethanol at 4°C. Samples were provided to the Human Tissue Acquisition and Pathology (HTAP) Core at BCM for paraffin embedding, cutting of 5μM thick sections, and staining by H&E or Masson's trichrome. The PFA-fixed lung cryosections were obtained in similar manner, but after the overnight fixation step in PFA at 4°C, the lungs were equilibrated with 30% sucrose for 1-2 days, prior to embedding in OCT. Samples were then cut into 7-10uM thick sections for the analysis. β-galactosidase (X-gal) lung staining was performed on frozen lung sections using the Senescence Cells Histochemical Staining Kit (Sigma CS0030) following manufacturer instructions. Hydroxyproline Colorimetric Assay Kit (Abcam, ab222941) was used to measure collagen in lung lobes following kit instructions and optical density was measured using a Varioskan LUX plate reader (Thermo Fisher Scientific, VL0000D0).

Bronchoalveolar lavage and Prussian blue staining.

Bronchoalveolar lavage for measurements of total and differential cell counts was done as we described [24]. Prussian Blue staining of BALF cytospins was performed using the Iron Stain Kit (Sigma, HT20-1KT) following kit instructions.

Immunostaining.

For immunostaining of paraffin embedded lung tissues or organoid sections, antigen retrieval was performed with 1X BioCare Diva Decloaker RTU (BioCare, DV2004G1) buffer in a Decloaking Chamber NxGen (BioCare, DC2012) at 95°C or 110°C for 15mins. Slides were then allowed to cool in antigen retrieval buffer for 10 mins before washing in distilled water. Both paraffin embedded and thawed cryo-sectioned slides were permeabilized with 0.3% Triton X-100 in 1% Normal Goat Serum (Abcam, ab7481) for 10 mins. Antigen blocking was performed in 5% Normal Goat Serum for 1 hour. Samples were then incubated overnight at 4°C with the following primary antibodies diluted in 1% Normal Goat Serum: Rat anti-Ki67 conjugated FITC (1:50, Invitrogen, 11-5698-82), Mouse anti-active/pro-caspase3 (1:20, Invitrogen, MA1-91637), Mouse anti-EZH2 (1:100, Invitrogen, 14-9867-82), Rabbit anti-BACH1 (1:200, Novus Biological, NBP2-55113), Mouse anti-γH2AX (1:200, Novus Biological, NB100-74435), Rabbit anti-CLDN4 (1:200, Invitrogen, 36-4800), Rabbit anti-Phospho-p70 S6K (1:200, Cell Signaling, 9234), Rabbit anti-pro-SFTPC (1:200, Millipore, AB3786), Rabbit anti-RFP/tdTomato (1:200, Rockland, 600-401-379), Rat anti-Krt8/TROMA-I (1:20, DSHB, TROMA-I-s), Rat anti-Galectin-3/LGALS3 (1:500, Cedarlane, CL8942AP), Rat anti-RAGE/AGER (1:100, R&D Systems, MAB1179-100), Rabbit anti-H3K27me (1:200, Invitrogen, MA5-11198), Rabbit anti-Aqp5 (1:100, Invitrogen, PA5-36529), Rabbit anti-Caveolin-1 (1:200, Cell Signaling, 3267), and Mouse anti-HopX (1:200, Santa Cruz, sc-398703). The next day, samples were incubated for 1 hour at room temperature in the following secondary antibodies diluted in 1% Normal Goat Serum with DAPI (1:10000, Sigma, D9542): Alexa Fluor 488 goat anti-mouse (1:2000, Thermo Fisher Scientific, A11029),

Alexa Fluor 488 goat anti-Rabbit (1:2000, Thermo Fisher Scientific, A-32731), Alexa Fluor 488 goat anti-Rat (1:2000, Thermo Fisher Scientific, A11006), Alexa Fluor 555 goat anti-Mouse (1:2000, Thermo Fisher Scientific, A32727), Alexa Fluor 555 goat anti-Rabbit (1:2000, Thermo Fisher Scientific, A21428), Alexa Fluor 555 goat anti-Rat (1:2000, Thermo Fisher Scientific, A21434), Alexa Fluor 647 goat anti-Rabbit (1:2000, Thermo Fisher Scientific, A21244), Alexa Fluor 647 goat anti-Rat (1:2000, Thermo Fisher Scientific, A21247), Alexa Fluor 790 Goat anti-Rabbit (1:2000, Thermo Fisher Scientific, A11369). In some cases, paraffin embedded lung sections were treated after secondary staining with TrueBlack Lipofuscin Autofluorescence Quencher (Biotium, 23007) for 1 min following product instructions to block autofluorescence. Slides were mounted with ProLong Glass Antifade Mountant (Thermo, P36980).

Lung cell suspensions and purification of AT2 cells.

Isolation of *Sftpc-tdT*$^{+}$ lineage derived AT2s was performed as previously described with minor modifications [30,64]. Briefly, mouse lungs were perfused with cold PBS via right ventricle of the heart and then inflated with 1 mL digestion solution containing 5 units/mL dispase (Corning, 354235), 450 units/mL collagenase type I (Gibco, 17100-017), 0.33 units/mL DNase I (Sigma, 10104159001), and 1X antibiotic-antimycotic (Thermo Fisher Scientific, A5955-100ML) in Advanced DMEM/F-12 (Thermo Fisher Scientific, 12634028). Lung lobes were then minced into 1-2 mm^3 pieces, transferred into 4 mL digestion solution, and incubated at 37°C with rotation for 40 mins. Samples were dissociated by vigorously mixing by P1000 pipette tip until no large pieces remained. Samples were filtered through a 70µm strainer, pelleted, and then resuspended in 5 mL Advanced DMEM/F-12 with 1X antibiotic-antimycotic. We coated 100mm tissue culture dishes with 42 µg anti-Mouse CD45 (BioLegend, 103101) and 16 µg anti-Mouse CD32 (BioLegend, 156402) in 7 mL PBS for 24-48 hours at 4°C. Before use, antibody-coated dishes were washed twice with PBS and once with Advanced DMEM/F-12. Single cell suspensions were incubated on antibody-coated dishes at 37°C with 5% CO_2 for 45 mins to

remove adherent cells and immune cells. Supernatants were collected, pelleted, and treated with 1 mL RBC Lysis Buffer (Invitrogen, 00-4333-57) for 1-min before washing with 5 ml Advanced DMEM/F-12. Cells were then resuspended in 10% FBS, 25 mM HEPES-buffered DMEM (Corning, 10-013-CV). Cell suspensions were stained with Live/Dead Fix Blue (Invitrogen, L34962) followed by flow cytometry of live *Sftpc-tdT*$^{+}$ cells on a FACS Aria (BD Biosciences). For the bulk RNA-seq, sorted Sftpc-tdT^{+} cells (~700,000) were pooled from 3 mice per group (n = 3).

RNA extraction and quantitative real-time PCR (qPCR).

Cell pellets or lung lobes from individual mice were snap frozen in 350µl of 1:100 β-mercaptoethanol in the RLT Lysis Buffer from the RNAeasy mini plus Kit (Qiagen, 74134). Cell samples were thawed and vortexed while lung samples were homogenized for total RNA extraction intended for qPCR or bulk RNA-seq. The RNAeasy mini plus Kit (Qiagen, 74134) was used for removal of genomic DNA and RNA isolation. Quality control was performed with NanoDrop 2000C Spectrophotometer (Thermo Fisher Scientific). Equal amount of input RNA was used for preparation of cDNA with the High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Thermo Fisher Scientific, 4374966). Pre-designed Taqman probes were used with TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific, 4444557). *Gapdh* served to normalize gene expression by ΔΔCT method. The qPCR was conducted and analyzed on ViiA 7 System (Applied Biosystems). The following Taqman probes were used: *Bach1* (Mm01344527_m1), *Casp3* (Mm01195085_m1), *E2f1* (Mm00432939_m1), *Ezh2* (Mm00468464_m1), *Foxp2* (Mm00475030_m1), *Kras* (Mm00517492_m1), *Myc* (Mm00487804_m1), *Nras* (Mm03053787_s1), *Trp53* (Mm01731287_m1), *Gapdh* (Mm99999915_g1). The validated and published *pri-let-7a1/d/f1* transcript Taqman probe[24] was custom designed by Applied Biosystems (44411114, arfvmhy).

AT2 3D organoid cultures and differentiation.

Culture of 3D organoids was carried out as previously described with minor modifications [30,65]. Lineage labeled *Sftpc*-tdT$^+$ AT2 cells were purified from 8-10 week old mice on day 6-day after iTAM treatment. AT2 cells were resuspended in thawed Cultrex UltiMatrix Matrigel (Biotechne, BME001-05) at 1,000 cells/µL. For each biological replicate, organoids were initiated in each well of 6-well plate into five 20 µl domes containing ~20,000 cells and incubated at 37°C with 5% CO_2 for 20 mins without media. When domes hardened, 2mL of alveolar maintenance media (AMM) without Interleukin-1β was added to each well. The organoids were fed AMM without Interleukin-1β every other day and grown for 10-14-days for subsequent analysis. For RNA-seq, the domes were washed with 1 mL sterile PBS and then incubated in 1 mL Cultrex Organoid Harvesting Solution (Biotechne, 3700-100-01) on ice for 1 hour. Domes were gently detached from 6-well plates by spatula, resuspended in 10 mL PBS, pooled for each biological replicate, and then pelleted at 300 x g for 10 mins at 4°C. Cells were washed and pelleted several times with 10 mL sterile PBS to remove Matrigel. In AT2 to AT1 differentiation experiments, organoids were switched from AMM into AT2-Differentiation medium (ADM) on day 10 followed by an additional 7-days as previously described [30,65]. For immunostaining analysis, the AT2 organoids were fixed and sectioned as in previous publication[65]. Organoid colony forming efficiency (CFE) and alveolosphere size diameters were quantified using FIJI (ImageJ, v2.14.0).

Bulk RNA-seq processing, pathway enrichment, and regulatory network analysis.

Purified RNA (> 800 ng) from each of the samples was used for the bulk RNA-seq analysis. RNA samples with RIN scores greater than 7 were selected for RNA-seq with the Agilent Bioanalyzer 2100. Samples were submitted to the Genomic and RNA Profiling Core at BCM for RNA-seq and library preparation. Nonstrand-specific, polyA+ -selected RNA-seq libraries were

generated using the Illumina TruSeq protocol. Libraries were sequenced to a median depth of 35-40 million 100-bp single-end reads on a HiSeq sequencer (Illumina). The FASTQ files underwent quality control metrics (trimming), read mapping, and gene annotation (GRCm38/mm10 genome) with default parameters with the RNA-seq Alignment v2.0.2 pipeline from Illumina's BaseSpace (https://basespace.illumina.com/). The RNA-Seq Differential Expression v1.0.1 workflow (Illumina BaseSpace) was used with default parameters to extract gene counts and to obtain differentially expressed genes (DEG)s with an adjusted p value (FDR) < 0.05 between control and experimental samples. Gene Ontology and pathway enrichment analysis was carried out with web based tools Enrichr (https://maayanlab.cloud/Enrichr/) [66] or with Metascape (https://metascape.org/) [67]. The gene set enrichment analysis was performed with GSEA v4.2.3 Mac application using standard settings. The epithelial cell marker gene sets used for GSEA were obtained from Mouse Lung Injury & Regeneration web tool (https://schillerlabshiny.shinyapps.io/Bleo_webtool/) [12]. The Jaccard overlap scores for the *let-7* targetome were calculated by taking the size of the intersect and dividing it by the size of the union from epithelial cell types defined in GSE141259 [12]. AT2 and ADI pseudotime trajectory values were obtained from Strunz et al.,[12] Heatmaps and plots were generated with R (v2022.12.0). To predict transcription factors whose function is impacted by *let-7*, ChIP-Seq regulatory network analysis was performed on the *let-7afd*$^{-/-}$ RNA-seq datasets as previously described [68].

Small RNA-seq data analysis.

The published raw data files corresponding to small RNA-seq from lungs of mice after vehicle or bleomycin treatment[27] were extracted with the NCBI SRA toolkit. The SRA data files were analyzed with miRNA Analysis v1.0.0 application workflow (Illumina BaseSpace). To obtain mature *let-7* miRNA normalized counts, standard settings were used corresponding to murine GRCm38/mm10 genome and miRbase DB v21 database. Plots were generated with R

(v2022.12.0) by normalizing bleomycin treated samples to untreated samples for each timepoint.

CUT&RUN sequencing, processing, and data analysis.

