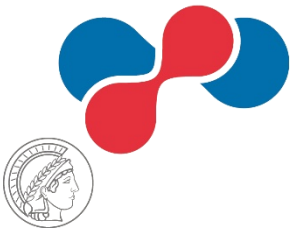
Max Planck Institute
for Heart and Lung Research
W.G. Kerckhoff Institute



JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

Transition-state, branch-point associated cells give rise to the majority of ciliated cells in the mouse lung

INAUGURAL-DISSERTATION
for the degree of Doctor rerum naturalium
(Dr. rer. nat.)

by

Keynoosh Khalooghi

submitted to the
Faculty of Biology and Chemistry
Justus-Liebig-University
Giessen, Germany

prepared in the
Department of Cardiac Remodeling
Max Planck Institute for Heart and Lung Research
Bad Nauheim, Germany

Bad Nauheim, 2022

PREFACE

The present work was carried out at the Max Planck Institute for Heart and Lung Research (W.G. Kerckhoff Institute) in Bad Nauheim.

First Reviewer: Prof. Dr. Dr. Thomas Braun

Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany

Second Reviewer: Prof. Dr. Soni Savai Pullamsetti

Justus-Liebig-University, Giessen, Germany

Declaration

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and referenced all text passages that are derived literally from or are based on the content of published or unpublished work of others, and all information that relates to verbal communications. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University of Giessen in carrying out the investigations described in the dissertation.

_____

Keynoosh Khalooghi

To “AS”

Table of contents

1 SUMMARY

The lung is constantly receiving environmental insults from pathogens in inhaled air, necessitating permanent repair processes. To acquire and maintain regional functionalities of the lung, several different cell types are needed. In the intralobar bronchioalveolar airways and at branch points, a subpopulation of Club cells exists, known as variant Club cell (^{v}Club), which express secretoglobin (Scgb1a1) but not cytochrome P450 (Cyp2f2). This subset is resistant to naphthalene and initiates repopulation of the airway at branch points in the vicinity of neuroepithelial bodies (NEBs). Due to the paucity of a unique molecular marker for ^{v}Club cells, the physiological role of these cells during both homeostasis and injury is poorly understood.

To identify a specific marker gene for ^{v}Club cells, 2000 individual Club cells were analyzed, which were isolated by Fluorescence activated cell sorting (FACS) from CCSP^{mCherry} SPC^{YFP} reporter knock-in mouse strains. Single cell RNA sequencing analysis identified a small subpopulation of Club cells defined by co-expression of CCSP and Foxj1, which showed hallmarks of ^{v}Club cells such as localization at branch points and resistance to naphthalene injury. Subsequent lineage tracing demonstrated that CCSP^{+} Foxj1^{+} cells give rise to almost 60% of all ciliated cells. Diphtheria toxin (DTA) mediated ablation of CCSP^{+} Foxj1^{+} cells reduced cellularity of the bronchiolar epithelium, causing a reduction in the number of ciliated cells and an even more substantial depletion of regular Club cells. The results indicate a high cellular plasticity of the airway epithelium during regeneration, allowing regular club cells to compensate for the absence of CCSP^{+} Foxj1^{+} cells. However, enforcement of ciliated cell formation by regular club cells comes with a price, resulting in exhaustion of club cells, which demonstrates some limitations of compensatory lung regeneration mechanisms.

2 ZUSAMMENFASSUNG

Die Lunge ist ständig Umweltbelastungen durch Krankheitserreger in der Atemluft ausgesetzt, was permanente Reparaturprozesse erforderlich macht. Um die regionale Funktionsfähigkeit der Lunge zu erreichen und aufrechtzuerhalten, werden verschiedene Zelltypen benötigt. In den intralobaren bronchoalveolären Atemwegen und an Verzweigungsstellen gibt es eine Subpopulation von Club Zellen, die als variante Club Zellen (^{V}Club) bezeichnet werden und Sekretoglobin (Scgb1a1), nicht aber Cytochrom P450 (Cyp2f2) exprimieren. Diese Untergruppe ist resistent gegen Naphthalen und initiiert die Neubesiedlung der Atemwege an Verzweigungspunkten in der Nähe der Neuroepithelkörperchen (= Neuroepithelial Bodies, NEBs). Da es keine eindeutigen molekularen Marker für ^{V}Club-Zellen gibt, ist die physiologische Rolle dieser Zellen sowohl in der Homöostase als auch bei Verletzungen nur unzureichend bekannt.

Um ein spezifisches Markergen für ^{V}Club-Zellen zu identifizieren, wurden 2000 einzelne Club Zellen analysiert, die durch fluoreszenzaktivierte Zellsortierung (= Fluorescence Activated Cell Sorting, FACS) aus CCSP^{mCherry} SPC^{YFP}-Reporter-Knock-in-Mausstämmen isoliert wurden. Die Einzelzell-RNA-Sequenzanalyse identifizierte eine kleine Subpopulation von Club Zellen, die durch die Koexpression von CCSP und Foxj1 definiert ist und Merkmale von ^{V}Club-Zellen wie Lokalisierung an Verzweigungspunkten und Resistenz gegen Naphthalinschäden aufweisen. Eine anschließende Abstammungsanalyse zeigte, dass aus CCSP^{+}Foxj1^{+}-Zellen fast 60 % aller Flimmerzellen hervorgehen. Die durch Diphtherietoxin (DTA) verursachte Ablation von CCSP^{+}Foxj1^{+}-Zellen reduzierte die Zellularität des bronchiolären Epithels, was zu einer Verringerung der Anzahl der Flimmerzellen und einer noch stärkeren Verarmung der regulären Club Zellen führte. Diese Ergebnisse deuten auf eine hohe zelluläre Plastizität des Atemwegsepithels während der Regeneration hin, die es regulären Club Zellen ermöglicht, das Fehlen von CCSP^{+}Foxj1^{+}-Zellen zu kompensieren. Die erzwungene Bildung von Flimmerzellen durch reguläre Club Zellen hat jedoch ihren Preis und führt zu einer Erschöpfung der Club Zellen, was die Grenzen kompensatorischer Lungenregenerationsmechanismen aufzeigt.

3 INTRODUCTION

Oxygen is crucial for survival of almost every living cell and the lungs play an essential role for terrestrial living by providing a constant supply of oxygen. Lungs are the main organ of respiratory system, taking in oxygen, performing gas exchange and expelling carbon dioxide [1, 2]. The respiratory epithelium is continuously exposed to varied extraneous insults such as pollutants, viruses, and bacteria due to the unique position of the lungs toward the external environment [3]. Therefore, the cellular composition of the conducting airway epithelium is designed to ensure an integrative and protective physical barrier, clearing inhaled agents and supplying several reparative mechanisms to cope with injury which are vital for normal lung function [4].

3.1 Development of the lung

The lung epithelium from trachea to the alveoli is a continuous epithelium and the lung goes through a series of developmental phases to create it. These phases can be separated into five different stages: embryonic stage, pseudoglandular stage, canalicular stage, saccular stage and alveolarization stage [5] (Figure 1 [6]). The lung epithelium initially arises from a small region of the anterior ventral foregut endoderm, marked by the transcription factor Nkx2-1.

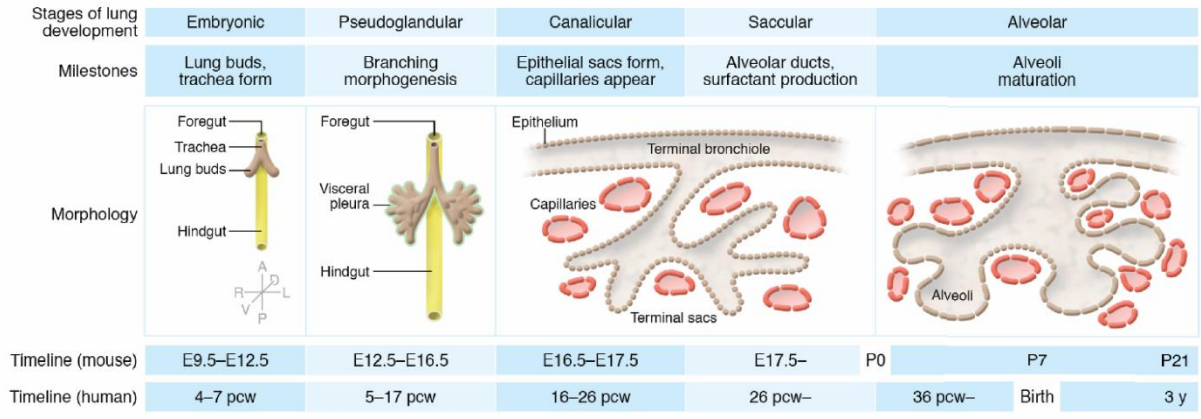


Figure 1: Timeline of five stages of lung development in mice and humans.

The five milestones of lung development are embryonic (formation of lung buds), pseudoglandular (branching morphogenesis), canalicular (formation of epithelial sacs and appearance of capillaries), saccular (production of alveolar ducts and surfactant protein), and alveolar (maturation of the alveoli). Mouse stages are represented in embryonic (E) or postnatal (P) days, and human stages are represented by post-conception weeks (pcw). Image reproduced from Bartman, C. M., A. Matveyenko, and Y. S. Prakash. *J Clin Invest* 130, no. 1 (Jan 2 2020).

The expression of Nkx2-1 transcription factor in endodermal cells is induced by morphogens from the surrounding mesoderm around E9.0 in the mouse, marking lung formation at embryonic stages [7]. Expression of Nkx2-1 is not only necessary for initiation of lung specification, but expression of this pioneering transcription factor is also essential for lung development. Expression of Nkx2-1 persists into the adult as well [3].

Evagination of the endoderm to the surrounding mesoderm results in formation of primitive trachea and two lung buds. The buds are composed of simple endodermal epithelium and a vascular plexus [3]. The primary buds extend around E12.5 during the pseudoglandular stage of the lung development and based on the mutual mesodermal-endodermal crosstalk branching occurs, generating the tree-like network of the airway. As the buds extend and branch, discriminative expression of Sox2 and Sox9 happens in the proximal/distal (stalk vs buds) regions, respectively. The proximal $Sox2^{+}$ progenitor cells become precursors of many epithelial cells of the trachea and conducting airway around E15.5, when branching morphogenesis progress toward canalicular stage, such as dome-shaped Club cells, neuroendocrine (NE) cells, ciliated cells and basal cells. Specification of $Sox9/ID2^{+}$ progenitor cells eventually leads to differentiation toward AT1 And AT2 cells [5]. Branching morphogenesis is followed by canalicular (E16.5-E17.5) and saccular stages (E18.5-birth), during which terminal branches narrow and form epithelial sacs that will later develop into alveoli. The final maturation of alveoli happens after birth during the alveolarization sage [3, 5, 7].

3.2 Anatomy and cellular organization of the adult mouse lung

The mammalian lungs consists of a branched network of airway tubes that at its most proximal part is connected to the trachea and terminating at it most distal part to millions of highly vascularized gas-exchange units known as alveoli. Air enters the respiratory tract through nasal and oral passages and pass down to the trachea, main stem bronchi and successively to more intricate conducting airways (bronchioles) and ultimately alveoli network [1, 3]. Trachea and intralobar airways are supported by cartilage while smaller bronchioles lack it. Conducting airways of respiratory epithelium are lined by pseudostratified epithelium that becomes columnar and cuboidal in small distal bronchioles until it finally terminates in cuboidal and

flattened alveolar cells [8]. The mammalian lung can be divided into two compartments, proximal and distal, which differ spatially, structurally, and functionally (Figure 2).

The lining of the epithelium in terms of its cellular constituents also varies along its proximal-distal axis. The proximal tracheobronchial zone contains ciliated cells, Club cells, mucus-producing goblet cells and basal cells. All of them reside on a continuous basement membrane, reaching the airway lumen, except basal cells. The proximal region is dominated by ciliated cells, which declines towards more distal airways where ciliated are substituted by Club cells. It is assumed that ciliated cells are terminally differentiated and derived from either basal or Club cells. Basal cells serve as a stem cell population necessary for epithelial maintenance and repair in this zone [8, 9]. In distal bronchioles, Club cells replace basal cells as the primary regenerative cells, which explains why club cells are often referred to as a second stem cell population in the airway epithelium. Club cells secrete several distinctive proteins, including Club cell secretory protein (CCSP) and metabolize xenobiotic compounds (such as naphthalene) through P450 monooxygenase system. Through the metabolization of such molecules, toxic metabolites can be generated that damage Club cells [10]. The club cell population is heterogeneous, including a subpopulation of variant type (here after called variant Club cell (^{V}Club cell)) that is resistant to naphthalene due to the absence of P450 cytochrome expression. ^{V}Club cells are capable of restoring ciliated and Club cell populations after injury and they colocalize with neuroepithelial bodies (NEBs) at branch points [11].

NEBs are clusters of 20-30 neuroendocrine (NE) cells in bronchiolar zone. NE are innervated and has play a role in oxygen sensing and modulation of immune responses [12, 13]. The bronchiolar airway terminates at the bronchoalveolar duct junction (BADJ), where transition into the alveolar zone occur. The BADJ also harbors another xenobiotic resistant cell type, named bronchioalveolar stem cell (BASC) (Figure 2). BASCs show marker expression genes of both bronchiolar Club cells and alveolar type II (AT2) cells [14] (see also section 3.4 Airway branch points and bronchoalveolar duct junction serve as two niches harboring adult lung stem cells). Finally, alveolar epithelial cells that are located at the end of bronchial tubes where gas exchange take place. In this region the epithelium consists of single squamous layer, forming gas exchange unit that consists of thin-walled alveolar cells and a dense network of capillary blood vessels to guarantee efficient gas exchange. The alveolar epithelia is composed of alveolar type I (AT1) and type II alveolar (AT2) cells. The cuboidal AT2 cells contains

numerous secretory vesicles called lamellar bodies and secrete large amount of surfactant protein to reduce surface tension and prevent alveolar collapse [15, 16]. Squamous AT1 cells cover almost 95% of the surface of alveoli and express aquapourin-5. AT2 cells are progenitors of AT1 cells, contribute to alveolar homeostasis and are capable to restore functional alveoli after injury by collaborating with niche cells [17-22].

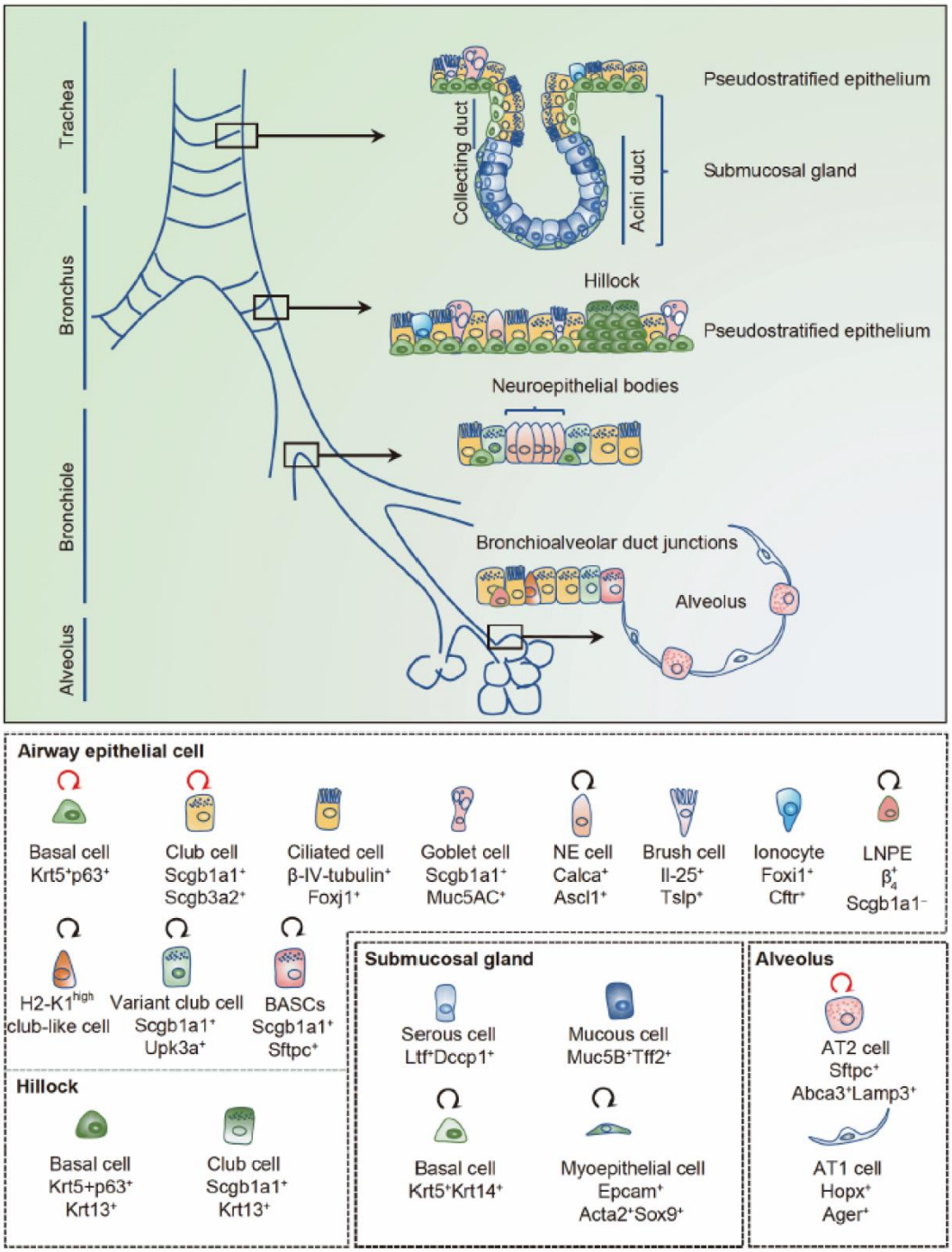


Figure 2: Cellular composition of adult mouse lung.

The adult lung consists of the trachea, bronchi, bronchioles, and alveoli in a proximal-distal pattern. Each component has different cellular constituents. The pseudostratified epithelium mainly includes Club cells, goblet cells, and multi-ciliated cells in the luminal side and basal cells in the basal side. Meanwhile, there are also specialized or rare cell types, including neuroendocrine cells (as solitary cells in the trachea and main bronchi or as neuroepithelial bodies at the bronchiolar branch junctions), brush cells, and ionocytes. The submucosal glands contain the proximal collecting duct and the distal acini. The architecture of pseudostratified epithelium is absent in the mouse bronchioles, which maintain the luminal Club cells and multi-ciliated cells but have occasionally rare atypical basal cells. Alveolar epithelia comprise two cell populations including AT2 cells and the AT1 cells. Adult lung harbor three main epithelial stem cell populations which are highlighted by red arrows: basal cells for the pseudostratified epithelium, Club cells for the bronchial epithelium, and AT2 cells for the alveolar epithelium. Other lung epithelial stem/progenitor cells are highlighted by black arrows. Image reproduced from Wang, Y., and N. Tang. *Sci China Life Sci* (Apr 30 2021).