The CUT&RUN sequencing protocol was done as previously described [69,70]. Lineage labelled *Sftpc*-tdT$^+$ AT2 cells (~100,000) were flow sorted from mice as indicated above. Cells were pooled from 2 mice per group for CUT&RUN sequencing (n = 1). Cells were washed with wash buffer (50 ml total H_2O with 20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, and one Roche Complete protein inhibitor tablet (Millipore, 11873580001). Cells were then incubated with activated concanavalin-A beads (Bang Laboratories, L200731C) for fifteen minutes at room temperature. Slurry was then separated with a Miltenyi MACSiMAG separator (Miltenyi 130-092-168), washed once, and resuspended in wash buffer containing 0.05% digitonin (Digitonin Wash buffer, Millipore, 300410). Digitonin wash buffer was used from here on out and cells kept on ice. The samples were placed on chilled tube rack kept on ice when not rotating for remainder of protocol. The samples were incubated with primary antibodies overnight at 4°C rotating and then washed three times and incubated on rotator with Digitonin wash buffer containing pAG-MNase (Epicypher 15-1016) for one hour at 4°C. Antibodies used were Rabbit anti-H3K27ac (Cell Signaling Technologies, 8173; 1:50 concentration), Rabbit anti-H3K427me3 (Cell Signaling Technologies, 9733; 1:50 concentration), and Rabbit-antimouse IgG (Jackson ImmunoResearch, 315-005-003; 1:100). After, samples were washed three times they were resuspended in 100uL Digitonin wash buffer and chilled on ice rack for two minutes. 2µL of 100mM $CaCl_2$ was added to each sample to catalyze MNase cleavage. Cells were then kept at 4°C for 40 minutes. Cleavage was stopped by adding equal volume stopping buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.05% Digitonin, 0.05 mg/mL glycogen, 5 ug/mL RNase A, 2 pg/mL heterologous spike-in DNA (Epicypher 18-1401)), and samples were heated to 37°C for 20 minutes for MNase/Histone/DNA complex release into supernatant. Slurry was separated

with magnetic stand, and supernatant with DNA was collected into new tubes. DNA was purified with phenol chloroform extraction (Thermo Fisher Scientific, 15593031) and resuspended in ultrapure water (Invitrogen, 10-977-023). The samples were then submitted for library generation and sequencing to Admerahealth (South Plainfield, NJ).

The amount of DNA was determined by Qubit 2.0 DNA HS Assay (ThermoFisher, Massachusetts, USA) and quality was assessed by Tapestation High sensitivity D1000 DNA Assay (Agilent Technologies, California, USA). The KAPA HyperPrep kit (Roche, Basel, Switzerland) was used for library preparation. Assessment of library quality and quantity was done with Qubit 2.0 DNA HS Assay (ThermoFisher, Massachusetts, USA), Tapestation High Sensitivity D1000 Assay (Agilent Technologies, California, USA) for sequencing on an Illumina NovaSeq (Illumina, California, USA) with a read length configuration of 150 paired-end (PE).

CUT&RUN sequencing data was trimmed for low quality reads and Illumina adapters using the trim galore package (https://github.com/FelixKrueger/TrimGalore). Data was mapped using bowtie2 against the mouse genome UCSC build mm10. Peaks were called using MACS2 with IgG serving as control in each sample. Differential peaks between control and *let-7afd*$^{-/-}$ samples were determined using DiffReps, with significance achieved for FDR adjusted q value < 0.05, and fold change exceeding 1.25x. Signal tracks were generated using DeepTools v2.1.0 https://deeptools.readthedocs.io/en/2.1.0/index.html) and BEDTOOLS v2.31.0 suite (https://bedtools.readthedocs.io/en/latest/content/bedtools-suite.html) for annotation of differential peaks overlapping genes +/- 3kb from its transcription start site and +/-10kb from its genebody [71]. GO terms for differential peaks were determined using GREAT v4.0.4 [72]. Enriched motifs were derived using the findMotifsGenome.pl HOMER tool [73]. The Cistrome data toolkit was used to compared differential histone peaks to transcription factors, chromatin regulators (http://dbtoolkit.cistrome.org) [74]. Data peaks were visualized with the UCSC genome browser.

AGO2 eCLIP and *let-7* interactome network analysis.

To obtain lung samples for AGO2-eCLIP+let-7, eight-week-old wild-type C57BL/6 mice were administered with constant flow of 3.5% isoflurane during intratracheal administration of 4U/kg bleomycin. The mice were then euthanized 6-days after treatment. Mouse lungs were cleared by cardiac perfusion and then inflated with 1mL RNAse free PBS prior to harvesting and snap freezing with liquid nitrogen. Samples were provided to EclipseBio (San Diego, California) for generation of AGO2-eCLIP+let-7 libraries. The AGO-eCLIP+let-7 was performed in duplicate from 2 lungs of mice with chimeric ligation of miRNA and mRNA as previously described but with following modifications for lung sample processing and the enrichment of *let-7* targets [31,32]. Briefly, approximately $20x10^6$ cells from the lung were plated on a dish then UV crosslinked, then pelleted and suspended in 1 ml volume of eCLIP lysis mix. The lysates were then sonicated (QSonica Q800R2) for 5 minutes, 30 seconds on / 30 seconds off with an energy setting of 75% amplitude, followed by digestion with RNase-I (Ambion). A primary mouse monoclonal AGO2/EIF2C2 antibody (sc-53521, Santa Cruz Biotechnology) was pre-coupled for 1hr with Sheep Anti-Mouse IgG Dynabeads (Thermo Fisher Scientific, 11202D) and then added to the homogenized lysate for a 2hr immunoprecipitation (IP) at 4°C. After the IP, a 2% portion of the sample was taken as paired Input and the remainder of lysate washed with eCLIP high stringency wash buffers. The miRNA:mRNA chimeric ligation step was done at room temperature for 1hr with T4 RNA ligase (NEB). The IP samples were then treated with alkaline phosphatase (FastAP, Thermo Fisher) and then T4 PNK (NEB) for ligation of barcoded RNA adapters and isolation of AGO2-RNA complexes [31,32]. To generate AGO2-eCLIP libraries enriched for *let-7* chimeras, 5' biotinylated ssDNA probes antisense to murine *let-7a-5p*, *let-7b-5p*, *let-7d-5p*, *let-7f-5p* miRNAs were ordered from Integrated DNA Technologies and resuspended in nuclease-free water to a final concentration of 100µM. Each probe was bound individually to Capture Beads (Eclipse Bioinnovations) and afterward mixed. The resulting probe-bound bead mixture was added to the reserved AGO2-eCLIP enrichment fractions and allowed to hybridize over a 1.5-hour period, with incrementally decreasing temperatures. Beads

were washed to remove nonspecific RNAs and afterward, bound RNAs were eluted from beads using a DNase treatment. Eluted RNA was purified using an RNA Clean and Concentrator-5 Kit (Zymo Research, cat. #R1015/R1016) and taken through reverse transcription, DNA adapter ligation, and PCR amplification to generate sequencing libraries. Library sequencing was performed as SE122 on the NextSeq 2000 platform. BedGraph files of *let-7* chimeric peaks were generated for visualization with UCSC genome browser.

The AGO2-eCLIP+let-7 processing, and bioinformatics analysis was done by Eclipsebio (San Diego, California) with a proprietary analysis pipeline v1 developed from several published eCLIP publications [31,32]. The original bioinformatic pipelines are available (https://github.com/yeolab/eclip; https://github.com/YeoLab/chim-eCLIP). A brief description follows: The umi_tools (v1.1.1) tool was used to prune Unique molecular identifier (UMIs) from reads. Next, cutadapt (v3.2) was used for trimming 3' adapters. Repetitive elements and rRNA sequences were removed and non-repeat mapped reads mapped to murine GRCm38/mm10 using STAR (v2.7.7a). PCR duplicates were removed with umi_tools (v1.1.1). miRbase (v22.1) miRNAs were reverse mapped to reads with bowtie (v1.2.3) and miRNA portion of each read trimmed. STAR (v2.7.7a) was used to map the miRNA portion of each read to the genome. PCR duplicates were trimmed, and miRNA target cluster peaks were identified with CLIPper (v2.1.0). Individual clusters were annotated with corresponding miRNAs names and peaks annotated to transcripts GENCODE release M25 (GRCm38.p6). Peaks present in both biological samples were selected for downstream analysis. De novo HOMER (v4.11) [73] motif analysis was also applied to confirm the *let-7* seed region as the most enriched motif in dataset. *Let-7* chimeric peaks with overlap to the 3'UTR regions of protein coding genes were designated as the AGO2-eCLIP *let-7* interactome. The Sylamer test was done as previously described[75]. Downregulated ($p < 0.05$) genes in *let-7afd*$^{-/-}$ *Sftpc*-tdT and/or cultured organoids were overlayed with AGO2-eCLIP *let-7* interactome with Excel software (Microsoft). Pathway enrichment analysis for the *let-7* interactome was carried out with Enrichr [66] and Metascape [67]. Protein-protein interaction

networks were derived from Metascape (using parameter of Combined Core) and then visualized and organized with EnrichmentMap/Cytoscape (v3.10.1).

Human IPF Atlas scRNA-seq integration.

A published scRNA-seq dataset deposited in IPF Cell Atlas from control and IPF patients was used[14]. The cell annotations for cell clusters were provided by the authors [14]. Cell-level expression for individual genes was obtained with R package Seurat [76]. Gene signature scores were generated using the UCell method. Violin plots for individual genes or gene signatures in specific cell types were generated using the R statistical package v4.4.0. Dotplot expression of individual gene in selected cell types were generated using Seurat [76]. The Hallmark EMT and ENCODE Consensus ChIP-X gene targets for EZH2 used for UCell enrichment were obtained from Enrichr [66].

Transmission electron microscopy (TEM) and analysis.

We utilized a published protocol for fixation [6], processing and embedding of lung tissues for preservation of AT2 cell lamellar body (LB) ultrastructure. Mice were euthanized, followed by an incision of the IVC artery and descending aorta to allow for cardiac draining of blood. Lungs were inflated with 1.2 mL of freshly prepared room temperature TEM fixative (4% paraformaldehyde, 2.5% glutaraldehyde, 0.02% picric acid in 0.1□M sodium cacodylate buffer, pH 7.3), tie sutured and then fixed overnight at 4°C. Strips of lung tissue were dissected, washed in 0.1M sodium cacodylate buffer, fixed in a solution of 1% OsO_4 and 1.5% potassium ferricyanide for 1hr with gentle agitation, washed several times with 0.1M sodium cacodylate buffer and then incubated overnight at 4 °C. Samples were then washed with ddH_20, stained with 1.5% uranyl acetate in the dark for 1hr, dehydrated with a graded series of ethanol, infiltrated with and embedded in EMbed812 epoxy resin (#14120, Electron Microscopy

Sciences) and then sectioned at 100 nm thickness with a Leica EM UC7 ultramicrotome. Images were collected with a JEOL JEM-1400Flash TEM operating at 120 kV of accelerating voltage and equipped with a high-contrast pole piece and an AMT NanoSprint15 Mk-II sCMOS camera. Images were examined with FIJI (ImageJ, v2.9.0) to quantify AT2 cell area (mm2), LB area (mm2), LB number/cell area (#/mm2) and AT2 cell circularity.

Microscopy, image acquisition, and quantification.

Slides were imaged on a Leica DM6000B (Leica, Wetzlar, Germany) using cameras CoolSNAP HQ2 (Teledyne Photometrics, USA) or Infinity3-6UR (Teledyne Lumenera, USA), and a Keyence BZ-X800 microscope (Keyence, USA). Brightfield color images captured by Leica microscope where white balance with FIJI (ImageJ, v2.14.0), and then stitched using AIVIA software (Leica, v11.0). Keyence images were stitched using BZ-X Series Advanced Analysis software (Keyence). Organoid domes were imaged on 6-well plates with a BioTek LionHeart LX microscope (Agilent, USA). Images were captured with 10X objective and stitched with built-in software. Cell counts were done manually based on markers and DAPI with FIJI software. Mean Fluorescence Intensity (MFI) was measured by FIJI software. Images of interest where then prepared using Adobe Photoshop and Illustrator. Mean linear intercept, Destructive index and Ashcroft scoring measurements of mouse lung morphometry were done as previously described [24,77-79].

Statistical analysis.

Values are represented as indicated in figure legends. P values were calculated using two-tailed unpaired or paired Student's t test. Nonparametric Rank test was used for Ashcroft Scoring. 1-way ANOVA with Sidak post-hoc test was used for the UCell enrichment analysis. A two-sided p value < 0.05 was statistically significant. Determination of sample size for animal experiments were based on pilot tests. Investigators were not blinded in the allocation of experiments. No

data was excluded from analysis. Mice of both genders were randomly assigned into experimental or control groups for experiments. Statistical analyses were performed using GraphPad Prism (v9.5.1) except for the RNA-seq and CUT&RUN which were done as described above.