3.3 Epithelial stem cell plasticity in the adult lung during homeostasis and repair

Adult stem cells own specific properties, enabling regeneration of various tissues and organs [3]. The lung epithelium is relatively stable under homeostatic condition, but depends on rapid and efficient regeneration to cope with constant threats from the environment [23]. The permanent exposure to external, potentially damaging agents explains why the lung had to develop and employs various partially redundant means to increase its regenerative capacity, including cellular plasticity [24, 25].

Several lines of evidence that lung epithelial cells can easily make and break identities, allowing one cell type to substitute for another one in case of damage or disappearance [24, 26-31]. In organs such as the lung that are constantly exposed to environmental insults, the existence of different regenerative strategies is particularly important [32]. To meet the crucial need of prompt and proper regeneration in response to injuries, the existence of multiple back-up and reserve cells or regenerative mechanisms is critical. In this respect, the lung might be called a 'democratic organ', since the chance to contribute to repair of the injured lung exist for many and is not only a privilege for few [25].

The first and foremost role of epithelial adult stem cells is the renewal of barrier formation. To accomplish this essential function, they must be able to replace lost cells and newly integrate into the tissue. Cell replacement happens under homeostasis condition but also during injury-induced regeneration [31]. During homeostasis, gradual loss and death of differentiated cells occur, while injury induces damage, activate immune/inflammatory responses and require rapid responses to replace large number of lost cell plus re-establishing the scaffold in damaged areas. During homeostasis, replacement of cells usually happens through progenitor cell expansion while injury-induced regeneration does not only involve proliferation and differentiation of resident stem cells but often also exploit other cellular mechanism such as dedifferentiation and/or transdifferentiation of differentiated cell types.

A typical example is the dedifferentiation of secretory Club cells to basal cells after DTA-mediated ablation of basal stem cells in lung epithelium [26]. This directly points to an alternative cellular mechanism through which tissues can regenerate after loss of stem cell. Dedifferentiating Club cells lose markers of secretory cell differentiation while acquiring

markers of basal stem cells. Another in vivo example, is the proliferation of $HOPX^{+}$ AT1 cells, a subpopulation of AT1 cells, which transdifferentiate to generate AT2 cells following pneumonectomy [27].

3.4 Airway branch points and bronchoalveolar duct junction serve as two niches harboring adult lung stem cells

As descried earlier, the composition of the lung epithelium varies along its proximal-distal axis, which is reflected by the diverse physiological functions of the lung [4]. A growing body of evidence, mainly exploiting animal injury models and genetically manipulated animals, employing the *Cre-Lox* system for lineage tracing analysis, suggests at least five putative stem cell niches in the adult lung. Starting from the upper most proximal region such niches comprise (i) duct of SMG [33], (ii) basal cells of tracheobronchial airways [34-37]. (iii) AT2 cells (most distal part) in the alveolar compartment [18, 38, 39], (iv) the microenvironment associated with neuroepithelial bodies (NEBs) [37, 40-42], and (v) bronchioalveolar duct junctions (BADJs) of the bronchiolar epithelium [43-48]. Distinct intrapulmonary airway stem cells reside in all these specific niches (Figure 2) [9], where they are ready for rapid and dynamic responses to injury and to regenerate necessary cell types [49].

Among these niches, airway branch points and BADJ have specific properties. Both niches host $Scgb1a1^{+}$ resistant cell types (^{v}Club cells at branch points and BASCs at BADJ) that survive xenobiotic insults and regenerate bronchiolar epithelium after massive sloughing of Club cells post naphthalene injury. Branch points usually host NEBs. Within intrapulmonary airways neuroendocrine cells are primarily found as a group of cells termed neuroepithelial bodies (NEBs), although pulmonary neuroendocrine cells (NE) are also dispersed throughout the epithelium of the upper and lower airways [50, 51]. Variant Club cells, characterized by naphthalene resistance in the mouse [52], were initially described adjacent to NEBs [41, 53]. ^{v}Club cells characteristically express regular Club cell marker (Scgb1a1) and express low levels of Cyp2f2. Despite this knowledge about ^{v}Club cells, specific markers defining this cell type are scarce, making it difficult to study their contribution to lung epithelium during homeostasis and repair.

In addition to the NEB microenvironment, BADJ also harbors special pollutant-resistant cell type (where the airway opens into the alveoli) [43, 44]. BASCs have been proposed to function as stem cells that contribute to regeneration of both bronchiolar and alveolar compartments, although stringent tracing with markers to confirm them as *bona fide* stem cells have only been performed recently [54, 55]. Contribution of BASC to both homeostasis and repair was described after development of novel tracing tools (see also section 23.5 Novel experimental tools for assessing dual-marker expressing cells *in vivo*).

3.5 Novel experimental tools for assessing dual-marker expressing cells *in vivo*

To harness the potential of endogenous stem cells for further therapeutic purposes, it is necessary to study them *in vivo*, identify them correctly, and thoroughly investigate their normal homeostatic behavior *in vivo* [56]. Lineage tracing studies are pivotal to study stem cell populations in vivo. Lineage tracing provides information about the abundance and localization of stem cell descendants and their differentiation potential [11]. In recent years, lung cell type-specific transgenic mouse lines have facilitated a better understanding of cellular processes underlying lung homeostasis and lung repair [57]. The need for prior knowledge about the identity of specific marker gene(s) sometimes complicates or even prevents the use of lineage tracing methods to study unknown or not well characterized cell populations. This problems gets worse when a distinct stem cell population of interest can be only distinguished from others by using more than one specific marker gene. Thus, characterization of BASC and analysis of their function during lung homeostasis and regeneration has remained controversial for more than a decade [14]. Recently, two independent studies by Liu *et al.* [54] and Salwig *et al.* [55], established genetic systems to specifically trace BASCs *in vivo*. Liu *et al.* [54] combined Dre-rox and Cre-loxP systems to generate a new dual-recombinase based reporter system. They used a *Rosa26–rox–Stop–rox–loxP–Stop–loxP–tdTomato (R26–RSR–LSL–tdTomato)* reporter, which only becomes active in BASCs that co-express two marker genes (therefore two recombinases Cre and Dre simultaneously). After subsequent recombination by Dre- and Cre-recombinases in BASCs, tdTomato expression gets activated. Salwig *et al*. also devised a novel cell tracing tool by developing a split-Cre and split-tTA system. This system is based on intein-mediated assembly

of newly engineered split-effectors which lead to selective targeting of dual-marker expressing cells (in this case BASC) in mouse lung. The split-Cre system was developed to overcome limitations of Dox and Cre/CreER system caused by the necessity for singular cell type specific promoters to restrict gene expression to cells of interest [57, 58]. In this system, the Cre recombinase protein is split into two halves that are expressed from different promoters (in case of BASC the two promoters are *CCSP* and *Spc*). Individual halves are not active but as soon as the two promoters are simultaneously active in the same cell of interest, intermolecular intein-mediated protein trans-splicing occurs and a fully functional Cre recombinase is generated. The same approach was employed for the tTA transactivator molecule. These genetic approaches circumvent problems caused by the lack of unique markers for manipulation, labeling and lineage tracing of cells of interest. They will be beneficial to advance the field of lineage relationships and stem/progenitor cells.

3.6 Cell state transition and transitory state during homeostasis and repair

Improved experimental techniques, such as single-cell RNA-seq (scRNA-seq) and computational analysis, enables a better understanding of cell states beyond what traditionally has been defined as “cell types” (i.e. cell states with well-defined functions) [59]. The identification of cells that acquire a prolonged “transition state” (also called hybrid or intermediate cell states) are of particular interest. The term ‘cell state transition’ refers to the process by which cells change states over time. By definition the transition state is an intermediate state during cell fate decisions, during which a cell show a hybrid identity between two or more states, which often represents the state of origin (that is, the initial state the cell is in) and the state of destination (that is, the identity that the cell is adopting) [60].

Transitions state are often acquired during embryonic development as cells progressively differentiate and during regeneration when stem/progenitor cells or even more mature cell types acquire new phenotype and act to fill the injury-induced void [61]. However, it is a relatively recent observation that cells can remain in a transition state for a long time. The reason for this behavior remains a mystery and the mechanism keeping transition-state cell for a prolonged time in a hybrid state are largely unknown.

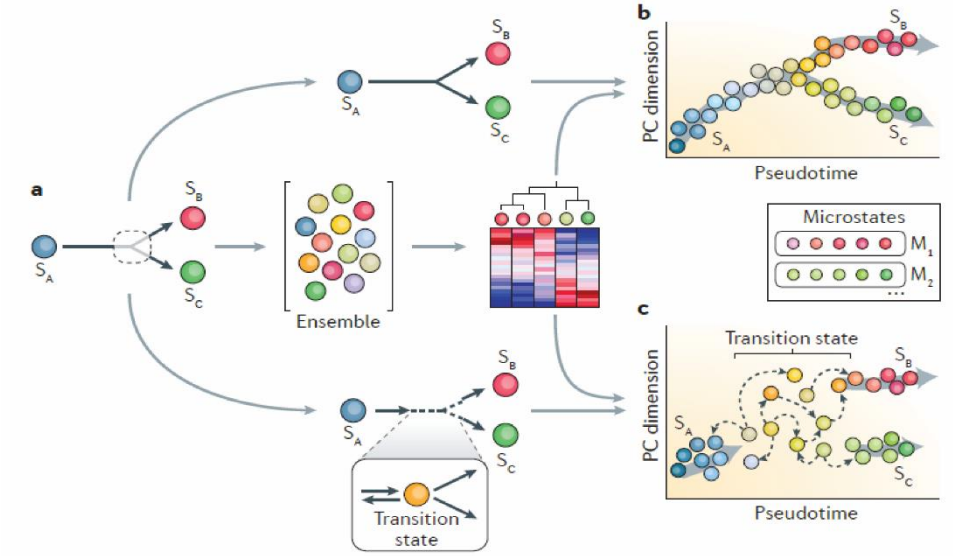


Figure 3: Discrete versus continuous cell states.

(**A**) During fate transition and cellular conversion (e.g. differentiation), a cell pass through trajectories before reaching destined faith. (**B**) These trajectories can be in a smooth and continuous manner; a sequence often referred to as 'pseudotime'. The more similar transition states, closer to each other are in pseudotime. (**C**) If transition between states is considered to be discontinuous, cells might cluster into a number of functional 'macrostates' (shown as M1 and M2). Each of these microstates has a different probability of transitioning to another state, but this is not necessarily directly correlated to similarities between a given microstate and the final state. Image reproduced from Moris, N., C. Pina, and A. M. Arias. Nat Rev Genet 17, no. 11 (Nov 2016): 693-703.

Characterization of transitory cell states from single-cell gene expression data is challenging. It is also unclear why cells either undergo smooth transitions or are locked for extended time periods in an intermediate states [62] (Figure 3).

Recently three different research group detected intermediate cellular states during the transition from AT2 cells to AT1 cells in the regenerating alveolar epithelium of the lung [63-65]. While investigate differentiation of AT2 to AT1 cells Kobayashi *et al.* identified a distinct transitional cell population termed "pre-alveolar type-1 transitional cell state" (PATS) [64]. In this transitional state, AT2 cells undergo extensive stretching to fill the gaps left by injury-induced alveolar epithelial cell loss. Two other studies also provided support for the existence of an intermediate cellular state during alveolar regeneration. Strunz *et al.* [63] used scRNA-seq data collected from bleomycin-injured mouse lungs, to identify a transitional state marked by expression of Krt8 that precedes replacement of AT1 cells and peaks during the fibrogenic phase of repair. They termed this state Krt8^{+} alveolar differentiation intermediate (ADI) state. In the report by Choi *et al.* [65], in bleomycin-injured mouse lungs, the authors identified a transitional state that was characterized by expression of Krt8 and Cldn4, which they termed damage-associated transient progenitors (DATPs).

A common feature of intermediate cell states is that the intermediate cell state allows bidirectional transition between cell types, while providing additional functions. With other words, intermediate AT2 cells are able to compensate for occurring gaps in the epithelial layer by extensive stretching, while transitioning toward AT1 cells, eventually acquiring a flat and squamous cell type and abandoning the initial secretory cuboidal properties.

Taking into account these new findings, the Waddington's epigenetic landscape with fixed and defined bifurcation and valleys for already known cell types, may have a lot of 'divots' that represent intermediate cell states (e.g. transitional PATS/ ADI/DATP states) necessary for transitioning from one cell state to another.

Transitory cell states also have been discussed in the context of transdifferentoation (direct conversion of one differentiated cell type into another identity). Transdiferentiaition (Td) is consider as a direct conversion process, in which cells do not go through intermediate transitory steps. However, the 'direct' aspect of transdifferentiation is not as clear-cut as previously thought [66]. Several reports suggest that cells pass through a continuum of intermediate states during Td that are not well understood, some resembling transitional states occurring during development [67, 68] (Figure 4).

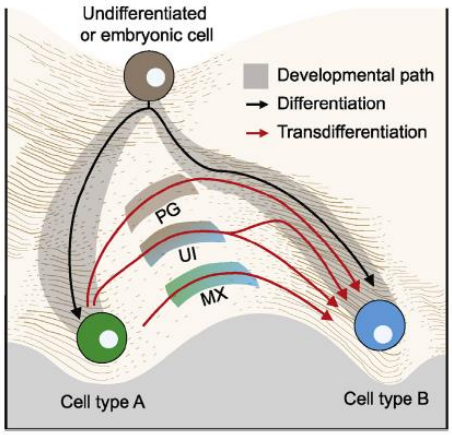


Figure 4: A summary of different modes of cell state transition in transdifferentiation.

Converting cells can deviate from the path of development. Cells undergoing Td can pass through mixed (MX), unspecific intermediate (UI) or progenitor-like (PG) transition states. In the MX state during Td, both fates (donor and target cell) are present. In the UI transition state, donor cells may lose original identity completely before acquiring new fate and the intermediate state does not resemble donor nor target fates. In the PG transition state, converting cells may revert to progenitor state with partial or complete pluripotency. Image reproduced from Reid, A., and B. Tursun. *Curr Opin Syst Biol* 11 (Oct 2018): 18-23.

4 AIM OF THE STUDY

Delineation of lung epithelial stem cell hierarchy and heterogeneity is a controversial and challenging topic. Although lung cell type-specific transgenic mouse lines caused tremendous progress in understanding cellular processes that underly lung development, homeostasis, and repair, main questions remain unanswered. The advent of new experimental tools for assessing dual-marker expressing cells *in vivo* opens the opportunity to characterize previously not accessible cell populations and understand their functions during homeostasis and regeneration.

The aim of this dissertation was to investigate heterogeneity among airway secretory cells in the adult mouse lung (Club cells) and shed more light on airway stem cells located within the branch point niche.

Starting with the identification of novel markers or combination of markers using single cell RNA sequencing, I pursued this goal by thoroughly characterizing the distribution of putative markers in mouse lung airways. Spatial resolution of candidate marker gene or combination of them was planned to be achieved by *in-situ* sequencing and by generation of new knock-in mouse strains. Finally, functional analysis of newly identified cell populations was on the agenda by specifically ablating distinct cell population and investigate the consequences for lung maintenance and regeneration.

5 MATERIALS AND METHODS

5.1 Materials

5.1.1 Chemical reagents

Name	Manufacturer (catalogue number)
LIVE/DEAD™ Fixable Near-IR Dead Cell Stain Kit, for 633 or 635 nm excitation	Thermo Fisher Scientific (L34976)
ampicillin sodium salt	Roth (K029.2)
4',6-diamidino-2-phenylindole (DAPI), dilactate	Sigma (D9564)
3,3'-diaminobenzidine tetrahydrochloride	Merck (Sigma-Aldrich) (D5905)
DMEM high glucose with L-glutamine and 25 mM HEPES	Gibco (42430)
doxycycline hyclate	Sigma (D9891)
Entellan® new	Merck (Sigma-Aldrich) (1.07961.0100)
Eosin G	Merck (Sigma-Aldrich) (921K11613835)
ethidium bromide	AppliChem (A1152.0100)
fetal bovine serum	Sigma (F7524)
formaldehyde solution (37 %)	Sigma (252549)
FSC 22® Clear frozen section compound	Leica (3801480)
gelatin from porcine skin, Type A	Sigma (G1890)
GeneRuler™ 1 kb Plus DNA Ladder	Thermo Scientific (SM1333)
GeneRuler™ 100 bp DNA Ladder	Thermo Scientific (SM0241)
glutaraldehyde solution (25 % in H_20)	Sigma (G5882)
Hank's BSS (HBSS)	PAA (H15 -009)
hematoxylin solution, Gill No. 3	Sigma (GHS332)
hydrogen peroxide solution (30 %)	Roth (8070.2)
Isoflurane (IsoFlo®)	Zoetis
MACS® buffer (autoMACS® Rinsing Solution)	Miltenyi Biotec (130-091-222)
Miglyol® 812	Caelo (3274)
Mowiol® 4-88	Millipore (457904)

naphthalene	Sigma (84679)
penicillin/streptomycin	Sigma (P0781)
saline solution (NaCl 0.9 %)	B Braun (3820084)
Potassium hydroxyde	Roth (6751.1)
D(+)-Saccharose	Roth (909701)
Glycerol	Roth (3783.1)
Xylol	Roth (9713.3)
Ethanol	Roth (k928.3)
2-propanol	Roth (67524)

5.1.2 Media and solutions

Name	Component	Concentration
FACS buffer (in PBS, pH 7.4)	EDTA	5 mM
	HEPES	25 mM
loading buffer DNA (5x)	glycerol	50% (w/v)
	orange G	2%
phosphate-buffered saline (PBS), 10x	NaCl	80 g/1
1x PBS adjusted to pH 7. 4	Na_2HPO_4	6 g/1
	KH_2PQ_4	2 g/1
	KCl	2 g/1
staining solution (β-galactosidase)	Na_2HPO_4	77.4 mM
	NaH_2PO_4	22.6 mM
	$MgCl_2$	2 mM
	NP-40	0.02 %
	sodium deoxycholate	0.01 %
	$K_3Fe(CN)_6$	5 mM
	$K_4Fe(CN)_6$	5 mM
	5-bromo-4-chloro-3-indolyl-β-O-galactopyranoside	1 mg/ml
solution A (plasmid isolation), pH 8	glucose	50 mM
	EOTA	l0 mM
	Tris	25 mM

solution B (plasmid isolation)	NaOH	0.2 M
	SDS	1 %
solution C (plasmid isolation)	potassium acetate (5 M)	60 ml
	glacial acetic acid	11.2 ml
	H_2O	28.5 ml
TAE buffer, 50x, pH 8.5	Tris	2 M
	glacial acetic acid	1 M (57.1 ml/liter)
	EDTA	50 Mm
TENS buffer, pH 8	Tris-HCl	50 mM
	EDTA	40 mM
	NaCl	100 mM
	SDS	1 %
wash solution, pH 7.4	Na_2HPO_4	77.4 mM
(β-galactosidase staining)	NaH_2PO_4	22.6 mM
	$MgCl_2$	2 mM
	NP-40	0.02 %
	sodium deoxycholate	0.01 %

5.1.3 Kits

Name	**Manufacturer (catalogue number)**
electroporation cuvette (0.1 cm gap)	PeqLab (71-2010)
Filtropur V50 0.22 µm	Sarstedt (83.1823.001)
MS Columns	Miltenyi Biotec (130-042-201)
Superfrost Ultra Plus® microscope slides	Thermo Scientific (J3800AMNZ)
Tissue-Tek® Cryomold, Standard	Sakura (4557)
Whatman® 595 1/2 folded filters	Whatman (10311645)
NucleoSpin® Gel and PCR Clean-up Kit	Macherey-Nagel (740609)
pGEM®-T Easy Vector System	Promega (A1360)
Vectastain Elite ABC Kit	Vector Laboratories (PK-6100)
miRNeasy Micro Kit (50)	QIAGEN (217084)
RNase-Free DNase Set (50)	QIAGEN (79254)
Buffer RWT (80 ml)	QIAGEN (1067933)

5.1.4 Antibodies

Name	Manufacturer (catalogue number)
anti-acetylated tubulin	Sigma (T7451)
Alexa Fluor -conjugated secondary antibodies	Invitrogen
biotinylated secondary antibodies	Vector Laboratories
anti-CD31 microbeads, mouse	Miltenyi Biotec (130-097-418)
anti-CD45 microbeads, mouse	Miltenyi Biotec (130-052-301)
Anti-Ter-119 microbeads, mouse	Miltenyi Biotec (130-049-901)
Anti-CC10 Antibody (E-11)	Santa Cruz Biotechnology (sc-365992)
Anti-RUNX1 / AML1 + RUNX3 + RUNX2 antibody	abcam (ab92336)
Recombinant Anti-FoxJ1 antibody [EPR21874]	Abcam (ab235445)

5.1.5 Enzymes

Name	Manufacturer (catalogue number)
Dispase	Corning (354235)
DNase I	Roche (10104159001)
T4 DNA Ligase	Fermentas (EL0011)
RNase Inhibitor, Murine	New England Biolabs (M0314S)

5.1.6 Plasmids

Name	Use/Application	Source/Reference
pGEM-T easy	TA cloning of PCR products	Promega
pL452		From Dr. Neal G. Copeland [69]

5.1.7 Genotyping primers

	Name	Nucleotide sequence (5'- 3')
CCSP knock-in strains	CCSP tga wt gt s	TGGGGGTGTGGGGTGTTTCT
	CCSP tga wt gt as	CCAGCTCCAATGCTCTGGCTCC
	CCSP tga sfY gt as	GGCGAAGCACATCAGGCCGT
	CCSP tga mCherry gt as	ACATGAACTGAGGGGACAGG
SPC knock-in strains	SPC tag wt gt s	TACCAGCTCTCAGGTGGCCCT
	SPC tag wt gt as	CCTCCATCCTCCGAGCTCCCG
	SPC tag sfY gt as	CAGCTCCTCGCCCTTGCTCAC
	SPC tag iCre N gt s	GGCTGCCTGGTGCAAGCTGAA
	SPC tag tTA N gt s	TCCGCCGTGGGCCACTTTAC
	SPC tag iCre gt s	GACATGGCCAGGGCTGGTGTG
Responder/deleter	lacZ s	GTTGCAGTGCACGGCAGATACACTTGCTGA
strains	lacZ as	GCCACTGGTGTGGGCCATAATTCAATTCGC
	Rosa26 for	AAAGTCGCTCTGAGTTGTTAT
	Rosa26 rev	GGAGCGGGAGAAATGGATATG
	Rosa26 spliAC	CATCAAGGAAACCCTGGACTACTG
	Cre s	GACCAGGTTCGTTCACTCATGG
	Cre as	AGGCTAAGTGCCTTCTCTACAC
	Dta207 s	TAA TGA AAA CCC GCT CTC TG
	Dta578 as	CTG ACA CGA TTT CCT GCA C
Foxj1 knock-in strain	KK17-F	TCAGGAGTTTGAGGAGGCCACAGG
	KK20-F	GATCTGATGCGCGTGGATAACCTGC
	KK14-R	CTGATGCTCTTAGCCTCAAGTC
Cyp2f2-CFP knock-in	KK01-F	ATAGTCGACTTGGGCTCCGTAGCCAGATCAC
strain 5'	CFP gt as	GTGCGCTCCTGGACGTAGCC
Cyp2f2-CFP knock-in	CFP gt s	GAG AGT GAT CCC GGC GGC G
strain 3'	KK02-R	TATGCGGCCGCATTCCTCAAGTTGGGGTGAT CCTACAC

5.1.8 In situ sequencing probes

All of the probes for in situ sequencing were ordered form Integrated DNA Technologies (IDT).

Gene symbol	Padlock Probe sequence
Nme9-p351	CTTCCTGGATCTTCATGATTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCGCATCTATCTTCTTTTTTAATAGTATGTCAAATCCTG
Nme9-p194	CCTCATCAGTAGCTGTATATCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCGCATCTATCTTCTTTTGAGTGCACCCAGAAGTT
Nme9-p446	CTCAAATGCCTCCTCTCTCCTCTATGATTACTGACTGCGTCTATTTAGT GGAGCCGCATCTATCTTCTTTACATGTGATGTACTAGCCT
Nek5-p1368	ATGTTCCTCCTTACATCTGTCCTCTATGATTACTGACTGCGTCTATTTAG TGGAGCCCGTACTATCTTCTTTCTCTGTCTGTATAGTCTTC
Nek5-p709	CCAGATGGTGGAAGTTGTTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCCGTACTATCTTCTTTTCCTTGACAAATCTTCAGAA
Nek5-p2010	TTAGCCACACAGATCAAACTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCCGTACTATCTTCTTTCATTTGATGAACCTTGGTC
Stk33-p1663	CTCATCACTTTCTGGGTTATCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCAGCTCTATCTTCTTTCTCATCTGTGTTGGTCTC
Stk33-p949	GGTTGACAGTCTTCAGGATCCTCTATGATTACTGACTGCGTCTATTTAG TGGAGCCAGCTCTATCTTCTTTAGGTGGATGATGTGTTGAT
Stk33-p2219	AGTAGAACTAAGCGAGCATTCCTCTATGATTACTGACTGCGTCTATTT AGTGGAGCCAGCTCTATCTTCTTTGTGCCTAAATATAAGGATTTG
Fank1-p290	TGTATAGTGTCCTTGGTTCTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCTCGACTATCTTCTTTGTGACCTTCAGTCGAAATT
Fank1-p755	CTCACAACCGTCCTTTATTCCTCTATGATTACTGACTGCGTCTATTTAGT GGAGCCTCGACTATCTTCTTTCAGTGTCTACAACATCTAC
Fank1-p1368	AACACCAATGCGTAAGTATTTCCTCTATGATTACTGACTGCGTCTATTT AGTGGAGCCTCGACTATCTTCTTTCATGCACTCTCGTTTATGT
1700001L19Rik-p1498	GAGAGATGCTTTCTTCCAATCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCCAGCCTATCTTCTTTGAAATTCAATCGCAGTAGAT
1700001L19Rik-p320	ATAATCTAGTCTCCTCTTAAATCCTCTATGATTACTGACTGCGTCTATTT AGTGGAGCCCAGCCTATCTTCTTTCTCGGTGTAATTTCTGGTT
1700001L19Rik-p752	CAACATAAGGAAAGGTGCTTCCTCTATGATTACTGACTGCGTCTATTT AGTGGAGCCCAGCCTATCTTCTTTAATCCTTACCCTAAAGCTTT
Cfap53-p85	GCTGTACATCTTCAAGTCTTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCCAAGCTATCTTCTTTAACAGTGCCAAACTGCTG
Cfap53-p620	TTCCGTTCCTCAACCACTCCTCTATGATTACTGACTGCGTCTATTTAGT GGAGCCCAAGCTATCTTCTTTGTTGAATTCTATCTGCGCA
Cfap53-p966	TCTTATCCTGCACAGCTTTTCCTCTATGATTACTGACTGCGTCTATTTAG TGGAGCCCAAGCTATCTTCTTTCCGCTGCATAGACTCTA

Fam47e-p29	TTTCTACAGAGAGAGTCTCTCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCCGCGCTATCTTCTTTATGCAGGACAGCATCCAA
Fam47e-p1010	ACACAGCTGTACATAACTAATCCTCTATGATTACTGACTGCGTCTATTT
	AGTGGAGCCCGCGCTATCTTCTTTCTACCCACATGCAGACAT
Fam47e-p1407	GGACTACACACGAATATTGTCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCCGCGCTATCTTCTTTATAACACGAAAGCGTGAGA
Fam92b-p671	AAGTCAGAGAAGATTCTCCTCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCGCCACTATCTTCTTTACCATCTCGATGGTCACA
Fam92b-p181	TCAGCATTTGTCACCGTGTCCTCTATGATTACTGACTGCGTCTATTTAG
	TGGAGCCGCCACTATCTTCTTTAACTGACCGAAATACCTC
Fam92b-p931	GTCTCACTGTCCTCTTCTCCTCTATGATTACTGACTGCGTCTATTTAGTG
	GAGCCGCCACTATCTTCTTTTCTCCTCCACGGAATCA
Stmnd1-p176	TGACATCAGCCTTGGATCTCCTCTATGATTACTGACTGCGTCTATTTAG
	TGGAGCCACCGCTATCTTCTTTTTTCTTTGGAAGAAGTTACTC
Stmnd1-p874	AGCTCCAAACCTGTACAGTCCTCTATGATTACTGACTGCGTCTATTTAG
	TGGAGCCACCGCTATCTTCTTTATACATGTCTGACTCCACA
Stmnd1-p1249	TCATAAAGGACTCATCTGCTCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCACCGCTATCTTCTTTTAATATACATCTATGCAGTTTG
Adgb-p1833	CTGTGACTTGTCCAGTGTTCCTCTATGATTACTGACTGCGTCTATTTAG
	TGGAGCCACACCTATCTTCTTTTCGTTTATTGCTAACTCTTC
Adgb-p2672	CTCATTTATTGGTGTAGAATATTCCTCTATGATTACTGACTGCGTCTAT
	TTAGTGGAGCCACACCTATCTTCTTTCTCGGAGGTATACTCCTT
Adgb-p4328	TCCTTCCTCTGAGGATTTATCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCACACCTATCTTCTTTATATTTCCCAGTATCTAGTTC
Cetn4-p680	ACCTAAACTAAACGTGGAAATTCCTCTATGATTACTGACTGCGTCTATT
	TAGTGGAGCCGACACTATCTTCTTTATTAGACACCAAGGGACTA
Cetn4-p182	CAATATCGAATAAGTCAAAGGTCCTCTATGATTACTGACTGCGTCTATT
	TAGTGGAGCCGACACTATCTTCTTTTCTATGGTTCCAGATCCAT
Cetn4-p478	TTCATCTTCGGTTAAGTTTTCTCCTCTATGATTACTGACTGCGTCTATTT
	AGTGGAGCCGACACTATCTTCTTTTTCATCAAGCATTTCCTGTAA
Nme5-p420	TCTCCTTGGCTACTAAACTTCCTCTATGATTACTGACTGCGTCTATTTA
	GTGGAGCCAAGCCTATCTTCTTTTAAGCTGTCCGGGTGTG
Nme5-p524	TGAACCTGATCTCTCGCTCCTCTATGATTACTGACTGCGTCTATTTAGT
	GGAGCCAAGCCTATCTTCTTTATAATCACGGCTGGAAACA
Nme5-p761	TTGGCTCTCGCTCTGAGTCCTCTATGATTACTGACTGCGTCTATTTAGT
	GGAGCCAAGCCTATCTTCTTTGCTTCTTACAGTAAGGTAAC
Drc3-p447	AGCCAGACGAGGTGTATTCCTCTATGATTACTGACTGCGTCTATTTAG
	TGGAGCCGATCCTATCTTCTTTTGTTGTTAAAGGACAAATCC
Drc3-p779	TCTTCGCCATCTCTCTCATCCTCTATGATTACTGACTGCGTCTATTTAGT
	GGAGCCGATCCTATCTTCTTTAATGCTGTACTGATGTTTCA

Drc3-p1305	TCCAGCTGTTCCACCAGTCCTCTATGATTACTGACTGCGTCTATTTAGT GGAGCCGATCCTATCTTCTTTTCAAATATGTTTATGGTCTCC
Gm1673-p493	AAGATCTTCAGCGTCCAGTCCTCTATGATTACTGACTGCGTCTATTTAG TGGAGCCCTCACTATCTTCTTTAAGGAATATCACTGGCCA
3300002A11Rik-p887	ACTCAGCAGCAGACAGATTCCTCTATGATTACTGACTGCGTCTATTTA GTGGAGCCATACCTATCTTCTTTAATCCGATGATGTCTGATC
3300002A11Rik-p713	GTCCCTCCATTTGTGTGATCCTCTATGATTACTGACTGCGTCTATTTAG TGGAGCCATACCTATCTTCTTTACGAGAGAGCTGAGACT

5.2 Methods

5.2.1 Generation of the *Cyp2f2*^{CFP} knock-in mouse line

The *Cyp2f2*$^{-T2A\ CFP}$ targeting construct was designed to insert coding regions of CFP fused with T2A before the endogenous stop codon in the mouse *Cyp2f2* Exon10 on chromosome 7. A targeting vector was designed with a T2A CFP and Neomycin resistance (NEO) cassette. The NEO cassette, flanked by LoxP sites, in the targeted allele was excised by crossing F1 progeny with CMV-Cre deleter mice resulting in the *Cyp2f2* $^{-T2A\ CFP}$ knock-in allele.

5.2.2 Generation of the split *Foxj1*$^{-2A\text{-}YFP\text{-}tTA\ N}$ knock-in mouse line

A targeting vector was designed with T2A YFP T2A tTA-N and Neomycin resistance (NEO) cassettes inserted at the endogenous stop codon in the Foxj1 locus. NEO cassettes, flanked by loxP sites, in the targeted allele was excised by crossing F1 progeny with CMV-Cre deleter mouse strain, resulting in the *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ knock-in allele.

5.2.3 Naphthalene injury and EdU incorporation

Adult mice between 8 to 12 weeks of age were selected (both genders included) and either naphthalene dissolved in MIVIGDOL (oil vehicle) or the vehicle alone was administered intraperitoneally at the dose of 200 mg/kg body weight to the treated and control groups respectively. The lungs were harvested 3 and 7 days post injection or the mice were allowed to recover 21 and 60 days and then killed for evaluation of LacZ reporter gene expression.

5.2.4 Influenza injury

For IV infections, mice were instilled intratracheally after anesthetization with 2000 plaque-forming units (pfu) of influenza A/H1N1/Puerto Rico/8/34 (PR8; Peteranderl *et al.*, 2016).

5.2.5 Lung tissue collection

At time points of interest, mice were euthanized followed by exsanguinate blood form inferior vena cava. By performing a small cut at the upper part of the diaphragm, lungs effortlessly collapsed and both the diaphragm and thoracic cavity were cut open without harming the lungs. Blood was flushed out of the lung by perfusion with PBS through the right ventricle. To insert 30G needle (made blunted) into the trachea, a half thorough incision was made in the proximal region of the trachea and lung was inflated with 1% formaldehyde using a 30G needle, fixed in position by a thread in the form of a knot. The lung were inflated and initially

fixed with PFA 1% under constant pressure for 5 minutes and the inflated harvested lung was subjected to further fixation of 2 hours in PFA 1% on ice. After 2 hours the knot was removed from trachea, allowing the PFA come out of the lung and then the lung were kept for 2 more hours in PBS on ice, followed by overnight dehydration in 20% Sucrose. The next day, the lobes were separated and embedded in Tissue-Tek O.C.T compound, solidified on dry ice and cut using a cryotome (Leica Biosystems Cryostat) of 10µm thickness.

5.2.6 Lung-cell isolation and FACS sorting

Mice were prepared for organ harvest as described above. Here as well blood was flushed out of the lung by perfusion of PBS trough the right ventricle. Lung was cannulated by making small incision in the trachea and insertion of the insulin syringe plugged with blunt 30G needle. The needle was fixed in place by a knot and 1ml of Dispase (Corning, 354235) was pushed through the needle. The inflated lung was then harvested from a thoracic cavity, the heart was then removed and the lung was chopped to smaller fragments by a scissors. All the fragments were then collected with 1ml more of the Dispase and then incubated in 37°C with frequent agitation for 20 to 30 minutes. 1µl/ml DNase (20 U/ml, Roche) was added to digestion solution too. Half way through the digestion time, tissue fragments were more broken down by several times pipetting. Tissue fragments were loaded on a 100µm cell strainer. Any remaining tissue from the digestion were also collected by DMEM containing 10% FBS and loaded on the filter. By pressing the tip of the pipet on the mesh, tissue was forced to pass through the mesh and half way through, the membrane was washed two times with 800µl of medium. The flow-through from the first step of passing through the filter was collected and passed through the second filter (40µm) which yield a final single cell suspension. Cells were centrifuged 300 *x g* for 5 minutes. The supernatant was soaked off and the pellets were re-suspended in 1ml FACS buffer followed by centrifugation 300 *x g* for 5 minutes. The pellets were then re-suspended in 70µl MACS® buffer and 10µl CD31 microbeads, 10µl CD45 microbeads and 10µl Ter119 microbeads were added to the cell suspension followed by incubation for 15 minutes in 4°C. After incubation time, 1ml MACS® buffer was added to cell suspension and centrifuged 300 *x g* for 5 minutes. Supernatant was discarded and the pellets were re-suspended in 500µl MACS® buffer and loaded on pre-equilibrated MACS® columns. Columns were washed two times more each time with 500µl MACS® buffer and the collected flow-through was sorted at BD FACSAria™ III.