COMPETING INTEREST STATEMENT

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We thank Drs. David Corry and Farrah Kheradmand for valuable scientific suggestions and for generous sharing of lab equipment. M. Sayeeduddin for histology at the Tissue Acquisition and Pathology Core (funded in part by P30 CA125123); and Joel M. Sederstrom for flow cytometry at the BCM and Cell Sorting Core (funded in part from NIH (CA125123 and RR024574); we thank the Texas Children's Hospital William T Shearer Center for Human Immunobiology and Dr. Ronald Parchem for access to microscope equipment. Bioinformatics analysis was partially supported by The Cancer Prevention Institute of Texas (CPRIT) grants RP210227 and RP200504, NIEHS grants P30 ES030285 and P42 ES027725. This work was supported by grants from the NHLBI (R01HL140398, HL167814 to AR), and NIGMS (T32 GM136554 to MS). functional pathway and protein interactome"

AUTHOR CONTRIBUTIONS

M.J.S. and A.R. conceptualized experiments, interpreted data, and wrote the manuscript. M.J.S., M.S., P.A.E., B.T.T, S.L.L. carried out experiments. M.D.M. carried out the TEM imaging. A.F.C. provided microscope image acquisition expertise. M.E.R-E., R.S.K. L.M.S., S.C.W., S.A.O. performed bioinformatic data analysis. K.Y.K. provided expertise and

supervision of epigenomics experiment. A.E., I.O.R., N.J.M., C.C., provided expertise and co-supervised bioinformatic data analysis.

REFERENCES

1 Basil, M. C. *et al.* The Cellular and Physiological Basis for Lung Repair and Regeneration: Past, Present, and Future. *Cell Stem Cell* **26**, 482-502 (2020). https://doi.org:10.1016/j.stem.2020.03.009

2 Hogan, B. & Tata, P. R. Cellular organization and biology of the respiratory system. *Nat Cell Biol* (2019). https://doi.org:10.1038/s41556-019-0357-7

3 Basil, M. C., Alysandratos, K. D., Kotton, D. N. & Morrisey, E. E. Lung repair and regeneration: Advanced models and insights into human disease. *Cell Stem Cell* **31**, 439-454 (2024). https://doi.org:10.1016/j.stem.2024.02.009

4 Wuyts, W. A. *et al.* Idiopathic Pulmonary Fibrosis: Best Practice in Monitoring and Managing a Relentless Fibrotic Disease. *Respiration* **99**, 73-82 (2020). https://doi.org:10.1159/000504763

5 Katzen, J. & Beers, M. F. Contributions of alveolar epithelial cell quality control to pulmonary fibrosis. *J Clin Invest* **130**, 5088-5099 (2020). https://doi.org:10.1172/JCI139519

6 Chung, K. P. *et al.* Mitofusins regulate lipid metabolism to mediate the development of lung fibrosis. *Nat Commun* **10**, 3390 (2019). https://doi.org:10.1038/s41467-019-11327-1

7 Barkauskas, C. E. *et al.* Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* **123**, 3025-3036 (2013). https://doi.org:10.1172/JCI68782

8 Nabhan, A. N., Brownfield, D. G., Harbury, P. B., Krasnow, M. A. & Desai, T. J. Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* **359**, 1118-1123 (2018). https://doi.org:10.1126/science.aam6603

9 Zacharias, W. J. *et al.* Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* **555**, 251-255 (2018). https://doi.org:10.1038/nature25786

10 Choi, J. *et al.* Inflammatory Signals Induce AT2 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration. *Cell Stem Cell* **27**, 366-382 e367 (2020). https://doi.org:10.1016/j.stem.2020.06.020

11 Kobayashi, Y. *et al.* Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis. *Nat Cell Biol* **22**, 934-946 (2020). https://doi.org:10.1038/s41556-020-0542-8

12 Strunz, M. *et al.* Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. *Nat Commun* **11**, 3559 (2020). https://doi.org:10.1038/s41467-020-17358-3

13 Wang, F. *et al.* Regulation of epithelial transitional states in murine and human pulmonary fibrosis. *J Clin Invest* **133** (2023). https://doi.org:10.1172/JCI165612

14 Adams, T. S. *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* **6**, eaba1983 (2020). https://doi.org:10.1126/sciadv.aba1983

15 Perera, U. E. *et al.* Evaluation of microRNA expression in a sheep model for lung fibrosis. *BMC Genomics* **22**, 827 (2021). https://doi.org:10.1186/s12864-021-08073-4

16 Pandit, K. V. *et al.* Inhibition and role of let-7d in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* **182**, 220-229 (2010). https://doi.org:10.1164/rccm.200911-1698OC

17 Kim, S. *et al.* Integrative phenotyping framework (iPF): integrative clustering of multiple omics data identifies novel lung disease subphenotypes. *BMC Genomics* **16**, 924 (2015). https://doi.org:10.1186/s12864-015-2170-4

18 Yang, G. *et al.* Discovery and validation of extracellular/circulating microRNAs during idiopathic pulmonary fibrosis disease progression. *Gene* **562**, 138-144 (2015). https://doi.org:10.1016/j.gene.2015.02.065

19 Kim, H. H. *et al.* HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev* **23**, 1743-1748 (2009). https://doi.org:10.1101/gad.1812509

20 Kong, D. *et al.* Loss of let-7 up-regulates EZH2 in prostate cancer consistent with the acquisition of cancer stem cell signatures that are attenuated by BR-DIM. *PLoS One* **7**, e33729 (2012). https://doi.org:10.1371/journal.pone.0033729

21 Johnson, S. M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635-647 (2005). https://doi.org:10.1016/j.cell.2005.01.014

22 Ma, Y., Shen, N., Wicha, M. S. & Luo, M. The Roles of the Let-7 Family of MicroRNAs in the Regulation of Cancer Stemness. *Cells* **10** (2021). https://doi.org:10.3390/cells10092415

23 Jin, B. *et al.* Let-7 inhibits self-renewal of hepatocellular cancer stem-like cells through regulating the epithelial-mesenchymal transition and the Wnt signaling pathway. *BMC Cancer* **16**, 863 (2016). https://doi.org:10.1186/s12885-016-2904-y

24 Erice, P. A. *et al.* Downregulation of Mirlet7 miRNA family promotes Tc17 differentiation and emphysema via de-repression of RORgammat. *Elife* **13** (2024). https://doi.org:10.7554/eLife.92879

25 Batool, A., Jin, C. & Liu, Y. X. Role of EZH2 in cell lineage determination and relative signaling pathways. *Front Biosci (Landmark Ed)* **24**, 947-960 (2019). https://doi.org:10.2741/4760

26 Comet, I., Riising, E. M., Leblanc, B. & Helin, K. Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. *Nat Rev Cancer* **16**, 803-810 (2016). https://doi.org:10.1038/nrc.2016.83

27 Strobel, B. *et al.* Time and phenotype-dependent transcriptome analysis in AAV-TGFbeta1 and Bleomycin-induced lung fibrosis models. *Sci Rep* **12**, 12190 (2022). https://doi.org:10.1038/s41598-022-16344-7

28 Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-1910 (2004). https://doi.org:10.1101/gr.2722704

29 Erice, P. A. *et al.* Downregulation of Let-7 miRNA promotes Tc17 differentiation and emphysema via de-repression of RORgammat. *bioRxiv* (2024). https://doi.org:10.1101/2023.10.12.562059

30 Konishi, S., Tata, A. & Tata, P. R. Defined conditions for long-term expansion of murine and human alveolar epithelial stem cells in three-dimensional cultures. *STAR Protoc* **3**, 101447 (2022). https://doi.org:10.1016/j.xpro.2022.101447

31 Krivdova, G. *et al.* Identification of the global miR-130a targetome reveals a role for TBL1XR1 in hematopoietic stem cell self-renewal and t(8;21) AML. *Cell Rep* **38**, 110481 (2022). https://doi.org:10.1016/j.celrep.2022.110481

32 Van Nostrand, E. L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* **13**, 508-514 (2016). https://doi.org:10.1038/nmeth.3810

33 Rodriguez, A. *et al.* Requirement of bic/microRNA-155 for normal immune function. *Science* **316**, 608-611 (2007). https://doi.org:10.1126/science.1139253

34 Huang, H. Y. *et al.* miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* **50**, D222-D230 (2022). https://doi.org:10.1093/nar/gkab1079

35 Dhanasekaran, R. *et al.* The MYC oncogene - the grand orchestrator of cancer growth and immune evasion. *Nat Rev Clin Oncol* **19**, 23-36 (2022). https://doi.org:10.1038/s41571-021-00549-2

36 Punekar, S. R., Velcheti, V., Neel, B. G. & Wong, K. K. The current state of the art and future trends in RAS-targeted cancer therapies. *Nat Rev Clin Oncol* **19**, 637-655 (2022). https://doi.org:10.1038/s41571-022-00671-9

37 Zhang, X. *et al.* Bach1: Function, Regulation, and Involvement in Disease. *Oxid Med Cell Longev* **2018**, 1347969 (2018). https://doi.org:10.1155/2018/1347969

38 Huang, C. L. *et al.* E2F1 overexpression correlates with thymidylate synthase and survivin gene expressions and tumor proliferation in non small-cell lung cancer. *Clin Cancer Res* **13**, 6938-6946 (2007). https://doi.org:10.1158/1078-0432.CCR-07-1539

39 Hou, W., Tian, Q., Steuerwald, N. M., Schrum, L. W. & Bonkovsky, H. L. The let-7 microRNA enhances heme oxygenase-1 by suppressing Bach1 and attenuates oxidant injury in human hepatocytes. *Biochim Biophys Acta* **1819**, 1113-1122 (2012). https://doi.org:10.1016/j.bbagrm.2012.06.001

40 Shen, Y. *et al.* c-Myc promotes renal fibrosis by inducing integrin alphav-mediated transforming growth factor-beta signaling. *Kidney Int* **92**, 888-899 (2017). https://doi.org:10.1016/j.kint.2017.03.006

41 Qin, H. *et al.* C-MYC induces idiopathic pulmonary fibrosis via modulation of miR-9-5p-mediated TBPL1. *Cell Signal* **93**, 110274 (2022). https://doi.org:10.1016/j.cellsig.2022.110274

42 Shields, M. A. *et al.* Snail cooperates with KrasG12D to promote pancreatic fibrosis. *Mol Cancer Res* **11**, 1078-1087 (2013). https://doi.org:10.1158/1541-7786.MCR-12-0637

43 Newbury, L. J., Wang, J. H., Hung, G., Hendry, B. M. & Sharpe, C. C. Inhibition of Kirsten-Ras reduces fibrosis and protects against renal dysfunction in a mouse model of chronic folic acid nephropathy. *Sci Rep* **9**, 14010 (2019). https://doi.org:10.1038/s41598-019-50422-7

44 Liu, Y. & Zheng, Y. Bach1 siRNA attenuates bleomycin-induced pulmonary fibrosis by modulating oxidative stress in mice. *Int J Mol Med* **39**, 91-100 (2017). https://doi.org:10.3892/ijmm.2016.2823

45 Zhang, Y. *et al.* E2F1 is a novel fibrogenic gene that regulates cholestatic liver fibrosis through the Egr-1/SHP/EID1 network. *Hepatology* **60**, 919-930 (2014). https://doi.org:10.1002/hep.27121

46 Le, H. Q. *et al.* An EZH2-dependent transcriptional complex promotes aberrant epithelial remodelling after injury. *EMBO Rep* **22**, e52785 (2021). https://doi.org:10.15252/embr.202152785

47 Yang, R. *et al.* E2F7-EZH2 axis regulates PTEN/AKT/mTOR signalling and glioblastoma progression. *Br J Cancer* **123**, 1445-1455 (2020). https://doi.org:10.1038/s41416-020-01032-y

48 Igarashi, K., Nishizawa, H., Saiki, Y. & Matsumoto, M. The transcription factor BACH1 at the crossroads of cancer biology: From epithelial-mesenchymal transition to ferroptosis. *J Biol Chem* **297**, 101032 (2021). https://doi.org:10.1016/j.jbc.2021.101032

49 Wang, T. *et al.* Transcription factor E2F1 promotes EMT by regulating ZEB2 in small cell lung cancer. *BMC Cancer* **17**, 719 (2017). https://doi.org:10.1186/s12885-017-3701-y

50 Yao, C. *et al.* Senescence of Alveolar Type 2 Cells Drives Progressive Pulmonary Fibrosis. *Am J Respir Crit Care Med* **203**, 707-717 (2021). https://doi.org:10.1164/rccm.202004-1274OC