5.2.7 FACS based quantification followed by β-galactosidase staining

Epithelial cells of interest were FACS isolated based on endogenous fluorescent proteins. Sorted cell fractions were seeded in gelatin coated 96 well plate and were fixed by glutaraldehyde (0.25%) for 5 minutes. Next step the fixed cells were pelleted by centrifugation 300 *x g* for 5 minutes and the supernatant was discarded. LacZ staining was performed by adding 150µl of warm and filtered LacZ solution to the cells and incubated overnight at 37°C followed by post fixation (4% PFA) next day.

5.2.8 Histology and immunofluorescence/immunohistochemistry staining of lung sections

Lungs were embedded in FSC 22 Frozen Section Media (3801480, Leica Biosystems), and 10-µm cryosections were used for staining. Sections were then washed with PBS to remove Frozen Section Compound and were incubated overnight at 4°C with anti-Foxj1 (1:1000, Abcam, ab235445) primary antibody. Sections were then washed with PBS (3x 5 minutes) to remove unbound antibody and incubated with Alexa Fluor-conjugated secondary antibody and nuclear stain DAPI (1:1,000) at room temperature for 1 hour. Immunostained slides were then washed with PBS (3x 5 minutes) to remove unbound secondary antibody and mounted with MOVIOL. Stained sections were imaged using a Zeiss AXIO Observer/Imager.Z1 or Zeiss LSM 880.

For immunohistochemistry, following the incubation of the primary antibody (anti-Acetylated Tubulin, 1:500, Sigma, T7451) and washing with PBS to remove the unbound antibody as described above, sections were treated with 0.3% hydrogen peroxide in methanol for 15 minutes and then washed in PBS. Sections were incubated at room temperature for 1.5 hour with a biotinylated goat anti-mouse IgG secondary antibody (Vector Laboratories). After rinsing (3x 5 minutes with PBS), sections were incubated with Vectastain EliteABC peroxidase kit according to the manufacturer's protocol (Vector Laboratories).

5.2.9 β-galactosidase staining

Defection of LacZ signal was performed by overnight incubation of slides in LacZ solution ((5 mmol/L K3Fe[CN]6, 5 mmol/L K4Fe[CN]6, 2 mmol/L MgCl2, 0.02% Nonidet P-40, 0.01% sodium deoxycholate and 1 mg/mL 5-bromo-4-chloro-3-indolyl-D-galactopyranoside [X-gal]), 37 degree followed by next day rinsing with PBS and mounting with MOVIOL and placing a cover slip on top.

5.2.10 scRNA-seq (ICELL8)

The ICELL8 system enables the flexibility to analyze multiple parameters per experiment and uses the power of imaging to distinguish true single from multiple cells and allow live/dead cell discrimination as well; therefore allows user to draw more meaningful conclusions from data. ICELL8 system is composed of a chip with 5184 pre-dispensed barcoded wells, a source plate capable of loading 8 different samples, multi sample nano-dispenser, imaging station and a cell select software (Figure 14). Dispensing the cells on a chip follows the probability of Poisson distribution meaning that from defined concentration of cells and with fixed rate of cell dispensation onto the multi nano-well plate, almost 35% of the wells meaning 1872 wells will be occupied by single cells (in optimum condition) (Figure 13). After dispensation, the chip should be fast covered with the film and get imaged as quickly as possible to minimize the possibility of changes in the cells transcriptome. Imaging will be done in 2 channels (blue and red) and since before dispensation PI and Hoechst were added to cell suspension, live/dead discrimination would be later possible in the cell select software. The cell select software allow the user to go through every individual well and exclude all the wells containing more than one cell for downstream analysis as well as excluding all the dead cells and wells containing cell debris. The critical step in single cell analysis with iCell8 technique is to consider minimum time gap between FACS sorting and dispensing step on the chip to get the maximum viable cells before freezing the chip for later downstream analysis and cell selection by cell select© software. The concentration of the sorted cells for dispensing also should be calculated in a way that while dispensing maximum number of wells loaded with single cell could be achieved. After imaging with automated fluorescence microscope, the chip was immediately frozen in -80. Proper cells were selected by cell select software and all PI positive, multi cell and wells containing debris were excluded before RNA-seq analysis.

5.2.11 *In situ* sequencing and pad lock probes design

42 pad lock probes were designed based on the 15 candidate genes selected from comparison of single cell analysis and ciliated cell transcriptome. To design pad lock probes for in situ sequencing, the mRNA sequence of the gene was downloaded from NCBI website and the online tool was used to obtain Taqman Probe sequences (30-40bp) as sequences to design for Padlock Probes.

(https://eu.idtdna.com/PrimerQuest/Home/Index?Display=AdvancedParams).

Obtained sequences were validated (BLAST) to ensure that the probes specifically and solely bind to the transcript and/or transcript variants of interest.

To prepare the tissue for in situ sequencing, lung tissue was fixed in PFA 4% for 2-12hours (depending the size), washed briefly with PBS 1X and embedded in Cryomatrix (Leica Cat No. 3801480 FSC22 Frozen Section Comp Clear) and get frozen with isopentane and dry ice. They were stored in -80^{o}C until sectioning. 10µm sections were placed on polysine-coated slides and kept in -80^{o}C upon use. Slides were removed from -80^{o}C and incubate 45^{o}C for 15 minutes to prevent moisture accumulation on the tissue. Samples were post fixed with PFA 4% in PBS 1x pH7.4 (freshly prepared) for 5 minutes. 2 x 5 minutes wash with PBS-Tween 0.05% (450ml DEPC H2O + 50ml PBS 10x + 250ul Tween20). To permeable and dehydrate the tissue, 3 minutes incubation in 0.1M HCl in DEPC H2O at RT followed by dehydration in descending ethanol series was performed. All of the following reactions were performed in Secure-Seal hybridization chambers. In the next step padlock probes were hybridized and ligated by SplitR ligase at RT overnight. Amplification of now ligated and circular padlock probes were carried out by an RCA mix containing 0,5 U/µl phi29 polymerase (Lucigen), 1× phi29 polymerase buffer, 0.25 mM dNTPs, 0.2 µg/µl BSA and 5% glycerol in DEPC H_2O was added to the reaction chamber and incubated for 30°C overnight to carry out the RCA. After the incubation, the slide was washed 6 times 1 minutes with PBS-Tween 0.05% and leave in last solution until next step. Finally, 0.02U/ul of each corresponding detection probe in 2× SSC and 20% formamide was applied to the slide and incubated at 37 °C for 30 minutes. Excess amounts of detection probes were then eliminated by washing three times with DEPC-PBS-T. In this step tissue should be dehydrated with ascending series of ETOH, air-dry followed by applying Slowfade mounting medium and coverslip. Then store in dark, at 4^{o}C or RT until imaging. For the second round of hybridization, coverslip and the mounting medium were removed by dipping the slide in a series of ETOH 70%, and place vertically at 45^{o}C until coverslip is detached, 85% ETOH for 2 minutes and 100% ETOH for 2 minutes. Air-dry and mount chamber. To strip of the first probes, the UNG-reaction solution was placed in the chamber and seal the two holes with PCR adhesive membrane. Incubation at 37^{o}C for 30-45min followed by 3 times 5 minutes 65% formamide in H2O at 30^{o}C and wash 2 times with PBS-Tween 0.05% and leave in last solution. Next step continue as with the 1st hybridization. Similar round were carried out for third and fourth round of hybridization.

6 RESULTS

6.1 Runx1 is enriched in Club cells of mouse lung epithelium

To identify proteins specifically present in either Club or AT2 cells we took advantage of a mouse strain that expresses the fluorescent reporter proteins mCherry and YFP in Club and AT2 cells [55]. Club and AT2 cells were isolated by FACS and subjected to mass spectrometry-based proteomics. Analysis of the resulting datasets yielded a list of proteins that appeared to be specific for either Club or AT2 cells. Identified proteins are depicted in Figure 5. We observed a strong enrichment of Runx1 in Club cells, which motivated me to study the role of Runx1 in the lung epithelium, since Runx1 has been implicated in epithelial stem cell biology and pathology, beside of its function in hematopoietic stem cells [70, 71].

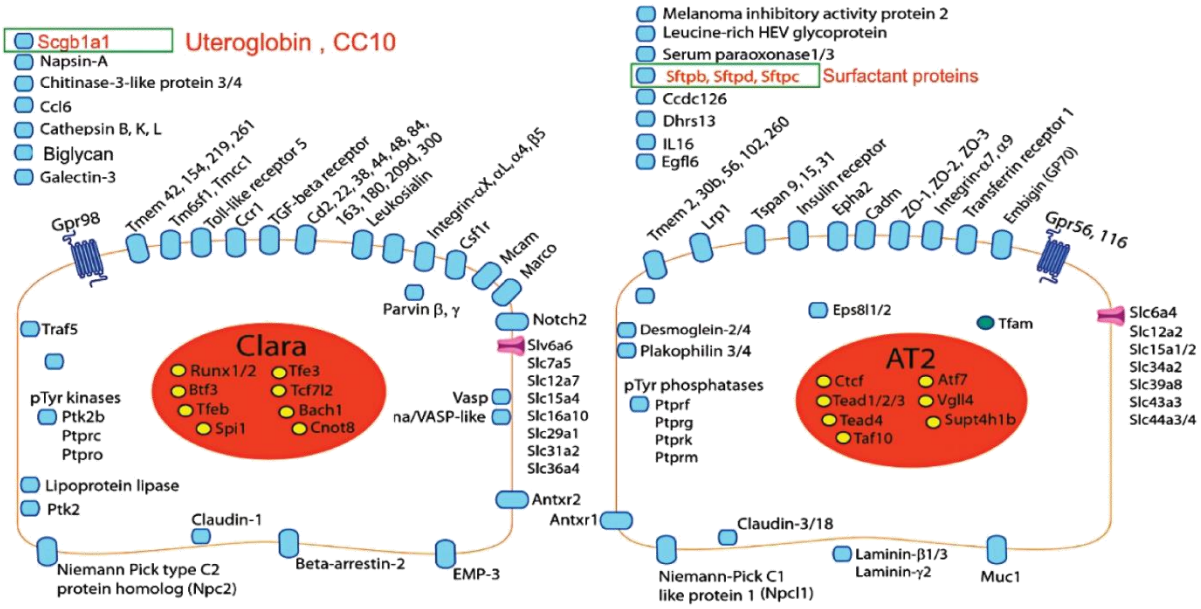


Figure 5: Results from proteomic profiling of FACS-sorted Club and AT2 cells using CCSP^{mCherry} SPC^{YFP} knock-in reporter mice.

Proteomic profiling resulted in the identification of unique expression patterns of membrane, cytoplasmic and nuclear proteins in Club and AT2 cells.

6.1.1 Visualization and cell-fate tracing of Runx1^{+} cells in adult mouse lung

To confirm findings from the proteome analysis, immunofluorescence staining for RUNX1 was performed, followed by tracing of *Runx1*-expressing cells and inactivation of *Runx1* in the respiratory epithelium.

To visualize Runx1-expressing cells, *Runx1*$^{-2A\ YFP-2A\ tTA}$ knock-in mice were crossed with the *tetO*$_{bi}^{EGFP/lacZ\ pos}$ strain to generate Runx1-viewer mouse (Figure 6A). Expression of the LacZ reporter gene observed in the adult lung epithelium of Runx1-viewer mice partially recapitulated the outcome of the proteomic analysis, although β-galactosidase staining showed a more widespread expression of Runx1 in lung epithelia as expected (Figure 6B).

To simultaneously visualize actual expression of Runx1 and mark descendants of Runx1^{+} cells, the *Runx1*$^{-2A\ YFP-2A\ tTA}$ knock-in allele was crossed to *tetO*$_{bi}^{luc/Cre\ pos}$ and the Cre-activatable *Rosa26-lacZ* reporter to establish a triple transgenic Runx1-Tracer mouse (Figure 6C). Runx1^{+} descendants (blue cells) appeared ubiquitously in the airway epithelium, suggesting that expression of Runx1 is not limited to a specific epithelial cell lineage (Figure 6D).

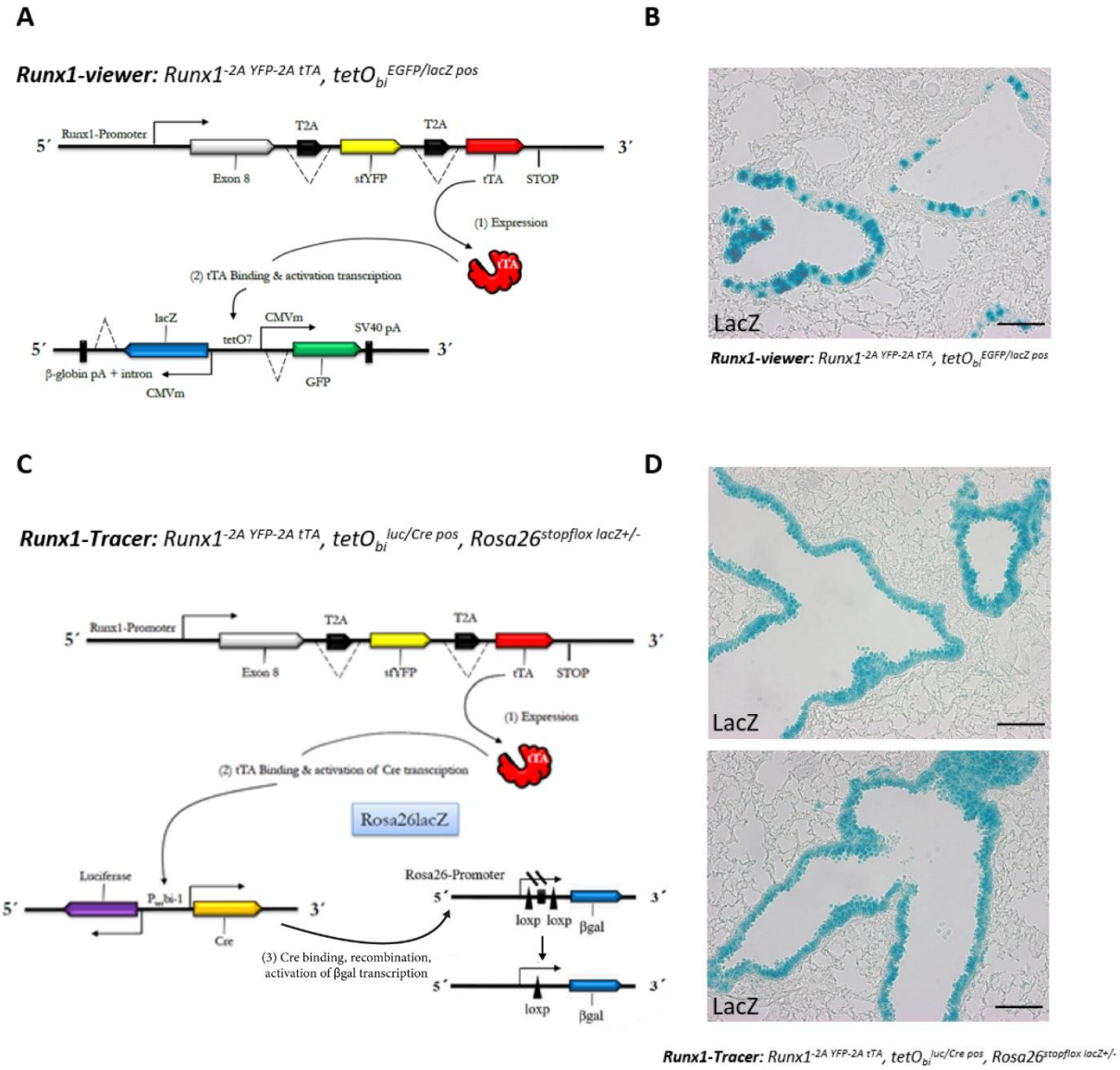


Figure 6: Actual expression and genetic lineage tracing of Runx1⁺ cells in adult lung epithelia.

Schematic description of the generation of Runx1-viewer (**A**) and Runx1-Tracer (**C**) knock-in mice. To generate the *Runx1*$^{-T2A\ YFP-\ T2A\ tTA}$ knock-in allele, a targeting vector was designed with a T2A YFP- T2A tTA, inserted before the stop site in the Runx1 locus and crossing with appropriate reporters to produce viewer and tracer strains. (**B**) β-galactosidase staining of lung sections from Runx1-viewer (59 weeks of age, male) and (**D**) Runx1-tracer (49 weeks of age, male) animals; scale bar: 100µm. Schematic images provided by Dr. Jennifer Kulhei.

To identify cell types expressing Runx1⁺ cells, immunostaining for Runx1 was performed n combination with antibodies against CCSP (Club cell marker) (Figure 7A) and α-tubulin (ciliated cell maker) (Figure 7B). The staining indicated heterogeneity among Club cells in terms of Runx1 expression and the presence of Runx1 expression in ciliated cells. The results from

lineage tracing and immunostaining indicated that Runx1 is not a suitable marker to identify distinct stem/progenitor subsets in the adult lung epithelium.

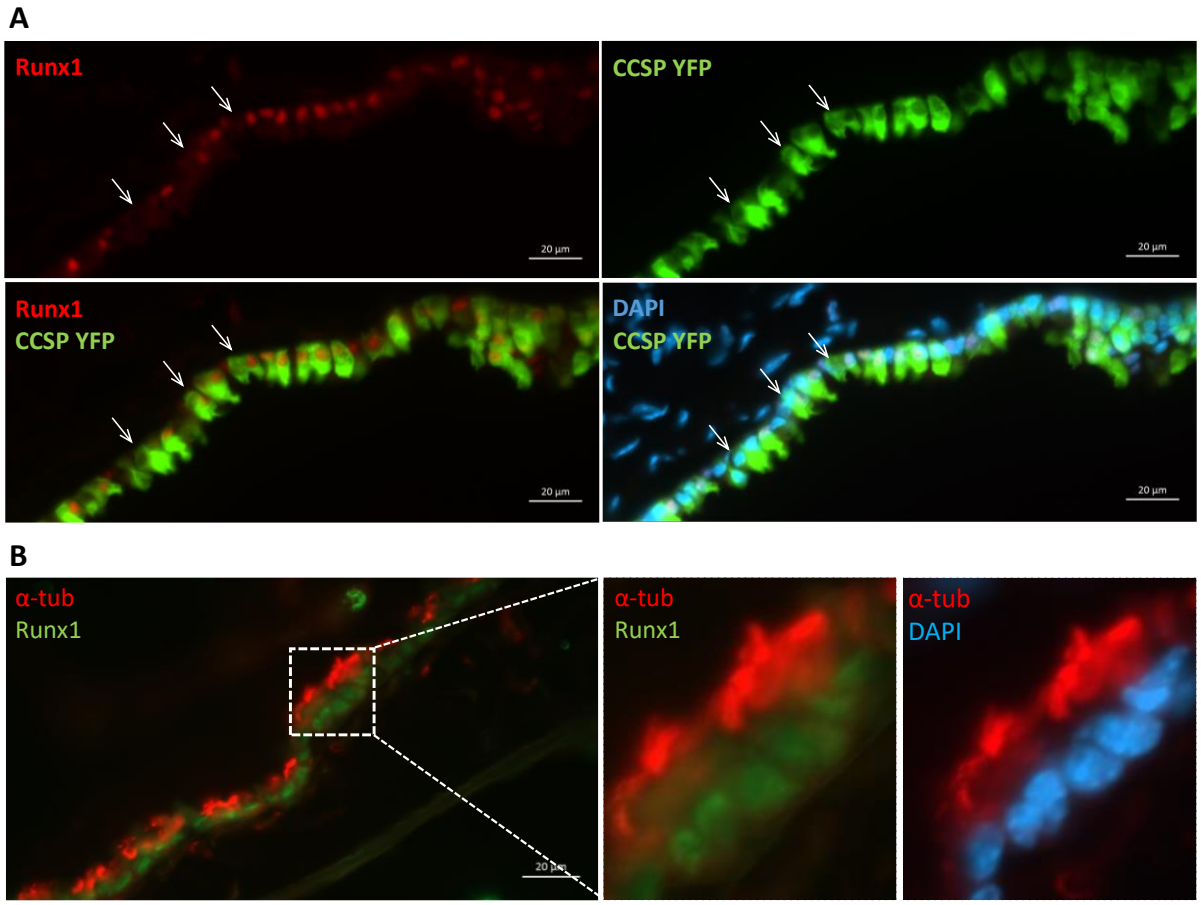


Figure 7: Distribution of Runx1 expression in both Club (A) and ciliated cells (B) of mouse lung epithelium.

(**A**) Immunostaining for Runx1 (red) on lung sections. White arrows indicate Club cells lacking expression of Runx1 in the nucleus. (**B**) Immunostaining for Runx1 (green) and α-tubulin as a marker of ciliated cells in (red).

6.1.2 Runx1 is dispensable for airway regeneration after naphthalene-induced injury

To further understand the role of Runx1 in the airway epithelium, conditional Runx1 mutants were generating using Nkx-2.1-Cre mice [72], which express Cre in the lung anlagen. Analysis of lung-specific Runx1 null mice (Runx1 KO) revealed that Runx1 is not critical for lung development. No embryonic lethality was observed and mutant animals lived through adulthood. Immunostaining of Runx1 on lung sections of WT and Runx1 KO mice (Figure 8) verified the complete absence of Runx1 expression in KO samples and confirmed widespread expression of Runx1 in WT lung epithelium. As expected, some cells in the lung parenchyma (alveolar compartment) stained positive for Runx1, probably representing immune cells,

which is in line with previous studies, reporting expression of Runx1 in myeloid progenitor cells [73, 74] (Figure 8).

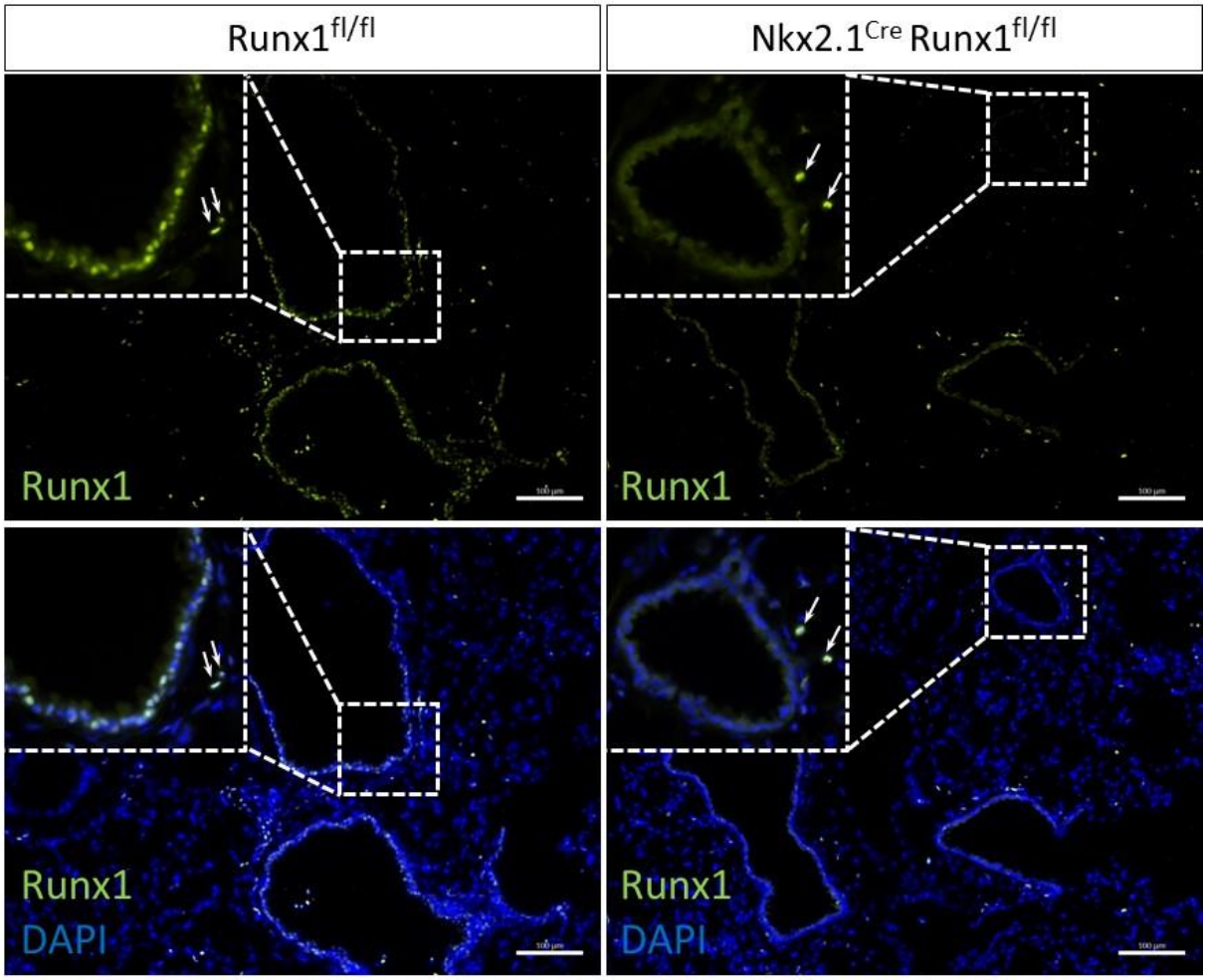


Figure 8: Immunofluorescence staining of Runx1 on lung sections of WT and Runx1 knockout mice.

Runx1 (green) is expressed abundantly in lung epithelium of WT mice, but absent in Nkx2.1^{Cre} Runx1$^{fl/fl}$ mutants. A few Runx1+ cells are visible outside the epithelium, presumably representing myeloid cells. Scale bar: 100µm.

To understand whether Runx1 might play a role for lung regeneration but not for lung development, Runx1 KO animals were treated with naphthalene to ablate Club cells. Morphological analysis did not uncover any significant differences of regenerated lung tissue between naphthalene-treated Runx1 KO and WT animals (Figure 9). Regeneration of the lung epithelium was normal 21 days post naphthalene treatment. No differences between WT and KO mice were seen in the extent of patches, which were still devoid of Club cells at this time point, indicating absence of any delay in airway regeneration in Runx1 KO mice (Figure 10).

This observation suggests that Runx1 is dispensable both for lung development and regeneration, despite its abundant expression in bronchiolar epithelial cells.

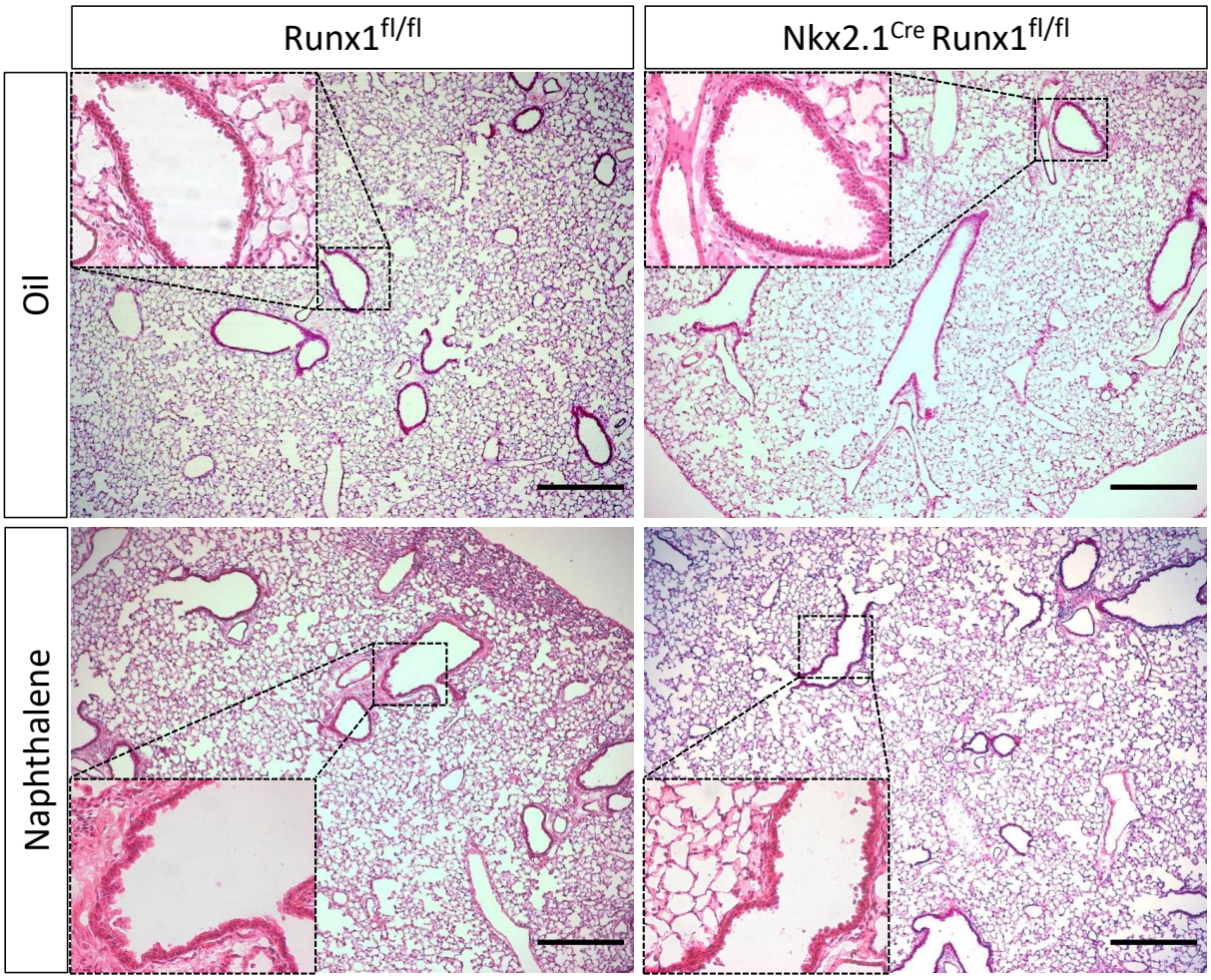


Figure 9: Naphthalene-induced injury in WT and Runx1-null mice.

Adult mice (10-12 weeks old) were either untreated or treated with a single injection of naphthalene in oil at 200 mg/kg i.p. for WT and Runx1 KO ($Nkx2.1^{Cre}Runx1^{fl/fl}$) mice. Lung tissues were collected for histopathological examination 21 days post naphthalene treatment. Representative hematoxylin and eosin-stained sections are shown. Scale bar: 100µm.

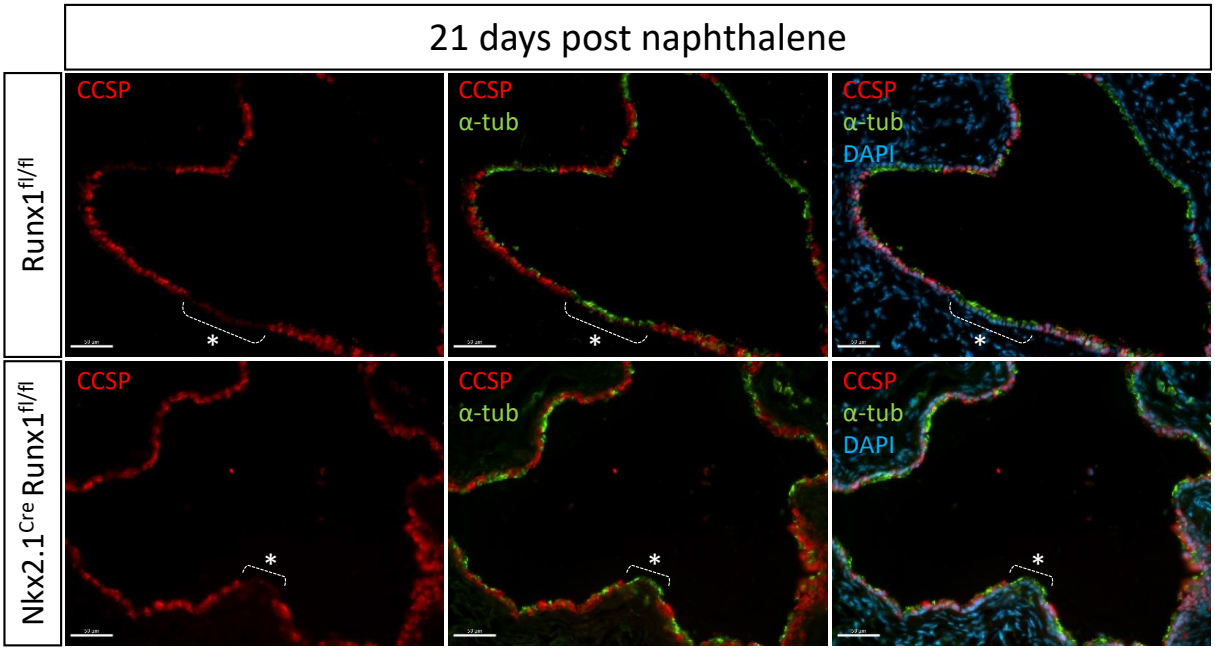


Figure 10: Immunofluorescence staining of WT and Runx1 KO for Club cell marker (CCSP) and Ciliated cell marker (α-tubulin) 21 days post naphthalene.

Stars label incomplete epithelial regeneration after naphthalene treatment, indicated by absence of CCSP expression. No differences between WT and Runx1 KO mice are apparent. Scale bar: 50µm.

6.2 Single cell RNA sequencing identifies a distinct sub population of CCSP^{+} cells, expressing ciliated cell marker genes.

In searching for unique markers of ^{v}Club cells, mice were treated with naphthalene and naphthalene-resistant Club cells were isolated by FACS. In pursuit of this strategy, mice were used in which the fluorescent reporters mCherry (CCSP^{mCherry}) and YFP (SPC^{YFP}) had been inserted into the CCSP and SPC loci, respectively (Figure 11). The rationale behind this approach was to enrich for resistant ^{v}Club cells, since all regular Club cells will be depleted by naphthalene treatment and only resistant ones will remain that can be visualized by mCherry expression in FACS experiments. Due to the co-expression of mCherry and YFP reporter proteins, BASCs, which are also naphthalene resistant, can be excluded in this mouse model.

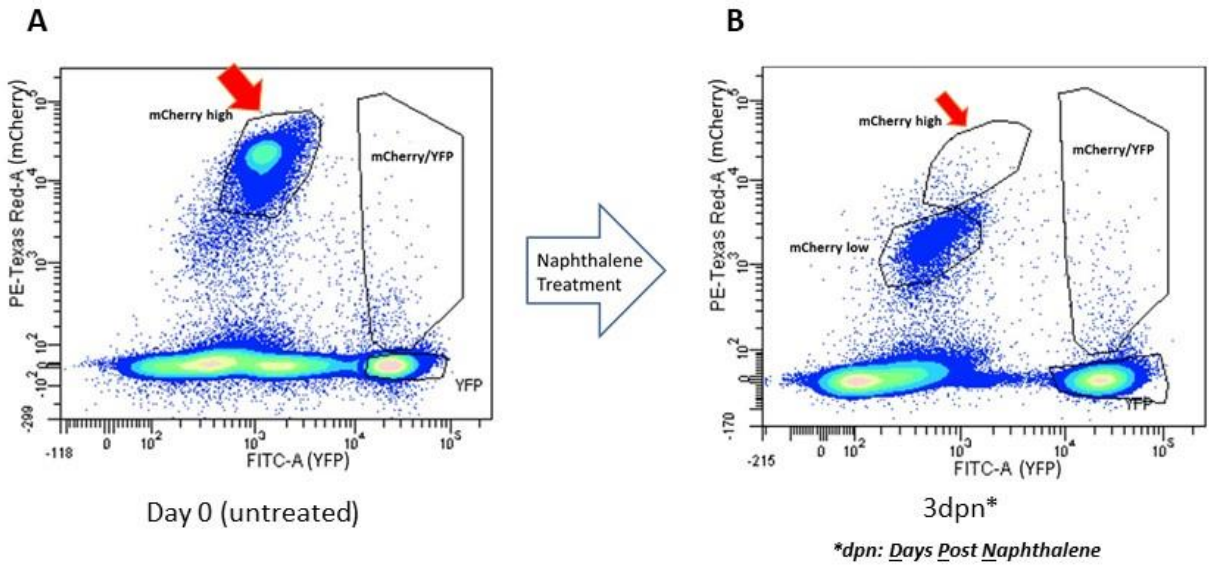


Figure 11: Depletion of Club cell population three days after naphthalene treatment.