51 Sato, M. *et al.* BACH1 Promotes Pancreatic Cancer Metastasis by Repressing Epithelial Genes and Enhancing Epithelial-Mesenchymal Transition. *Cancer Res* **80**, 1279-1292 (2020). https://doi.org:10.1158/0008-5472.CAN-18-4099

52 Byrd, A. L. *et al.* Dysregulated Polycomb Repressive Complex 2 contributes to chronic obstructive pulmonary disease by rewiring stem cell fate. *Stem Cell Reports* **18**, 289-304 (2023). https://doi.org:10.1016/j.stemcr.2022.11.009

53 Serresi, M. *et al.* Functional antagonism of chromatin modulators regulates epithelial-mesenchymal transition. *Sci Adv* **7** (2021). https://doi.org:10.1126/sciadv.abd7974

54 Shyam, S., Ramu, S., Sehgal, M. & Jolly, M. K. A systems-level analysis of the mutually antagonistic roles of RKIP and BACH1 in dynamics of cancer cell plasticity. *J R Soc Interface* **20**, 20230389 (2023). https://doi.org:10.1098/rsif.2023.0389

55 Luond, F. *et al.* Distinct contributions of partial and full EMT to breast cancer malignancy. *Dev Cell* **56**, 3203-3221 e3211 (2021). https://doi.org:10.1016/j.devcel.2021.11.006

56 Konkimalla, A. *et al.* Transitional cell states sculpt tissue topology during lung regeneration. *Cell Stem Cell* **30**, 1486-1502 e1489 (2023). https://doi.org:10.1016/j.stem.2023.10.001

57 Liu, Y., Fan, X., Jiang, C. & Xu, S. SPOCK2 and SPRED1 function downstream of EZH2 to impede the malignant progression of lung adenocarcinoma in vitro and in vivo. *Hum Cell* **36**, 812-821 (2023). https://doi.org:10.1007/s13577-023-00855-0

58 Kulshrestha, R. *et al.* Caveolin-1 as a critical component in the pathogenesis of lung fibrosis of different etiology: Evidences and mechanisms. *Exp Mol Pathol* **111**, 104315 (2019). https://doi.org:10.1016/j.yexmp.2019.104315

59 Sun, L. *et al.* Exosomal miRNA Let-7 from Menstrual Blood-Derived Endometrial Stem Cells Alleviates Pulmonary Fibrosis through Regulating Mitochondrial DNA Damage. *Oxid Med Cell Longev* **2019**, 4506303 (2019). https://doi.org:10.1155/2019/4506303

60 Huleihel, L. *et al.* Let-7d microRNA affects mesenchymal phenotypic properties of lung fibroblasts. *Am J Physiol Lung Cell Mol Physiol* **306**, L534-542 (2014). https://doi.org:10.1152/ajplung.00149.2013

61 Duan, R., Du, W. & Guo, W. EZH2: a novel target for cancer treatment. *J Hematol Oncol* **13**, 104 (2020). https://doi.org:10.1186/s13045-020-00937-8

62 Bao, X. *et al.* Inhibition of EZH2 prevents acute respiratory distress syndrome (ARDS)-associated pulmonary fibrosis by regulating the macrophage polarization phenotype. *Respir Res* **22**, 194 (2021). https://doi.org:10.1186/s12931-021-01785-x

63 McGovern, T. K., Robichaud, A., Fereydoonzad, L., Schuessler, T. F. & Martin, J. G. Evaluation of respiratory system mechanics in mice using the forced oscillation technique. *J Vis Exp*, e50172 (2013). https://doi.org:10.3791/50172

64 Chen, Q. & Liu, Y. Isolation and culture of mouse alveolar type II cells to study type II to type I cell differentiation. *STAR Protoc* **2**, 100241 (2021). https://doi.org:10.1016/j.xpro.2020.100241

65 Katsura, H. *et al.* Human Lung Stem Cell-Based Alveolospheres Provide Insights into SARS-CoV-2-Mediated Interferon Responses and Pneumocyte Dysfunction. *Cell Stem Cell* **27**, 890-904 e898 (2020). https://doi.org:10.1016/j.stem.2020.10.005

66 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016). https://doi.org:10.1093/nar/gkw377

67 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019). https://doi.org:10.1038/s41467-019-09234-6

68 Ochsner, S. A. *et al.* The Signaling Pathways Project, an integrated 'omics knowledgebase for mammalian cellular signaling pathways. *Sci Data* **6**, 252 (2019). https://doi.org:10.1038/s41597-019-0193-4

69 Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**, 1006-1019 (2018). https://doi.org:10.1038/nprot.2018.015

70 Le, D. T. *et al.* BATF2 promotes HSC myeloid differentiation by amplifying IFN response mediators during chronic infection. *iScience* **26**, 106059 (2023). https://doi.org:10.1016/j.isci.2023.106059

71 Trevino, L. S. *et al.* Epigenome environment interactions accelerate epigenomic aging and unlock metabolically restricted epigenetic reprogramming in adulthood. *Nat Commun* **11**, 2316 (2020). https://doi.org:10.1038/s41467-020-15847-z

72 McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501 (2010). https://doi.org:10.1038/nbt.1630

73 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010). https://doi.org:10.1016/j.molcel.2010.05.004

74 Zheng, R. *et al.* Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res* **47**, D729-D735 (2019). https://doi.org:10.1093/nar/gky1094

75 van Dongen, S., Abreu-Goodger, C. & Enright, A. J. Detecting microRNA binding and siRNA off-target effects from expression data. *Nat Methods* **5**, 1023-1025 (2008). https://doi.org:10.1038/nmeth.1267

76 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, 495-502 (2015). https://doi.org:10.1038/nbt.3192

77 Saetta, M. *et al.* Destructive index: a measurement of lung parenchymal destruction in smokers. *Am Rev Respir Dis* **131**, 764-769 (1985). https://doi.org:10.1164/arrd.1985.131.5.764

78 Hubner, R. H. *et al.* Standardized quantification of pulmonary fibrosis in histological samples. *Biotechniques* **44**, 507-511, 514-507 (2008). https://doi.org:10.2144/000112729

79 Crowley, G. *et al*. Quantitative lung morphology: semi-automated measurement of mean linear intercept. *BMC Pulm Med* **19**, 206 (2019). https://doi.org:10.1186/s12890-019-0915-6

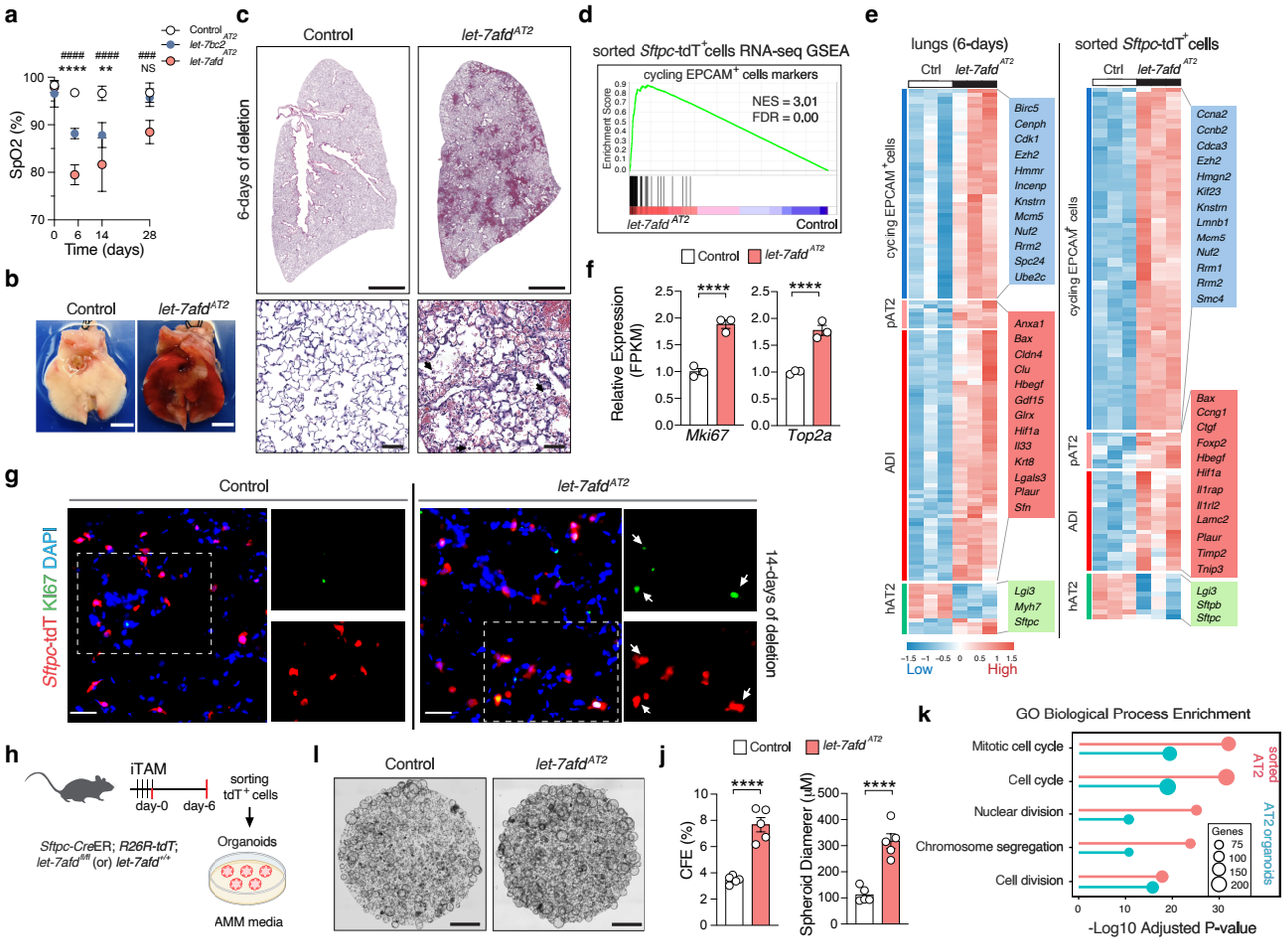

Fig. 1 Deletion of *let-7afd* in AT2 cells promotes spontaneous lung injury and AT2 transitional cells acutely. **a** Quantitative measurements of Sp02 from control, *let-7bc2*^{AT2}, *let-7afd*^{AT2} mice (n = 5 per group) obtained by pulse oxymetry following iTAM administration. Data are mean±s.d.**let-7bc2*^{AT2} vs. control, $^{\#}$*let-7afd*^{AT2} vs control; **, p < 0.01, $^{\#\#\#}$, p < 0.001, **** and $^{\#\#\#\#}$, p < 0.0001 by one way ANOVA with Tukey's correction. not significant (NS). **b**-**c** Representative gross whole lung dissections (**b**) and H&E-stained lung lobe sections (**c**) from indicated mice collected 6-days after iTAM (n = 16 mice per group). Scale bars: 5 mm (**b**); 2 mm upper panel; 50 μm lower panel (**c**). Arrowheads point to immune cells. **d** GSEA plot shows induction of proliferative/cycling epithelial genes following deletion of *let-7afd* in sorted *Sftpc*-tdT^{+} labeled AT2 cells (n = 3 samples per group from pools of 3 mice). **e** Heatmap shows differentially expressed AT2 transition markers obtained by bulk RNA-seq from whole lung (n = 3 mice per group) or sorted *Sftpc*-tdT^{+} labeled AT2 cells upon deletion of *let-7afd* vs controls after 6 days of iTAM. Adjusted p value < 0.05 vs control. **f** Transcript expression for *Mki67* and *Top2a* in *Sftpc*-tdT^{+} labeled AT2 cells obtained by RNA-seq after 6 days of iTAM (n = 3 samples per group). Data are mean±s.e.m. ****adjusted p < 0.0001. **g** Representative immunostainingdetection of KI67^{+}*Sftpc*-tdT^{+} cells in lungs of *let-7afd*^{AT2} mice compared to control mice on day 14 after iTAM (n = 4 mice per group). KI67 (green), *Sftpc*-tdT (red), DAPI (blue). White arrows indicate KI67^{+}*Sftpc*-tdT^{+} cells (scale bar 25 μm). **h** Schematic representation to establish AT2 organoid cultures on day 6 after iTAM from control and *let-7afd*^{AT2} mice. **i** Brightfield images show *let-7afd*$^{-/-}$ vs control AT2 organoids cultured for 14 days in AMM media (n = 5 mice per group). Scale bars 1 mm. **j** Quantification of CFE and spheroid diameters of organoids in panel (**i**). Data are mean±s.e.m (n = 5 mice per group). ****p < 0.0001 by unpaired Student's t test. **k** GO BP enrichment analysis for overrepresented terms associated with cell cycle following deletion of *let-7afd* in sorted *Sftpc*-tdT^{+} labelled AT2 cells (n= 3 samples per group) or organoid cultures (n = 2 mice per group) by bulk RNA-seq. Scale shows adjusted p values. **a**, **b**, **c**, **g**, **i**, **j** are representative of three independent experiments.