Representative FACS plots of cells isolated from lungs of untreated (**A**) versus treated (3dpn) (**B**) mCherry^{+}/YFP^{+} reporter mice. Red arrows in both FACS plots point to the mCherry high gate, containing ^{v}Club cells. Gatings for mCherry^{+}/YFP^{+} (BASC) and YFP^{+} (AT2) cell populations are shown as well (**A**,**B**).

BASCs, regular Club and AT2 cells and were obtained from untreated CCSP^{mCherry}/SPC^{YFP} mice, whereas ^{v}Club cells were collected 3 days post naphthalene treatment. Isolated cell populations were subjected to transcriptional profiling using Affymetrix DNA microarray hybridization, allowing identification of genes that are differentially expressed between lung epithelial cell populations before and after naphthalene exposure. Due to naphthalene-

induced epithelial damage several genes coding for regulators of immune/inflammatory responses were highly expressed in ^{v}Club cells (Figure 12). The induction of inflammatory responses rendered identification of unique and specific ^{v}Club cell marker genes difficult.

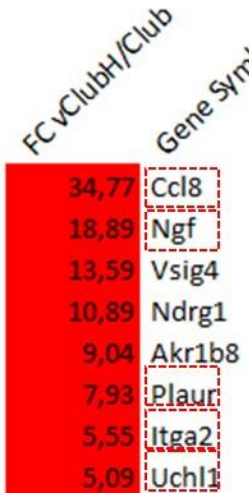


Figure 12: Genes in response to stimulus with the highest fold change comparing the resistant Club cells versus control Club cells from homeostasis.

The genes in red dashed squares show strong increase of expression in ^{v}Club cells after naphthalene treatment compared to untreated regular Club cells and are related to stress or inflammatory responses. Ccl8 (C-C Motif Chemokine Ligand 8), Ngf (Nerve Growth Factor), Plaur (Plasminogen Activator, Urokinase Receptor), Itga2 (Integrin Subunit Alpha 2) and Uchl1 (Ubiquitin C-Terminal Hydrolase L1) as top 5 representative genes.

To avoid artifacts resulting from the upregulation of stress-related genes in vClub cells after naphthalene treatment, single cells analysis of Club cells was performed under homeostatic conditions. This approach enables delineation of heterogeneity of Club cells in homeostasis and was supposed to reveal hidden subpopulations of Club cells and to identify unique markers for branch point-associated Club cells. ICELL8 system (Figure 14- 13) is one of the systems that allows unbiased analysis of large numbers of single cells while providing more control over the selection of the isolated cells.

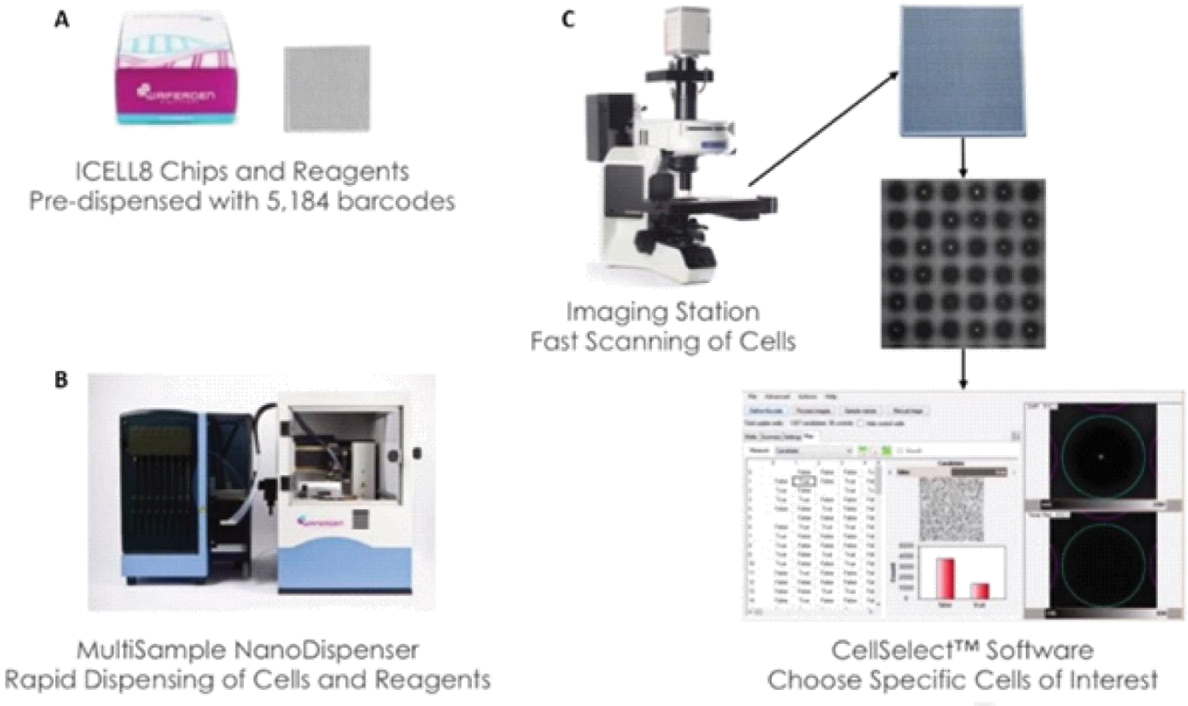


Figure 14: The ICELL8 system and its components.

(**A**) ICELL8 chip and reagents. The nano-well chip contains 5184 barcoded wells, which enables the user to identify wells corresponding to individual sequencing data sets and inspect the morphology of dispensed cells. (**B**) Multisample nano-dispenser. This part of the device takes up cells from the source plate and dispenses them onto the nano-well chip. Loading of each chip takes approximately 15 minutes. (**C**) Imaging station and cell select software. This microscope take images of every well of the nano-well chip in 2 different channels. The time period between the dispensing and the imaging is kept to a minimum. After imaging, the nano-well chip is frozen at -80. Cell selection was carried out via cell select software (adopted from https://www.takarabio.com/learning-centers/automation-systems/icell8-introduction).

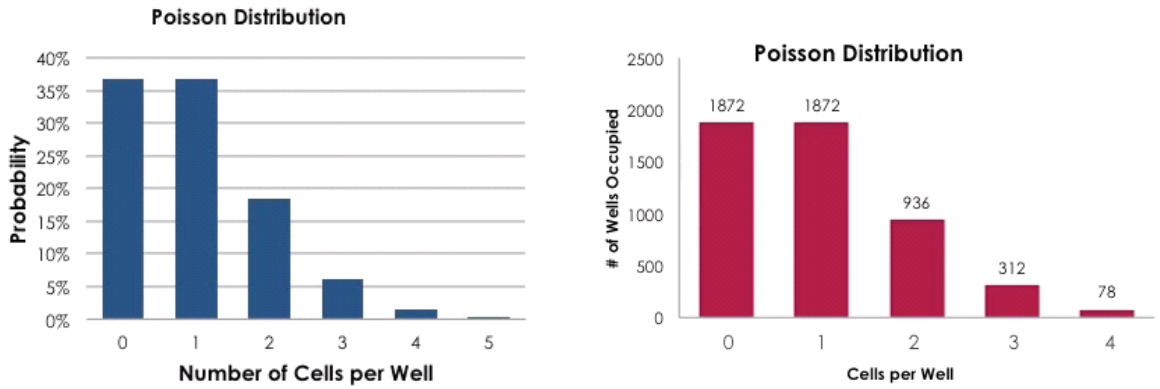
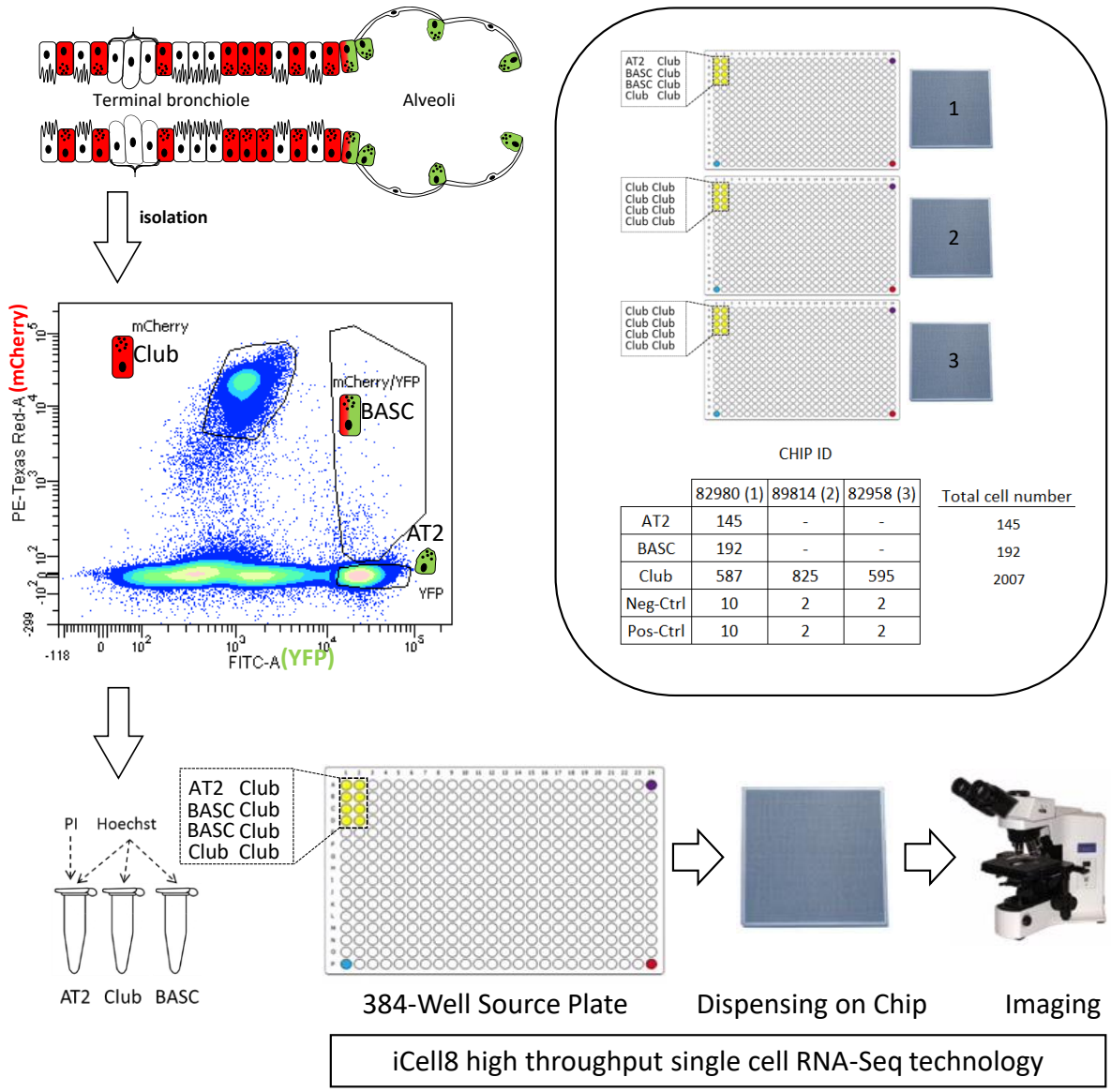


Figure 13: Number of wells containing single cell in the nano-well chip of ICELL8 system follow a Poisson distribution.

Poisson distribution is the probability of a number of events occurring in a fixed period of time if these events occur with a known average rate and are independent in the time since the last event. Cell distribution is the probability of the well containing a single cell, depending on the starting concentration of cells and the reagent dispense rate and volume to the chip. Under normal circumstances, around 37% of wells contain a single cell, which allows sequencing of approximately 1,728 single cells (adopted from WaferGen Biosystems-data sheet).

In total, 2372 lung epithelial cells were sequenced (Figure 15), isolated from $CCSP^{mCherry}SPC^{YFP}$ knock in mouse strain using ICELL8 full-length single-cell RNA-seq system [75] (Figure 15). As illustrated in Figure 15, Club, AT2 and BASC cells were FACS-isolated based on the fluorescent reporter gene expression in each cell type (Club cells $mCherry^{+}$, AT2 cells YFP^{+} and BASC positive for both mCherry and YFP). Since stem cells are rare, the number of chips was increased and some chips were only loaded t with Club cells to increase the chance to identify putative subpopulations of Club cells, (Figure 15). In total 145 AT2 cells, 192 BASC cells and 2007 Club cells were selected from 3 chips and subjected to RNA sequencing. t-distributed stochastic neighbor embedding (t-SNE) clustering of the expression profiles of the 2344 cells that passed quality controls identified three principal cell clusters. Two clusters were characterized by expression of canonical markers for Club and alveolar type 2 (AT2)/BASC (Figure 16). Proximal Club cells were identified by by higher expression of Scgb3a1 and Bpifa1 [76].

In addition, a rare subpopulation of Club cells (cluster number two composed of 72 cells) was identified, which was characterized by co-expression of CCSP (Club cells marker) and Foxj1 (ciliated cells marker), indicating a transitory cell state that share properties of both Club and ciliated cells. This cluster showed a relatively low expression of Cyp2f2 in comparison to the Club cell cluster.



	82980 (1)	89814 (2)	82958 (3)	Total cell number
AT2	145	-	-	145
BASC	192	-	-	192
Club	587	825	595	2007
Neg-Ctrl	10	2	2	
Pos-Ctrl	10	2	2	

Figure 15: Schematic experimental view of the ICELL8 single cell system.

Cells were FACS sorted and Club, AT2 and BASC cells were isolated from adult murine lung (10-12 weeks of age). AT2 cells were stained with PI for live/dead discrimination while this was not possible for Club and BASC cells since they already show red fluorescence due to the presence of mCherry. After calculating the proper concentration of cells to achieve the maximal number of wells loaded with single cells based on the Poisson distribution, cells were loaded to source plate and dispensed based on the defined pattern onto the nanowell chip. The Nanowell chip was covered (sealed with a film) for imaging and frozen immediately after imaging at -80. The inset shows the total number of 3 chips and related cell types that were loaded in addition to the total number of cells obtained from each chip and selected for RNA sequencing.

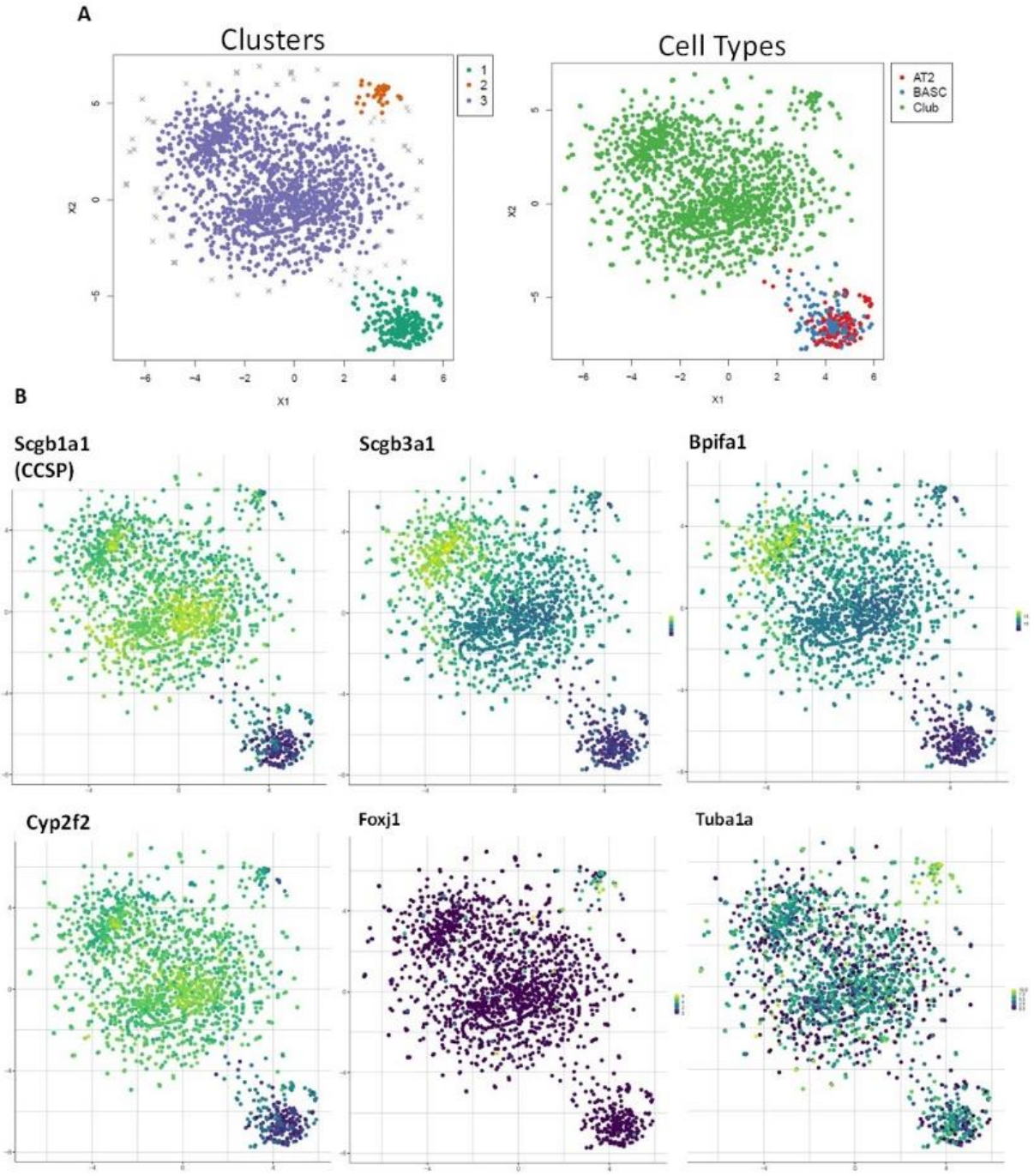


Figure 16: t-SNE plot of 2007 Club cells identifying a sub population CCSP-positive cells with expression of ciliated marker genes.

(**A**) Single-cell transcriptomic sequencing of 2344 cells. t-SNE plot of single cells displays 3 distinct clusters.

(**B**) Selected marker gene expression individually color-coded by the expression of (i) Scgb1a1 (CCSP), (ii) Scgb3a1, (iii) Bpifa1 (iv) Cyp2f2 (v) Foxj1 and (vi) Tuba1a.

6.3 Selection of genes for further characterization of the newly identified CCSP$^+$ Foxj1$^+$ cell cluster

To identify marker genes uniquely expressed in the newly identified cell cluster, gene expression in different cell subsets was compared. 195 genes were differentially expressed between the CCSP$^+$ Foxj1$^+$ cell cluster and the other clusters (Figure 17). Since expression of FoxJ1 in the CCSP$^+$ Foxj1$^+$ cell cluster indicated expression of genes of ciliated cells, the 195 differentially expressed genes (DEGs) from the CCSP$^+$ Foxj1$^+$ cell cluster were related to genes expressed during late (123 genes) and early ciliogenesis (238 genes). Genes characteristic for ciliated cells were derived from a published transcriptome [77] as well as from a data set of mature ciliated cells provided by Prof. Dr. Christos Samakovlis. To further limit the number of candidate genes potentially expressed in the CCSP$^+$ Foxj1$^+$ cell cluster, genes that are expressed during late ciliogenesis (in other words terminally differentiated ciliated cells) were excluded. The focus was on 15 genes that were expressed only during early ciliogenesis (Figure 17A). Furthermore, top marker genes expressed in ciliated cells were excluded, essentially eliminating 174 genes from the list that are highly expressed in mature ciliated cells with the $\log_2$ fold-change cutoff higher than 5. In addition, all marker genes related to both early and late ciliogenesis (from the published transcriptome) were removed, leaving 27 candidate genes that are exclusively expressed in the ICELL8 DEG cluster (Figure 17B). The selection was further narrowed, by excluding genes that were expressed in AT2/BASC or present in the cluster of regular Club cells, which left a final selection of 14 candidate genes. To further characterize the selected candidate genes, a spatial expression analysis was performed by using *in situ* sequencing in the adult mouse lung.

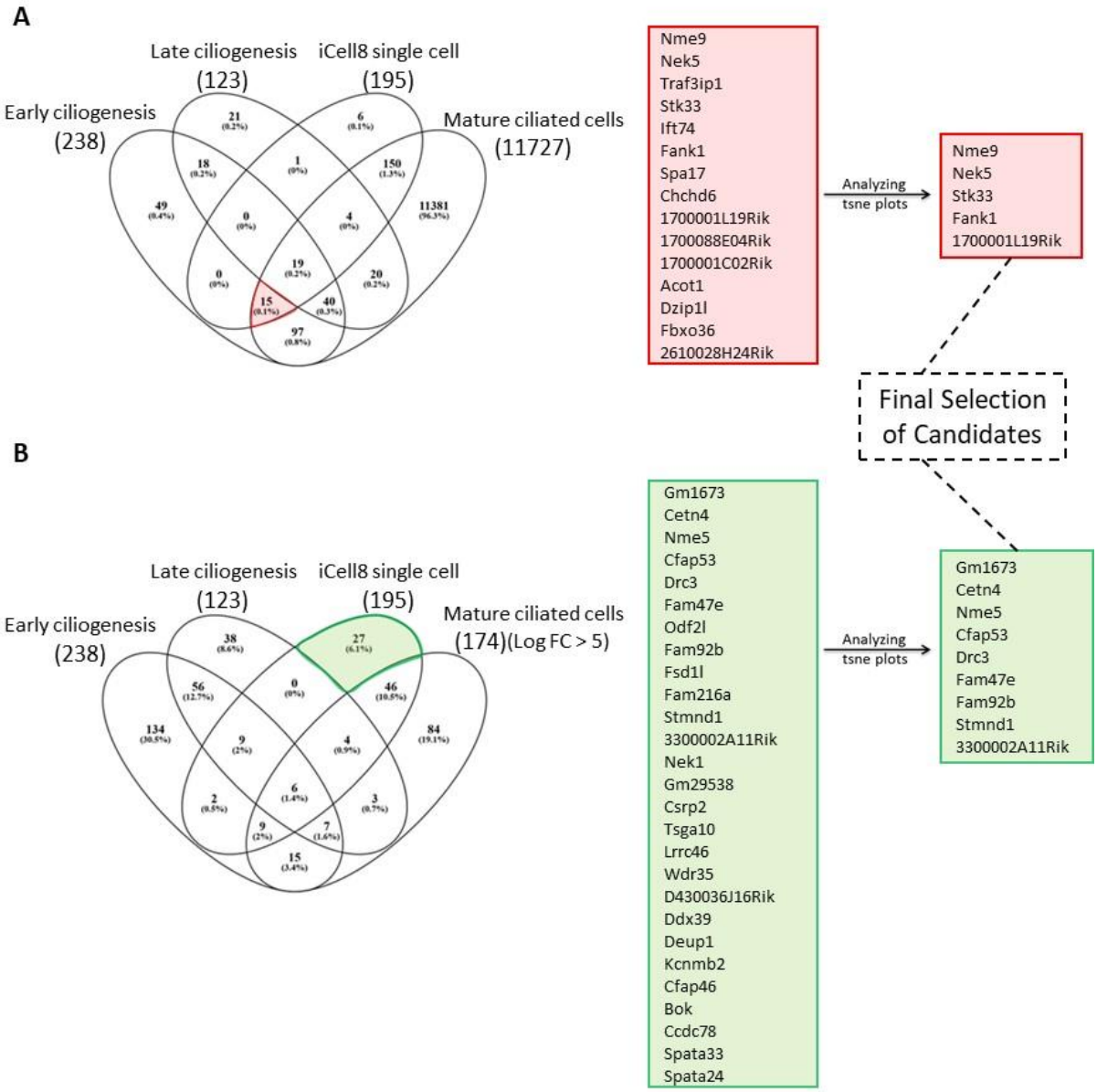


Figure 17: Selection of candidate genes potentially expressed in the CCSP^{+} Foxj1^{+} cell cluster.

(**A**) The Venn diagram shows selected markers based on the intersection of gene expression in different cell clusters. This segment contains 15 candidate gene that are listed in the red box next to the Venn diagram.(**B**) The Venn diagram shows marker gene selection based on the genes that are present only in the newly identified cluster of Club cells and not present in any of the ciliated lineage transcriptome data sets. This segment contains 27 candidate genes, which are listed in the green box next to the Venn diagram. These resulting 42 candidates were further reduced by excluding genes that showed expression in AT2/BASC or regular Club cells. Eventually, 14 candidate genes were selected for further analysis (smaller red and green boxes).

6.4 *In situ* sequencing to study spatial expression of selected candidate genes.

In situ sequencing is a versatile method to visualize the spatial expression of any gene of interest i, also offering the possibility of multiplexing [78]. In this technique, fluorescently-labeled oligonucleotide probes of each mRNA of interest is used to mark individual RNA molecules with a discrete, diffraction-limited punctum that can be quantitatively analyzed by fluorescence microscopy. 3 different padlock probes were designed for each of the 15 candidate genes, plus probes for Cyp2f2, Scgb1a1, Foxj1 and CGRP (57 probes in total) (Figure 18). All probes were multiplexed and the selected probed were added to sections from the

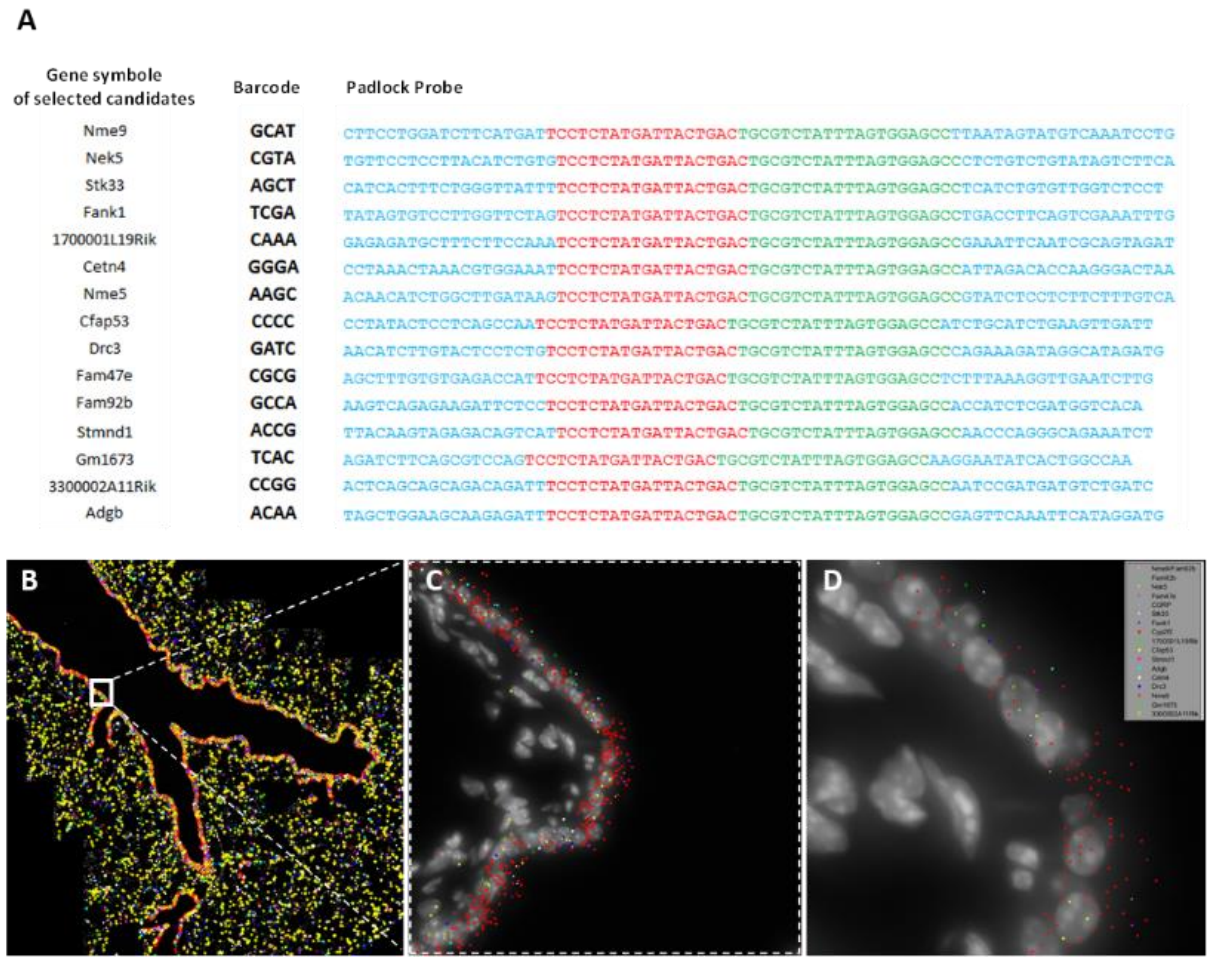


Figure 18: Pad lock probes sequence for the 15 marker gene candidates.

(**A**) Each probe contains a unique barcode and 3' and 5'hybridization arms (blue) in addition to an anchor probe sequence (green). (**B**) Left lobe section of the lung after hybridization, imaging and processing of all 15 candidates. (**C**) Higher magnification of the squared area in image B. Color coding for each signal is represented in the gray box in (**D**).

left lobe of the adult mouse lung and imaged after each round of hybridization. Results after interpreting the fluorescent signal are presented in Figure 18B and Figure 19.

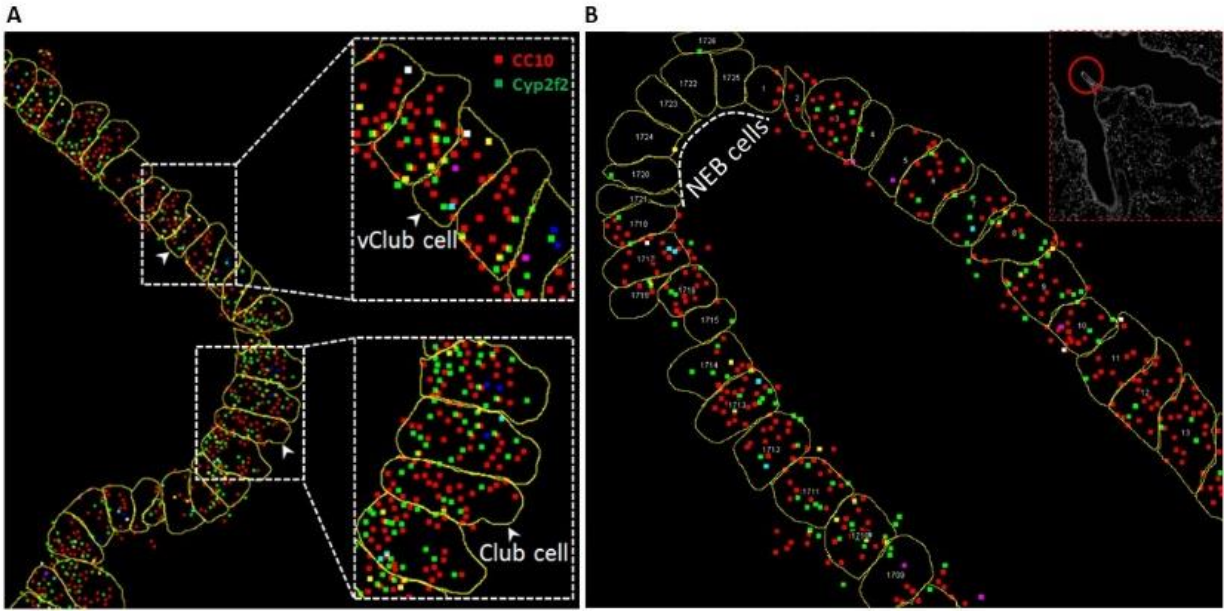


Figure 19: Representative images lung sections after lectin staining and in situ sequencing of a selection of gene of interest (CCSP and Cyp2f2).

(**A**) In this panel regular Club cells can be observed as they highly expressed both CCSP and cyp2f2 marker genes while in the upper white dashed square a ^{v}club cell can be observed with lower expression of Cyp2f2. (**B**) In this panel one branch point is circled in red and in the zoom is marked as NEB cell area. NEB cell area is denude of any CCSP, Cyp2f2 and Foxj1 expression. Color coding for both panels: CCSP (red), Cyp2f2 (green), Foxj1 (blue), gm1673 (white), Nme9 (cyan), 3300002A11Rik (magnta), Stk33 (yellow).

One of the indications that probes were highly selective, was the clear distinction between NEB and neighboring cells. None of the other probes, except for CGRP, violated this criterium (Figure 19B). At branch points, there was a clear distinction between Cyp2f2^{low} and Cyp2f2^{high} cells (Figure 19A). Analysis of candidate gene expression in Cyp2f2^{Low} expressing cells was hampered by the fact that the probe for Foxj1 did not work efficiently, which prohibited identification of CCSP^{+} Foxj1^{+} (and Cyp2f2^{Low}) co-expressing cells by *in situ* sequencing. No marker gene was recognized, which exclusively labelled Cyp2f2^{Low} cells, although such cells could be identified by a combination of two or more markers.

6.5 Generation and characterization of a knock-in Cyp2f2^{CFP} mouse to discriminate Cyp2f2^{Low} from Cyp2f2^{High} cells in the mouse lung epithelium

Since ^{v}Club cells show a low expression of Cyp2f2 compared to regular Club cells, which is the reason why ^{v}Club cells are resistant to naphthalene [79], a Cyp2f2^{CFP} knock-in mouse line was generated (Figure 20-22). Due to the differential expression of Cyp2f2 in ^{v}Club and regular Club cells, the Cyp2f2^{CFP} knock-in mouse should allow FACS-based isolation of CFP^{low} cells from the airway epithelium, which should comprise naphthalene-resistant BASCs and ^{v}Club cells. Cyp2f2^{CFP} knock-in mice were generated by inserting a *T2A-CFP-Neo^{R}* (Neomycin resistance) gene cassette immediately before the stop codon in the tenth exon of *Cyp2f2* gene (Figure 20). The targeting construct contained a *T2A-CFP-LoxP (Cre recombinase recognition target)-Neo^{R}-loxP* knock-in cassette flanked by homology arms as illustrated in Figure 20.

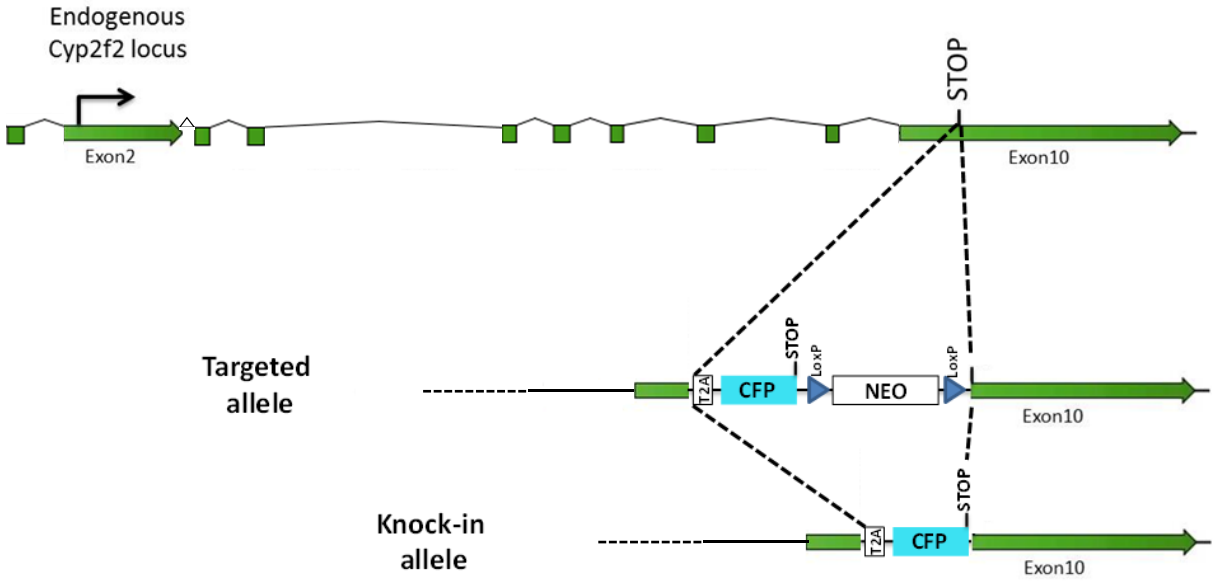


Figure 20: Generation of *Cyp2f2* $^{T2A\ CFP}$ knock-in mice.

The targeting vector carried T2A CFP and Neomycin (NEO) resistance cassettes for insertion at the endogenous stop codon in the Cyp2f2 locus. The NEO cassette, flanked by loxP sites, was excised in the targeted allele by crossing F1 progeny with CMV-Cre deleter mouse strain, resulting in the *Cyp2f2* $^{T2A\ CFP}$ knock-in allele.