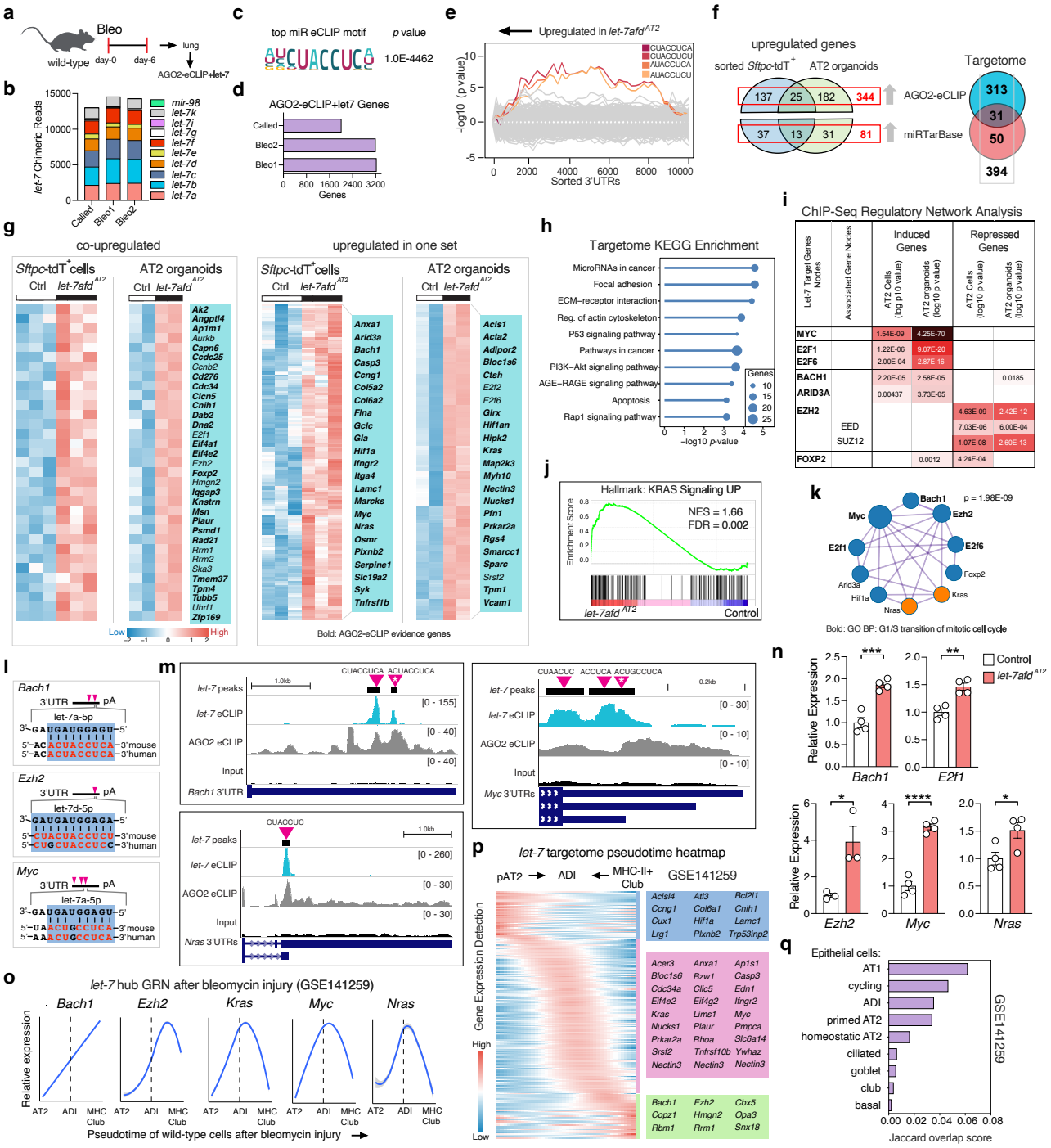
a
Bleo
wild-type
day-0
day-6
lung
AGO2-eCLIP+let-7
b
let-7 Chimeric Reads
15000
10000
5000
0
Called
Bleo1
Bleo2
mir-98
let-7k
let-7i
let-7g
let-7f
let-7e
let-7d
let-7c
let-7b
let-7a
c
top miR eCLIP motif
p value
1.0E-4462
d
AGO2-eCLIP+let7 Genes
Called
Bleo2
Bleo1
0 800 1600 2400 3200
Genes
e
Upregulated in let-7afd AT2
CUACCUCA
CUACCUCU
AUACCUCA
AUACCUCU
-log10 (p value)
Sorted 3'UTRs
f
upregulated genes
sorted Sftpc-tdT+
AT2 organoids
137
25
182
344
AGO2-eCLIP
37
13
31
81
miRTarBase
Targetome
313
31
50
394
g
co-upregulated
upregulated in one set
Sftpc-tdT+cells
AT2 organoids
Ctrl
let-7afd AT2
Ak2
Angptl4
Ap1m1
Aurkb
Capn6
Ccdc25
Ccnb2
Cd276
Cdc34
Clcn5
Cnih1
Dab2
Dna2
E2f1
Eif4a1
Eif4e2
Ezh2
Foxp2
Hmgn2
Iqgap3
Knstrn
Msn
Plaur
Psmd1
Rad21
Rrm1
Rrm2
Ska3
Tmem37
Tpm4
Tubb5
Uhrf1
Zfp169
Low
High
Anxa1
Arid3a
Bach1
Casp3
Ccng1
Col5a2
Col6a2
Flna
Gclc
Gla
Hif1a
Ifngr2
Itga4
Lamc1
Marcks
Myc
Nras
Osmr
Plxnb2
Serpine1
Slc19a2
Syk
Tnfrsf1b
Acls1
Acta2
Adipor2
Bloc1s6
Ctsh
E2f2
E2f6
Glrx
Hif1an
Hipk2
Kras
Map2k3
Myh10
Nectin3
Nucks1
Pfn1
Prkar2a
Rgs4
Smarcc1
Sparc
Srsf2
Tpm1
Vcam1
Bold: AGO2-eCLIP evidence genes
h
Targetome KEGG Enrichment
MicroRNAs in cancer
Focal adhesion
ECM–receptor interaction
Reg. of actin cytoskeleton
P53 signaling pathway
Pathways in cancer
PI3K–Akt signaling pathway
AGE–RAGE signaling pathway
Apoptosis
Rap1 signaling pathway
Genes
10
15
20
25
–log10 p-value
i
ChIP-Seq Regulatory Network Analysis
Let-7 Target Genes Nodes
Associated Gene Nodes
Induced Genes
AT2 Cells (log p10 value)
AT2 organoids (log10 p value)
Repressed Genes
AT2 Cells (log10 p value)
AT2 organoids (log10 p value)
MYC 1.54E-09 4.25E-70
E2F1 1.22E-06 9.07E-20
E2F6 2.00E-04 2.87E-16
BACH1 2.20E-05 2.58E-05 0.0185
ARID3A 0.00437 3.73E-05
EZH2 4.63E-09 2.42E-12
EED 7.03E-06 6.00E-04
SUZ12 1.07E-08 2.60E-13
FOXP2 0.0012 4.24E-04
j
Hallmark: KRAS Signaling UP
Enrichment Score
NES = 1.66
FDR = 0.002
let-7afd AT2
Control
k
Bach1
Myc
Ezh2
E2f1
E2f6
Foxp2
Arid3a
Hif1a
Kras
Nras
p = 1.98E-09
Bold: GO BP: G1/S transition of mitotic cell cycle
l
Bach1
3'UTR
pA
let-7a-5p
3'-GAUGAUGGAGU-5'
5'-ACACUACCUCA-3'mouse
5'-ACACUACCUCA-3'human
Ezh2
let-7d-5p
3'-GAUGAUGGAGA-5'
5'-CUACUACCUCU-3'mouse
5'-CUGCUACCUCC-3'human
Myc
let-7a-5p
3'-GAUGAUGGAGU-5'
5'-UAACUGCCUCA-3'mouse
5'-AAACUGCCUCA-3'human
m
let-7 peaks
let-7 eCLIP
AGO2 eCLIP
Input
Bach1 3'UTR
Nras 3'UTRs
CUACCUCA
ACUACCUCA
CUACCUC
ACCUCA
ACUGCCUCA
1.0kb
0.2kb
[0 - 155]
[0 - 40]
[0 - 30]
[0 - 10]
[0 - 260]
n
Relative Expression
Control
let-7afd AT2
Bach1
E2f1
Ezh2
Myc
Nras
o
let-7 hub GRN after bleomycin injury (GSE141259)
Bach1
Ezh2
Kras
Myc
Nras
Relative expression
AT2
ADI
MHC Club
Pseudotime of wild-type cells after bleomycin injury
p
let-7 targetome pseudotime heatmap
pAT2
ADI
MHC-II+ Club
GSE141259
Gene Expression Detection
High
Low
Acsl4 Atl3 Bcl2l1
Ccng1 Col6a1 Cnih1
Cux1 Hif1a Lamc1
Lrg1 Plxnb2 Trp53inp2
Acer3 Anxa1 Ap1s1
Bloc1s6 Bzw1 Casp3
Cdc34a Clic5 Edn1
Eif4e2 Eif4g2 Ifngr2
Kras Lims1 Myc
Nucks1 Plaur Pmpca
Prkar2a Rhoa Slc6a14
Srsf2 Tnfrsf10b Ywhaz
Nectin3 Nectin3 Nectin3
Bach1 Ezh2 Cbx5
Copz1 Hmgn2 Opa3
Rbm1 Rrm1 Snx18
q
Epithelial cells:
AT1
cycling
ADI
primed AT2
homeostatic AT2
ciliated
goblet
club
basal
GSE141259
Jaccard overlap score

Fig. 2 Chimeric AGO2 eCLIP and transcriptomics identify *let-7* targetome in AT2 cells. **a** Schematic of Experimental design for AGO2-eCLIP+let-7 analysis from lungs of mice after bleomycin treatment (n = 2 mice). **b** Fractional representation of *let-7*:mRNA chimeric reads for individual *let-7* family members. **c** The *let-7* family binding motif is highly enriched in corresponding *let-7*:mRNA chimeric read regions. **d** Total number of genes with *let-7*:mRNA chimeric peaks in 3'UTR. **e** Sylamer miRNA transcriptome enrichment analysis shows p value overrepresentation of *let-7* binding motifs in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells. **f** Venn Diagram shows gene totals and integration of AGO2-eCLIP+let-7 with miRTarbase and *let-7afd*$^{-/-}$ RNA-seq datasets to obtain the *let-7* AT2 cell targetome. **g** Heatmaps shows relative expression of significantly upregulated and validated *let-7* target genes upon loss of *let-7afd* in *Sftpc*-tdT^{+} AT2 cells (n= 3 samples per group) or AT2 organoids (n = 2 mice per group) bulk RNA-seq datasets. Adjusted p value < 0.05 vs controls. **h** KEGG pathway enrichment analysis for the *let-7* targetome. **i** ChIP-Seq cistromic analysis indicate significant enrichment of nodes corresponding to the hub targets of *let-7* on upregulated or downregulated genes after depletion of *let-7afd*$^{-/-}$ in sorted *Sftpc*-tdT^{+} cells and organoids. **j** GSEA plot predicts transcriptome activation of KRAS pathway in *let-7afd*$^{-/-}$ vs control *Sftpc*-tdT^{+} cells. **k** Functional pathway and protein interactome network indicates significant connectivity of hub targets of *let-7*. TFs (blue), enzymes (orange). P value obtained by Cytoscape/Enrichment Map. **l** Schematic alignments of *let-7* "seed" region with target mRNA sequences in mice and humans. **m** UCSC genome browser tracks from AGO2-eCLIP+let-7 indicates binding of *let-7* family to the 3'UTR of hub genes. Purple triangles indicate binding motifs of *let-7*. White asterisks indicate sequence alignment in panel (**l**). **n** qPCR analysis of indicated *let-7* targets from *let-7afd*$^{-/-}$ vs control sorted *Sftpc*-tdT^{+} cells on day 14 day after iTAM. (n = 3 per group for *Ezh2*; for rest n = 4 per group; from pools of 2 mice). Data are mean±s.e.m. ****p < 0.0001, ***p < 0.001, **p < 0.01, *p < 0.05 by unpaired Student's t test. **o** Line plots show smoothed relative expression levels of hub genes across the ADI pseudotime trajectories after bleomycin injury (GSE141259). Dashed line corresponds to *Krt8* peak and gray colors represent the 95% confidence interval derived from smoothing fit. **p** The expression pattern for 217 of 394 *let-7* targetome genes along the AT2 or MHC-II^{+} club cell trajectory into ADIs is based on inferred likelihood of detection in the heatmap. **q** Jaccard overlap scores were calculated between *let-7* targetome and epithelial cell types defined in GSE141259.

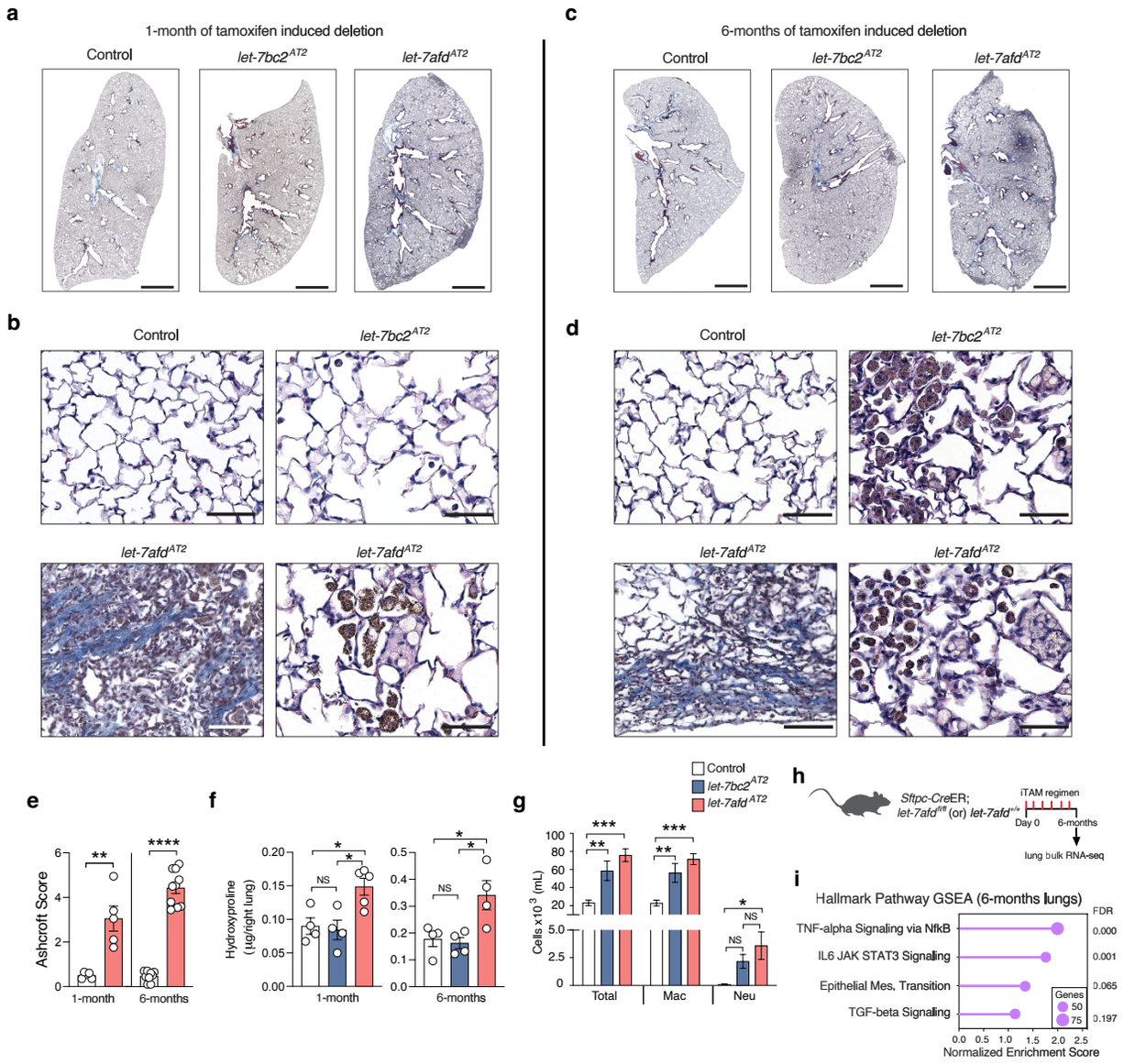

Fig. 3 AT2-specific deletion of *let-7* clusters induce spontaneous progressive chronic inflammation and ILD. **a-d** Representative Masson's trichrome-stained sections of lung lobes at 1- (**a**,**b**) or 6-months (**c**,**d**) post-iTAM. Scale bars: 2 mm upper panels; 50 µm lower panels (n = 16 mice per group). **e** Ashcroft score was used to evaluate lung injury after iTAM. (1-month iTAM, n = 5; 6-month iTAM, n = 10; mice per group). Data are mean±s.e.m. ****$p < 0.0001$,**$p < 0.01$, by unpaired Student's t test. **f** Hydroxyproline levels from lungs of control, *let-7bc2*^{AT2}, *let-7afd*^{AT2} mice at 1-month (n = 4-5 mice per group) or 6-months of iTAM (4 mice per group). Data are mean±s.e.m.*$p < 0.05$ by one way ANOVA with Tukey's correction. **g** Differential cell counts from BALF of mice is shown. Data are mean±s.e.m from individual mice (n = 7-8 mice per group). ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, by one way ANOVA with Tukey's correction. Macrophages (Mac), Neutrophils (Neu). **h** Schematic of lung bulk RNA-seq obtained from *let-7afd*^{AT2} and control mice at 6 months of iTAM (n = 4 mice per group). **i** GSEA Normalized enrichment scores and adjusted p values of significantly induced gene sets in lungs of *let-7afd*^{AT2} mice vs controls. **e**, **f** each dot represents a mouse. **a-g** are representative of three independent experiments. Not significant (NS).

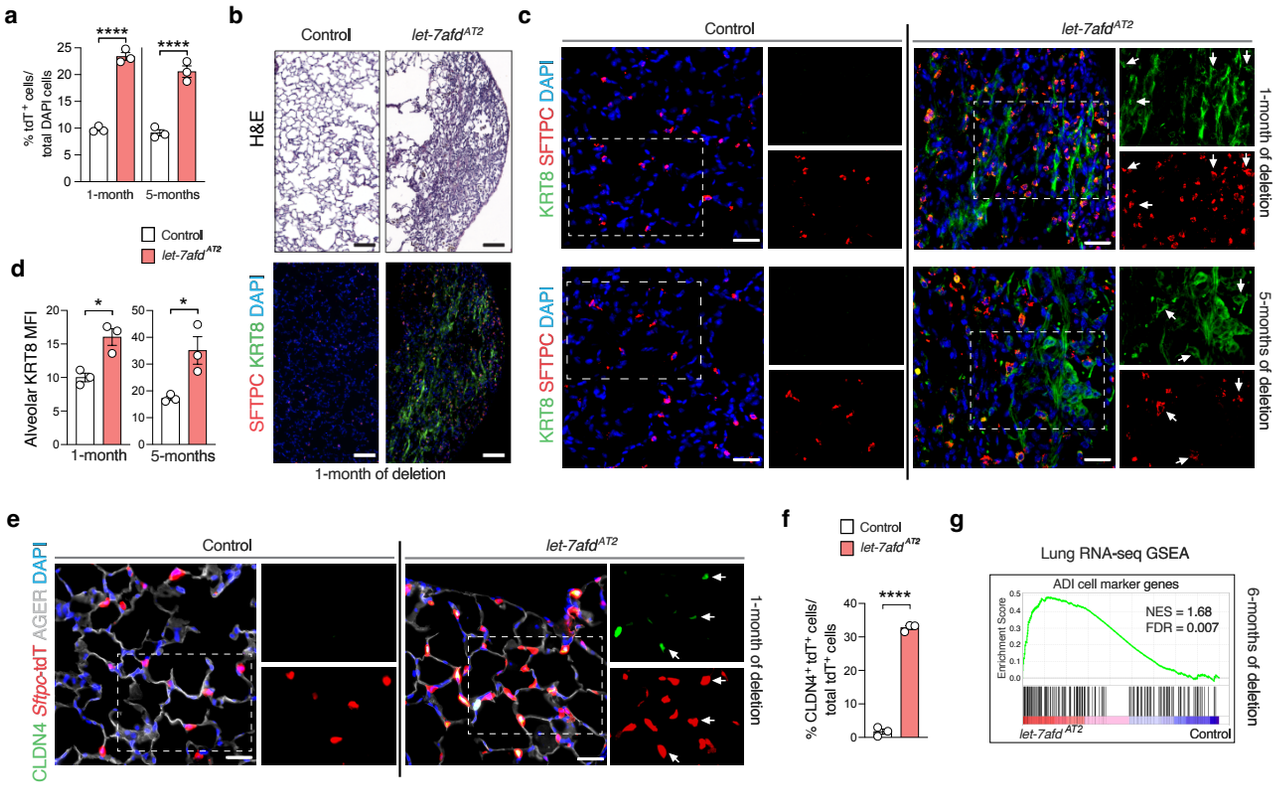

Fig 4 Ablation of *let-7afd* in AT2 cells stimulates the persistence of ADI intermediate cells in the lung. **a** Quantification of *Sftpc*-tdT$^+$ cells from total DAPI cells in alveolar regions of *let-7afd*^{AT2} mice compared to control mice at 1- or 5-months of iTAM induced deletion (n = 3 mice per group). Data are mean±s.e.m. ****$p < 0.0001$ by unpaired Student's t test. **b** Representative H&E-stained (upper panels) and immunostaining (lower panels) serial sections of lung are shown after 1 month of iTAM in *let-7afd*^{AT2} mice compared to control mice (n = 12 mice per group). IF shows KRT8 (green), SFTPC (red), DAPI (blue). Scale bars: 50 μm. **c** Representative immunostaining of KRT8 (green), SFTPC (red), DAPI (blue) at 1 month (upper panels) or 5 months of iTAM (lower panels) in *let-7afd*^{AT2} mice vs controls. Arrowheads point to KRT8$^+$SFTPC$^+$ cells. Scale bars: 25μm. **d** Quantification of KRT8 mean fluorescence intensity (MFI) from alveolar regions excluding airways of mice groups in panel (**c**) (n = 3 mice per group). Data are mean±s.e.m. *$p < 0.05$ by unpaired Student's t test. **e**,**f** Representative immunostaining (**e**) or quantification of *Sftpc*-tdT$^+$CLDN4$^+$ from total *Sftpc*-TdT$^+$ cells (**f**) in lungs of *let-7afd*^{AT2} mice compared to controls 1-month post-iTAM (n = 3 mice per group). CLDN4 (green), *Sftpc*-tdT (red), AGER (grey), DAPI (blue). Arrowheads point to *Sftpc*-tdT$^+$CLDN4$^+$ cells. Data are mean±s.e.m. ****$p < 0.0001$, by unpaired Student's t test. Scale bars: 25 μm. **g** GSEA plot shows induction of ADI genes in lungs of *let-7afd*^{AT2} mice vs controls following 6 months of iTAM. **a**, **b**, **c**, **d**, **e**, **f** are representative of three independent experiments.

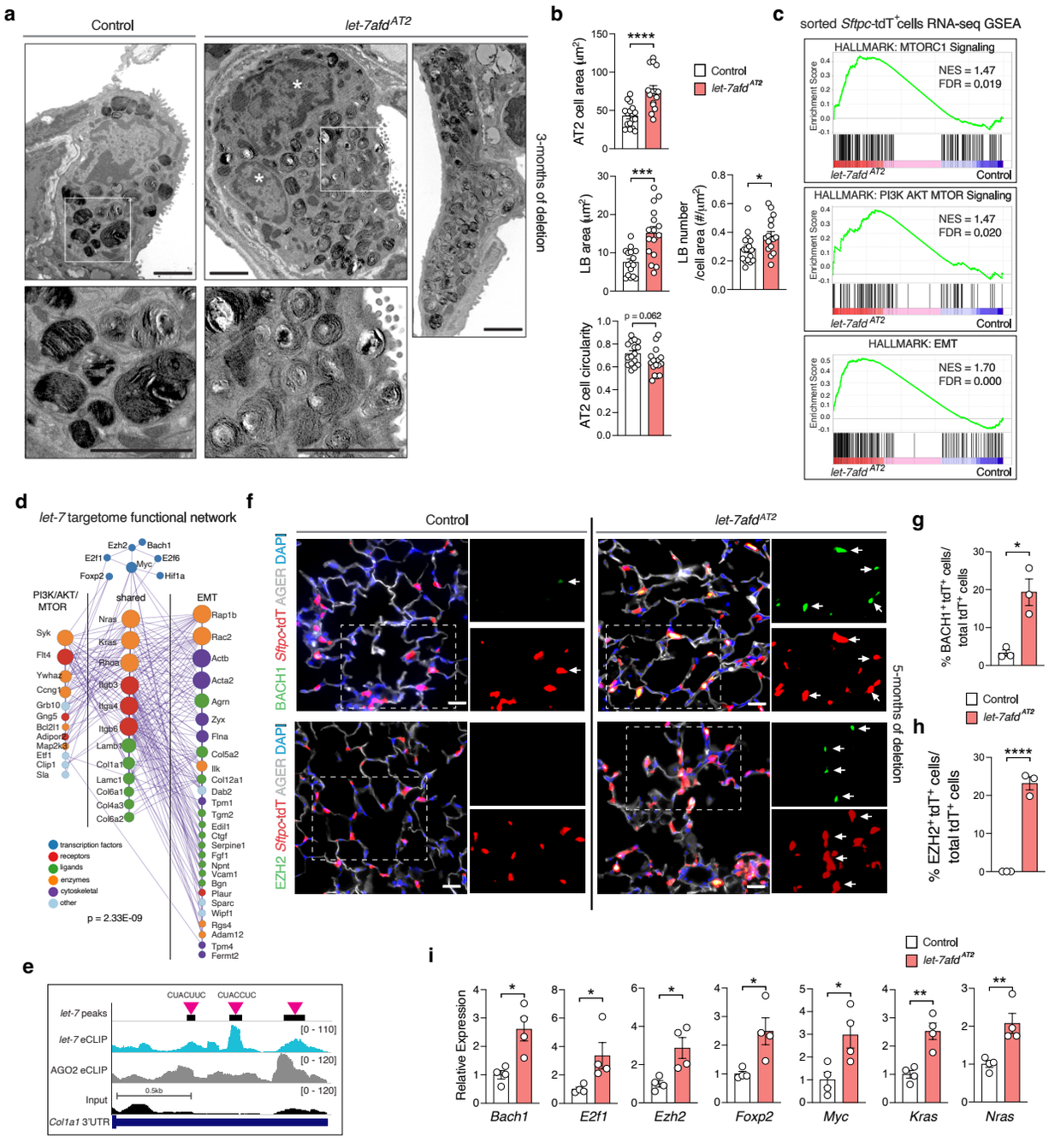

Fig 5 Loss of *let-7afd* enhances the size of AT2 cells and induces the expression of *let-7* target genes associated with promotion of ECM remodeling and fibrosis. **a** Representative TEM image shows an enlarged *let-7afd*$^{-/-}$ AT2 cell with increased abnormal lamellar bodies (LB)s and more than one nuclei (n = 3 mice per group). The white asterisks point to individual nuclei. An elongated *let-7afd*$^{-/-}$ AT2 cell is also shown on right. Scale bars: 2 μm. **b** TEM quantification of cell area and LB cell area (μm^2), LB number per μm^2 of cell area, and cell circularity index in each AT2 cell. Data are mean±s.e.m. Each dot represents one AT2 cell (control AT2 cells n = 16; *let-7afd*$^{-/-}$AT2 cells = 16, from a total of 3 mice per group; ****$p < 0.0001$, ***$p < 0.001$, *$p < 0.05$, and $p = 0.062$ by unpaired Student's t test. **c** GSEA plots indicate induction of PI3K/AKT/MTOR and EMT Hallmark pathways in *let-7afd*$^{-/-}$ *Sftpc*-tdT$^+$ cells vs control *Sftpc*-tdT$^+$ cells following 6 days of iTAM (n = 3 mice per group). **d** Functional interactome network of *let-7* target genes associated with PI3K/AKT/MTOR and EMT. Network p value was computed by Cytoscape/Enrichment Map. **e** UCSC genome browser track indicates binding of *let-7* to the 3'UTR of *Col1a1* gene. **f** Representative IF images show increased numbers of BACH1$^+$*Sftpc*-tdT$^+$ (top panels) and EZH2$^+$*Sftpc*-tdT$^+$ (lower panels) cells (indicated by white arrows) in lungs of *let-7afd*^{AT2} mice compared to control mice lungs after 5 months of iTAM. BACH1 (green); EZH2 (green); *Sftpc*-tdT (red); AGER (gray); DAPI (blue). Scale bars: 25 μm. **g-h** Quantification of BACH1$^+$*Sftpc*-tdT$^+$ cells (**g**) and EZH2$^+$*Sftpc*-tdT$^+$ cell (**h**) in total *Sftpc*-tdT$^+$ cells (n = 3 mice per group). Data are mean±s.e.m. ****$p < 0.0001$, *$p < 0.05$ by unpaired Student's t test. **i** *Bach1*, *E2f1*, *Ezh2*, *Foxp2*, *Myc*, *Kras*, and *Nras* mRNA expression levels were evaluated by qPCR 2 months post-iTAM from *let-7afd*$^{-/-}$ and control *Sftpc*-tdT$^+$ cells (n = 4 samples per group; from pools of 2 mice). Data are mean±s.e.m. **$p < 0.01$, *$p < 0.05$, by unpaired Student's t test. **f**, **g**, **h**, **i** are representative of three independent experiments.

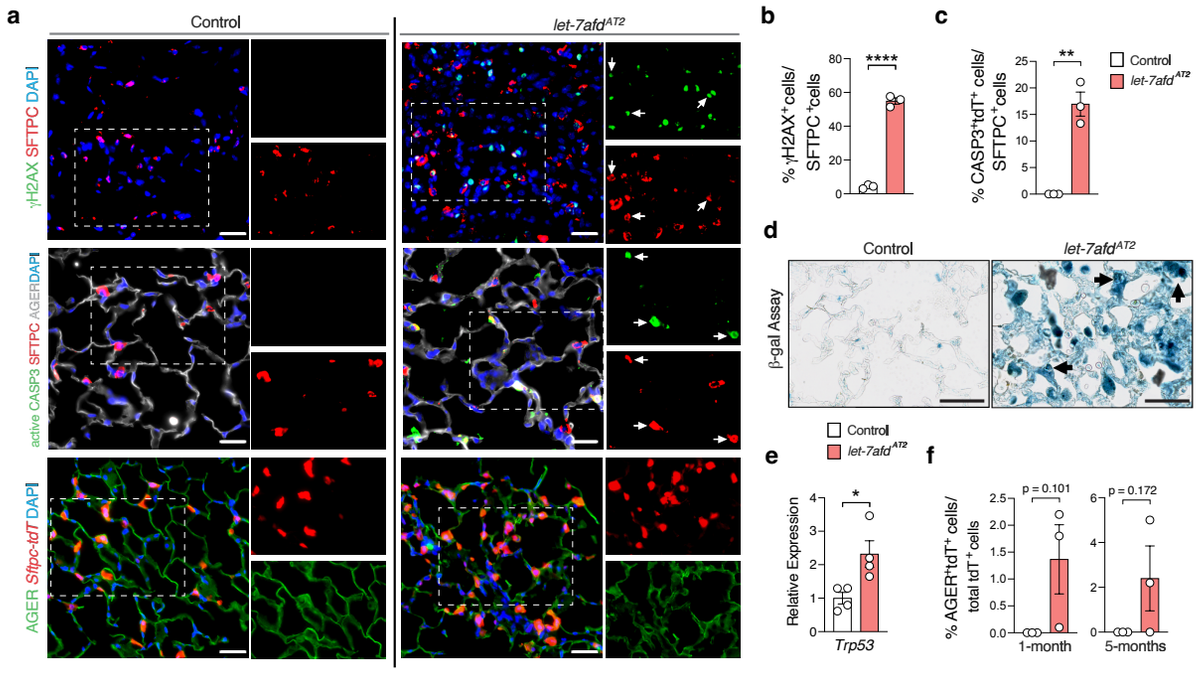

Fig 6 Loss of *let-7afd* is associated with AT2 cell DNA damage, senescence, and apoptosis. **a** Immunostaining for γH2AX (green), SFTPC (red), DAPI (blue) upper panels; active CASP3 (green), SFTPC (red), AGER (gray), DAPI (blue) middle panels; AGER (green), *Sftpc*-tdT (red), DAPI (blue) lower panels in lungs of *let-7afd*^{AT2} and control mice after 1 month of iTAM. Arrows indicate γH2AX$^+$SFTPC$^+$ cells (upper panel) or active-CASP3$^+$SFTPC$^+$ cells (middle panels). Scale bars: 25 μm. **b-c** Quantification of γH2AX$^+$SFTPC$^+$ cells (**b**) or active-CASP3$^+$SFTPC$^+$ cells (**c**) in total SFTPC$^+$ cells (n = 3 mice per group). Data are mean±s.e.m. ****$p < 0.0001$, **$p < 0.01$, by unpaired Student's t test. **d** Representative β-galactosidase staining in lungs of *let-7afd*^{AT2} and control mice after 1-month of iTAM. Arrowheads point to AT2 cells in thickened alveolar septa (n = 6 mice per group). Scale bars: 50 μm. **e** qPCR detection of *Trp53* from *let-7afd*$^{-/-}$ and control *Sftpc*-tdT$^+$ cells after 2 months of iTAM (n = 4 samples per group from pools of 2 mice). Data are mean±s.e.m. *$p < 0.05$, by unpaired Student's t test. **f** Quantification of AGER$^+$*Sftpc*-tdT$^+$ in total *Sftpc*-tdT$^+$ cells from lungs of *let-7afd*^{AT2} and control mice at 1- or 5-months after iTAM. Data are mean±s.e.m, by unpaired Student's t test. All panels are representative of three independent experiments.

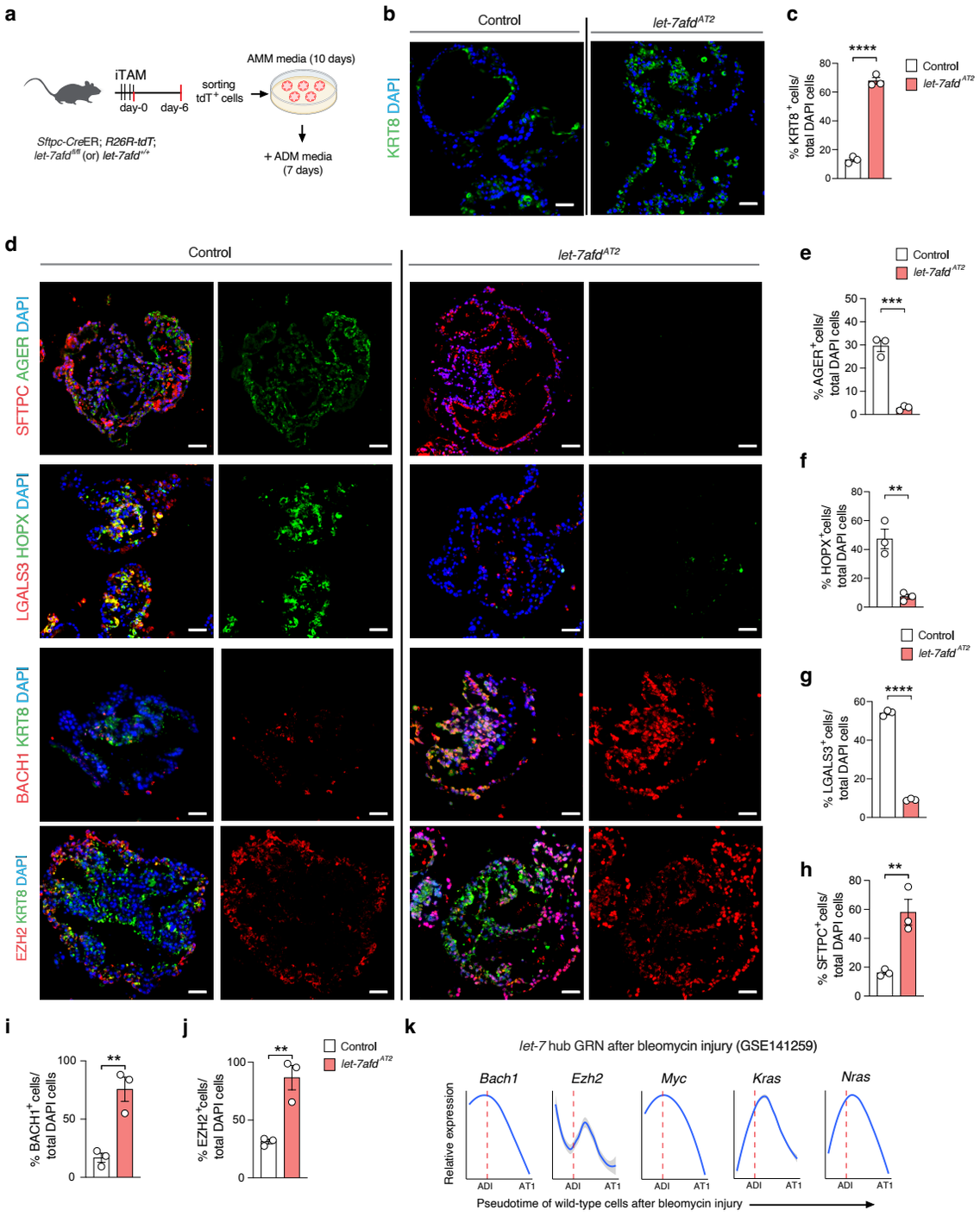

Fig 7 Absence of *let-7afd* impairs AT2 to AT1 differentiation. **a** Schematic representation of experimental design to examine the role of *let-7afd* in AT2 to AT1 differentiation in cultured AT2 cells. **b**,**d** Representative immunostaining of ADI, AT1 markers and BACH1/EZH2 expression in *let-7afd*$^{-/-}$ vs control AT2 organoids grown as in panel (**a**) (n = 3 mice per group). Panel (**b**) shows KRT8 (green), DAPI (blue). Panel (**d**) shows AGER (green), SFTPC (red), DAPI (blue) top row panels; HOPX (green); LGALS3 (red); DAPI (blue) second row panels; BACH1 (red), KRT8 (green), DAPI (blue) third row panels; EZH2 (red), KRT8 (green), DAPI (blue) bottom row panels. Scale bars: 25 μm. **e-j** Quantification of AGER$^+$ (**e**), HOPX$^+$ (**f**), LGALS$^+$ (**g**), SFTPC$^+$ (**h**), BACH1$^+$ (**i**), or EZH2$^+$ (**j**) cells respectively in total DAPI cells (n = 3 mice per group). Data are mean±s.e.m. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, by unpaired Student's t test. Each dot represents a mouse. Representative of three independent experiments. **k** Line plots show smoothed expression levels of selected *let-7* hub targets in the ADI to AT1 trajectory obtained from GSE141259 dataset. Grey line represents 95% confidence interval obtained from fit model. KRT8 peak represented by dashed line.

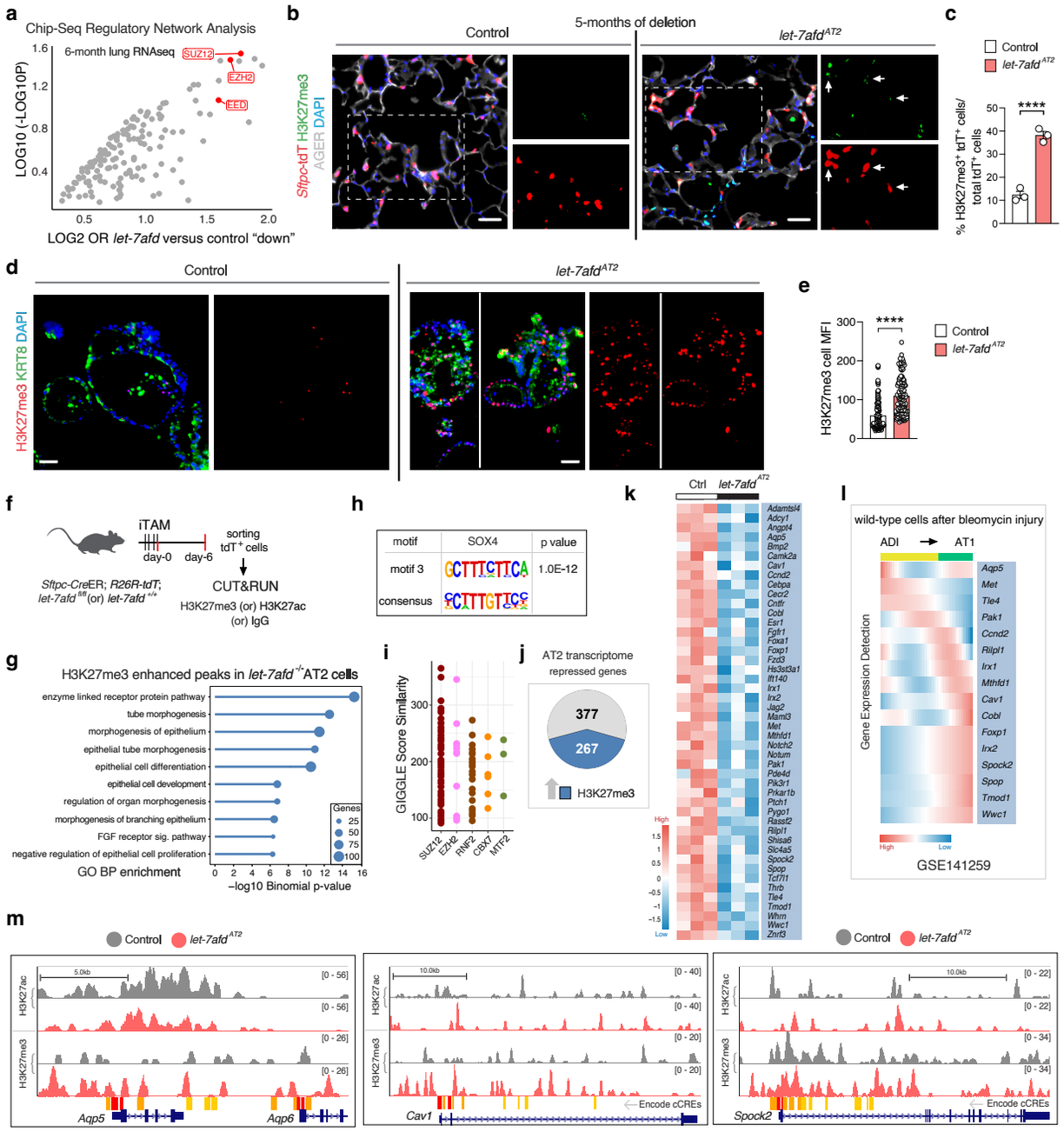

Fig 8 Loss of *Let-7afd* contributes to H3K27me3 mediated disruption of AT2 to AT1 differentiation. **a** ChIP-Seq regulatory network analysis on lung bulk RNA-seq dataset indicates significant representation of EZH2, EED, SUZ12 target genes in downregulated genes from *let-7afd*^{AT2} compared to control mice lungs at 6-months post-iTAM (n = 4 mice per group). Statistics were done with hypergeometric test. **b-c** Representative immunostaining (**b**) or quantification of H3K27me3^{+}*Sftpc*-tdT^{+} cells in total *Sftpc*-tdT^{+} cells (c) in lungs of *let-7afd*^{AT2} mice compared to controls at 5 months of iTAM (n = 3 mice). Scale bars: 25 μm. H3K27me3 (green), *Sftpc*-tdT (red), AGER (gray), DAPI (blue). Data are mean±s.e.m. ****$p < 0.0001$ by unpaired Student's t test. **d** Immunostaining for H3K27me3 in *let-7afd*$^{-/-}$ vs control AT2 organoids cultured in AMM for 10 days followed by ADM media for 7 days (n = 3 mice). H3K27me3 (red), KRT8 (green), DAPI (blue). Scale bars: 25 μm. **e** Quantification of H3K27me3 mean fluorescence intensity (MFI) in nuclei of AT2 organoids cultured as in panel (**d**) (n = 149 cells each group). Data are mean±s.e.m. Each dot represents single cell from a total of 3 mice per group. ****$p < 0.0001$ by unpaired Student's t test. **f** Experimental design scheme to carry out CUT&RUN from sorted *Sftpc*-tdT^{+} cells obtained from *let-7afd*^{AT2} mice or control mice (n = 1 from a pool of 2 mice per group). **g** GO biological pathway (BP) enrichment for induced peaks in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells compared to control *Sftpc*-tdT^{+} cells. **h** Transcription factor motif enrichment and alignment between motif #3 and SOX4 consensus sequence. **i** Increased H3K27me3 peaks in *let-7afd*$^{-/-}$ sorted *Sftpc*-tdT^{+} cells shows similarity to SUZ12/EZH2 ChIP-Seq datasets. **j** Pie chart shows that from a total of 644 repressed genes identified by bulk RNA-seq, 267 genes (blue color) exhibit enhanced H3K27me3 peaks in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells. **k** Heatmap shows significantly downregulated genes involved in cell differentiation, cell identity, cell growth inhibition, and cell-cell communication with increased H3K27me3 peaks in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells. **l** Heatmap indicates gene expression patterns for a subset of genes shown in panel (**k**) across the ADI to AT1 differentiation trajectory after bleomycin injury (GSE141259). **m** UCSC tracks show enrichment of H3K27me3 peaks in *Aqp5*, *Cav1,* and *Spock2* in *let-7afd*$^{-/-}$ *Sftpc*-tdT^{+} cells compared to controls.

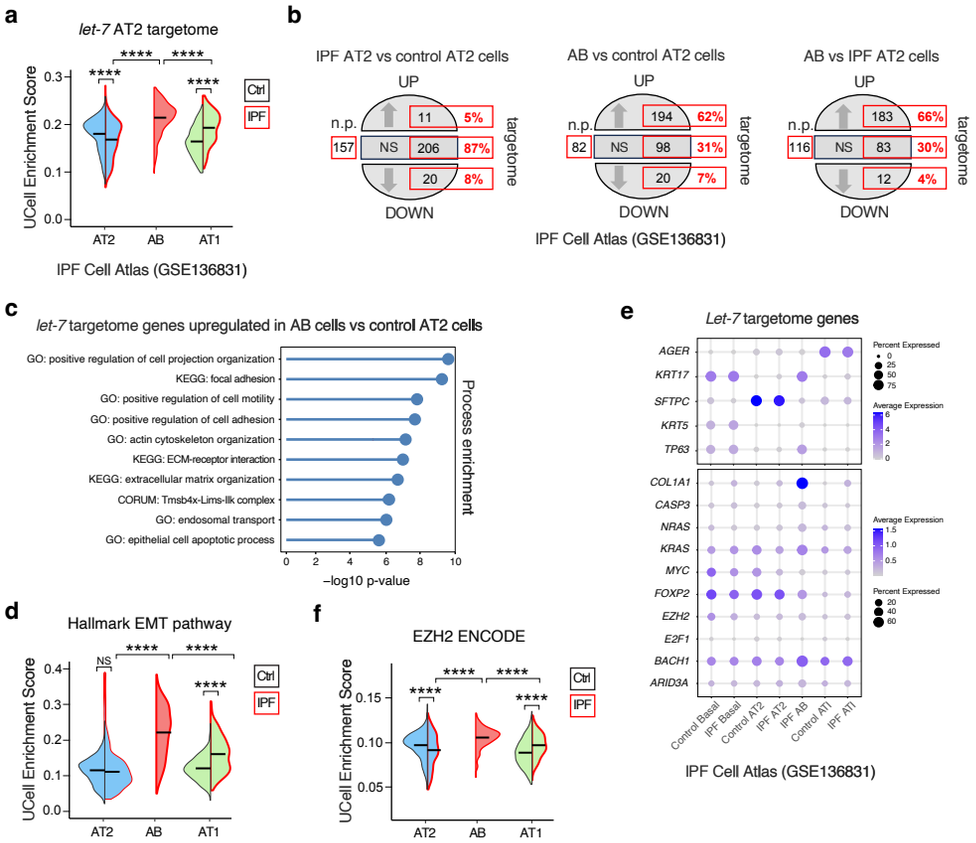

Fig 9 The *let-7* targetome is enriched in "aberrant basaloid" (AB) cells present in IPF patients. **a** UCell Enrichment scores for the murine *let-7* targetome in AT2, AB, and AT1 cells from control or IPF lungs. **b** Broken circle plots indicate the total number and proportions of *let-7* targets which are significantly upregulated (UP), downregulated (DOWN), or non significant (NS) between control (Ctrl) and IPF cells. Not present (n.p.) indicates gene targets missing in dataset GSE136831. **c** Gene process enrichment analysis for upregulated *let-7* targets in AB cells compared to control AT2 cells. **d** The Hallmark EMT pathway enrichment scores in control vs IPF AT2, AB and AT1 cell samples. **e** Dot plots show the relative expression of selected *let-7* targets and identity markers in AT2, AB, AT1 and basal cells from control and IPF patients. **f** ENCODE consensus targets of EZH2 were evaluated for UCell Enrichment Scores in AT2, AB and AT1 cells from control and IPF patients. **a**,**d**,**f** Data are mean±s.e.m half violin plots. ****$p < 0.0001$, by one way ANOVA with Sidak correction. Data was derived from published scRNAseq dataset GSE136831.