The linearized targeting construct was electroporated into embryonic stem cells (ESCs) of C57BL/6 mice. Neomycin resistant ESC clones were expanded and screened for homologous recombination at both 3' and 5' ends by polymerase chain reaction (PCR). Correctly targeted ES cells were injected into mouse blastocytes and the resulting chimeras were crossed to CMV-Cre deleter transgenic mice, which removed the *Neo^{R}* cassette to create the *Cyp2f2^{CFP}* knock-

in allele by Cre/LoxP recombination (Figure 20). No obvious abnormality was detected in the Cyp2f2^{CFP} knock-in mouse strain. Viability and fertility of heterozygous and homozygous Cyp2f2^{CFP} mice was normal as well.

Expression of CFP was assessed in adult lungs of Cyp2f2^{CFP} knock-in mice. Reporter expression was observed only in airways and not in the alveolar compartment by fluorescence microscopy on lung sections, which is in agreement with the known low gene expression level of Cyp2f2 in ^{v}Club, ciliated, BASC and AT2 cells [80], (Figure 21A).

To identify and separate Cyp2f2^{low} ^{v}Club cells from regular Club cells and other epithelial cells with low Cyp2f2 low expression, the Cyp2f2^{CFP} strain was crossed with a reporter strain, which contains both the SPC^{YFP} and CCSP^{mCherry} knock-in alleles (previously established and published from our lab by Salwig *et al*. [55]). In this way Cyp2f2^{CFPLow}SPC^{YFP} AT2 cells and Cyp2f2^{CFPLow}CCSP^{mCherry}SPC^{YFP} BASC cells can be excluded from total fraction of Cyp2f2^{CFPLow} cells, whereas CCSP^{mCherry}Cyp2f2^{CFPHigh} and/or CCSP^{mCherry}Cyp2f2^{CFPlow} cells can be discriminated and isolated (Figure 22). Verification of Cyp2f2 and Foxj1 expression by quantitative real-time PCR (qRT-PCR)-based analysis in FACS-isolated populations of Club^{CFPLow}, Club^{CFPHigh}, AT2 and BASC confirmed that different epithelial cell types can be distinguished based on differences of the endogenous fluorescence reporter proteins (CFP, YFP and mCherry) (Figure 22D,E). Club^{CFPHigh} showed the highest expression of Cyp2f2, consistent with a high expression of this cytochrome in regular non-naphthalene resistant Club cells. Club^{CFPLow} cells showed lower Cyp2f2 expression in comparison to Club^{CFPHigh} but Cyp2f2 expression was higher compared to BASC and AT2 cells. Cyp2f2 expression was lowest in alveolar AT2 cells (Figure 22D). Interestingly, Club^{CFPLow} showed the highest Foxj1 expression, supporting the conclusion that CCSP^{+} Foxj1^{+} cells represent a naphthalene-resistant subpopulation of Club cells (Figure 22E).

In contrast to the CCSP^{mCherry}Cyp2f2^{CFP} alleles, which can be robustly used for flow cytometry analysis and cell isolation by FACS, detection of CFP expression was challenging with fluorescence microscopy. No clear difference was visible between Cyp2f2^{CFPHigh} versus Cyp2f2^{CFPLow} in bronchiolar epithelial cells, preventing unambiguous identification of CCSP^{mCherry}Cyp2f2^{CFPlow} naphthalene-resistant cells by fluorescence microscopy (Figure 21B).

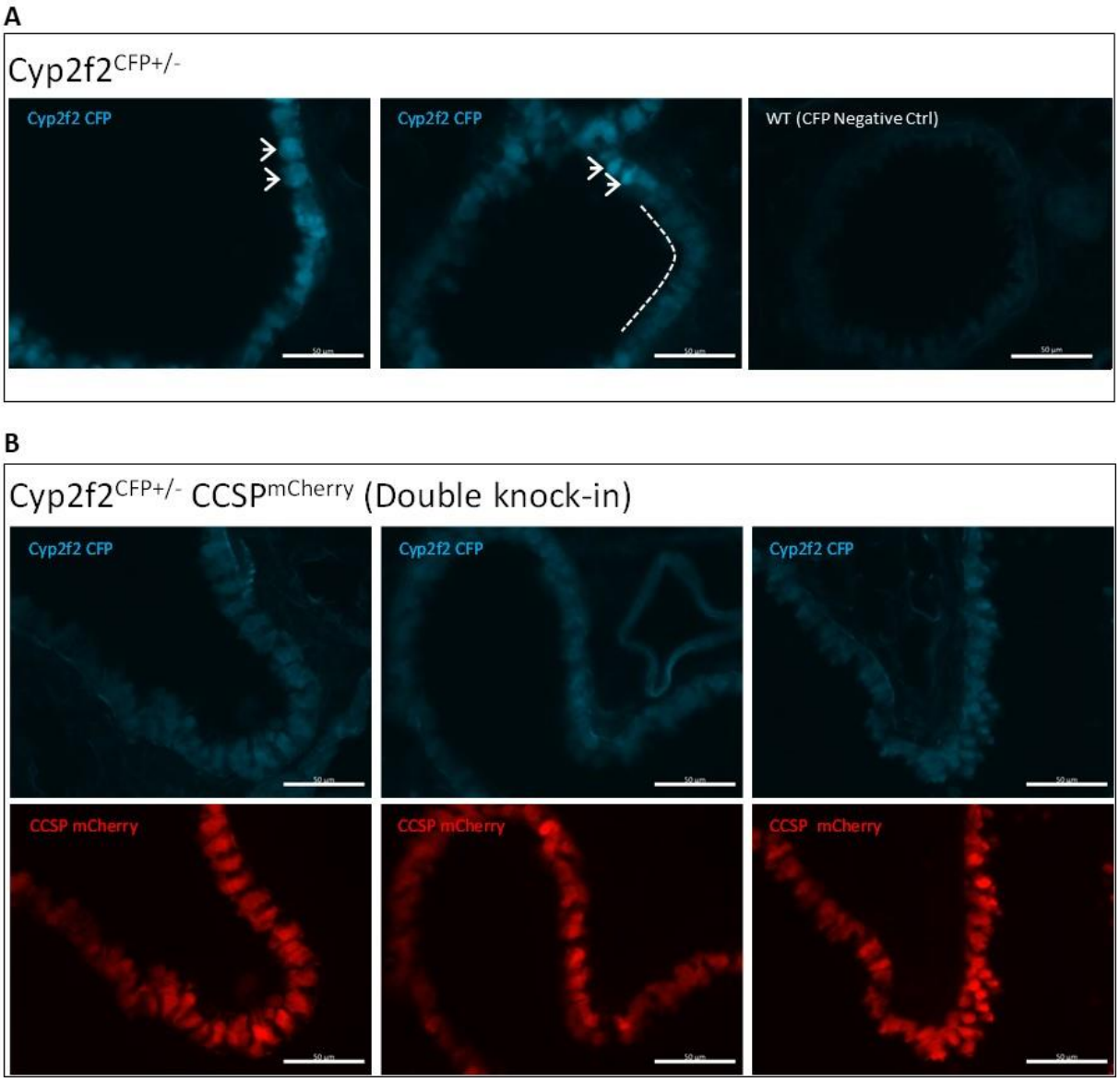


Figure 21: CFP expression from the Cyp2f2 locus in lung epithelial cells.

(**A**) Cross section of the lung epithelium showing CFP^{Pos} cells in the bronchiolar epithelium. White arrows label CFP^{High} cells and the dashed line indicates a group of CFP^{Low} cells. (**B**) Cross section of the lung epithelium of $CCSP^{mCherry}Cyp2f2^{CFP}$ (double knock-in) mice. A clear difference in the expression of CFP is not visible anymore. Scale bar: 50µm.

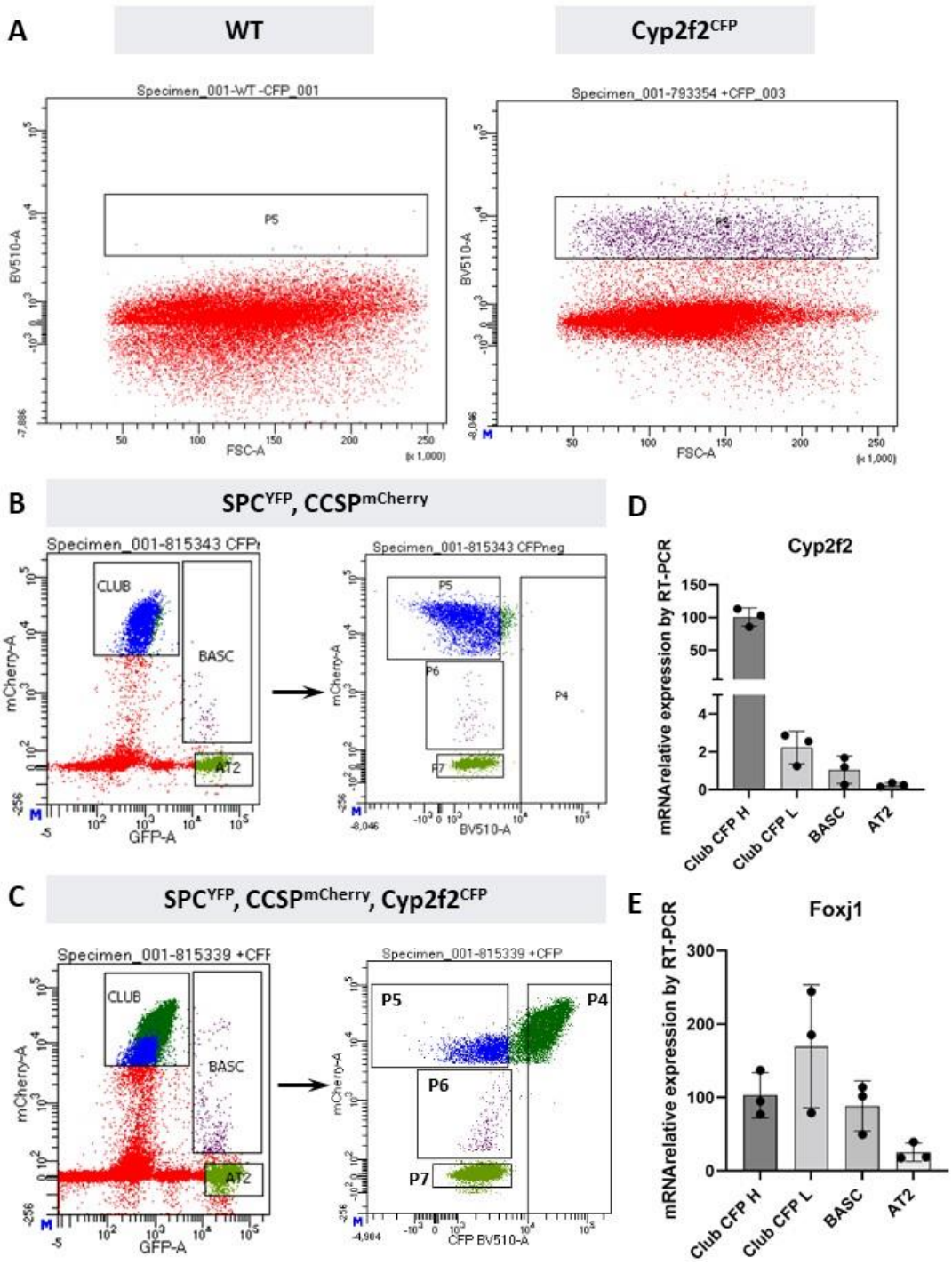


Figure 22: Flowcytometric analysis of airways cells from SPC^{YFP}CCSP^{mCherry}Cyp2f2^{CFP} mice.

(**A**) Flow cytometric analyses of Cyp2f2^{CFP} knock-in reporter from lung epithelial cells of the adult mouse compared to WT (**B**) Flowcytometric analysis of epithelial cells from SPC^{YFP}CCSP^{mCherry} bigenic mouse lacking CFP allele as a negative control showed no expression of CFP. First Club, AT2 and BASC cells are each gated separately based on the expressions level of mCherry and YFP (GFP in the diagram). Second (right panel) are same populations separated based on the expression levels of mCherry and CFP (**C**) Flowcytometric analysis of epithelial cells from SPC^{YFP}CCSP^{mCherry}Cyp2f2^{CFP} triplegenic mouse. In the right panel, same populations are gated based on the expression level of CFP. P4 is gated for CFP^{High} and P5 is gated for CFP^{Low}. P6 and P7 are gated for BASC and AT2 cells respectively. (**D,E**) Expression of Cyp2f2 and Foxj1 was analyzed by RT-qPCR in all the four populations of Club (CFP^{High} and CFP^{Low}), BASC and AT2.

6.6 Transcriptional profiling of *Cyp2f2^{Low}* and *Cyp2f2^{High}* epithelial cell populations

RT-qPCR analysis of Club^{CFPLow} FACS-sorted epithelial cell population uncovered a high expression Foxj1 in the new subpopulation of CCSP^{Pos}Cyp2f2^{Low} cells (Figure 22E), confirming that Club^{CFPLow} cells express marker genes of ciliated cells (Figure 16). To further study transcriptional differences and similarities between Club^{CFPLow}, Club^{CFPHigh}, ciliated, BASC and AT2 cells bulk RNA sequencing (bulk RNA-seq) experiments were performed. To isolate the aforementioned different epithelial cell populations (except ciliated cells), the triple reporter mouse line *SPC^{YFP}CCSP^{mCherry}Cyp2f2^{CFP}* was used, following the sorting strategy depicted in (Figure 22).

To isolate ciliated cells, which may also include Foxj1^{high} Club^{CFPLow} cells, the *Foxj1$^{CreERT2::GFP}$* strain [81] was crossed with *Rosa26$^{stopflox\ YFP+/-}$* reporter mice. To induce activation of reporter genes tamoxifen (tmx) was injected to induce nuclear translocation of Cre recombinase, mediating excision of the stop codon flanked by loxP sites to achieve YFP expression (Figure 23). Adult mice received daily intraperitoneal injection of tmx (once per day for five consecutive days) [81].

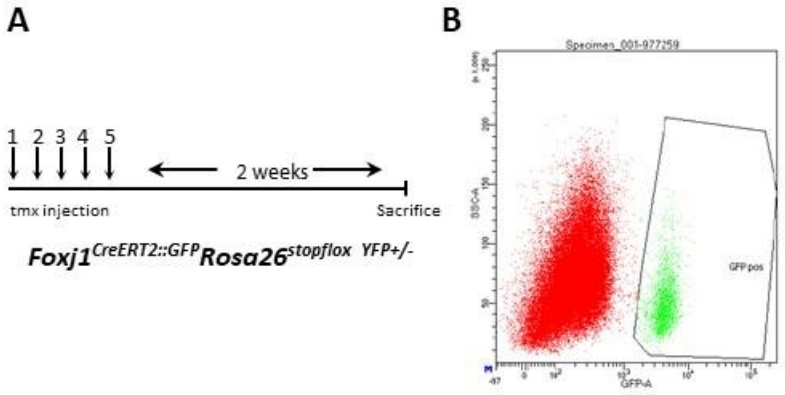


Figure 23: Activation of YFP reporter gene expression in *Foxj1$^{CreERT2::GFP}$Rosa26$^{stopflox\ YFP+/-}$* mice after tmx induction and isolation of YFP-labeled ciliated cells by FACS.

(**A**) Time course of tmx injections. Adult mice, harboring the *Foxj1$^{CreERT2::GFP}$Rosa26$^{stopflox\ YFP+/-}$* knock-in allele, were injected with tmx for 5 consecutive days. Isolation of cells was done 2 weeks after the last tmx injection. (**B**) Representative FACS plot, showing the YFP-based FACS isolation strategy of ciliated cells for transcriptional profiling.

An unbiased, global Spearman correlation analysis of all expressed genes was performed, resulting in a heatmap that is shown in (Figure 24A). Biological replicates clustered well together, confirming reproducibility of the approach. Notably, Cyp2f2^{Low} showed the highest similarity to Club and ciliated cells. Principal component analysis (PCA) confirmed these results, demonstrating that the Cyp2f2^{Low} cell cluster is situated between Cyp2f2^{High} and ciliated cells (Figure 24B).

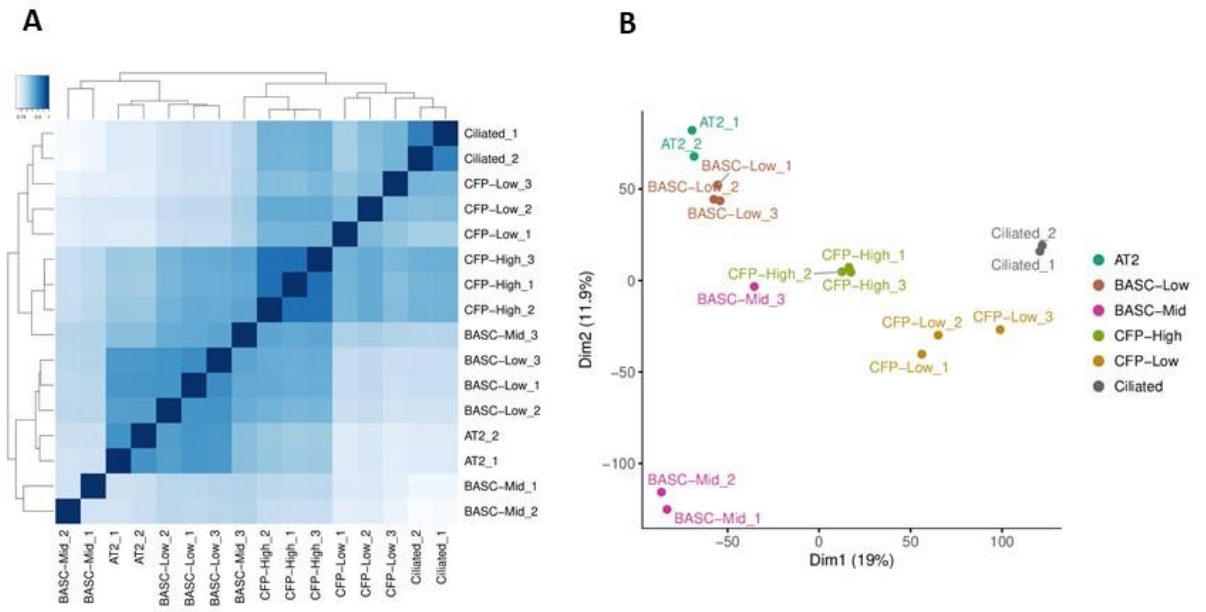


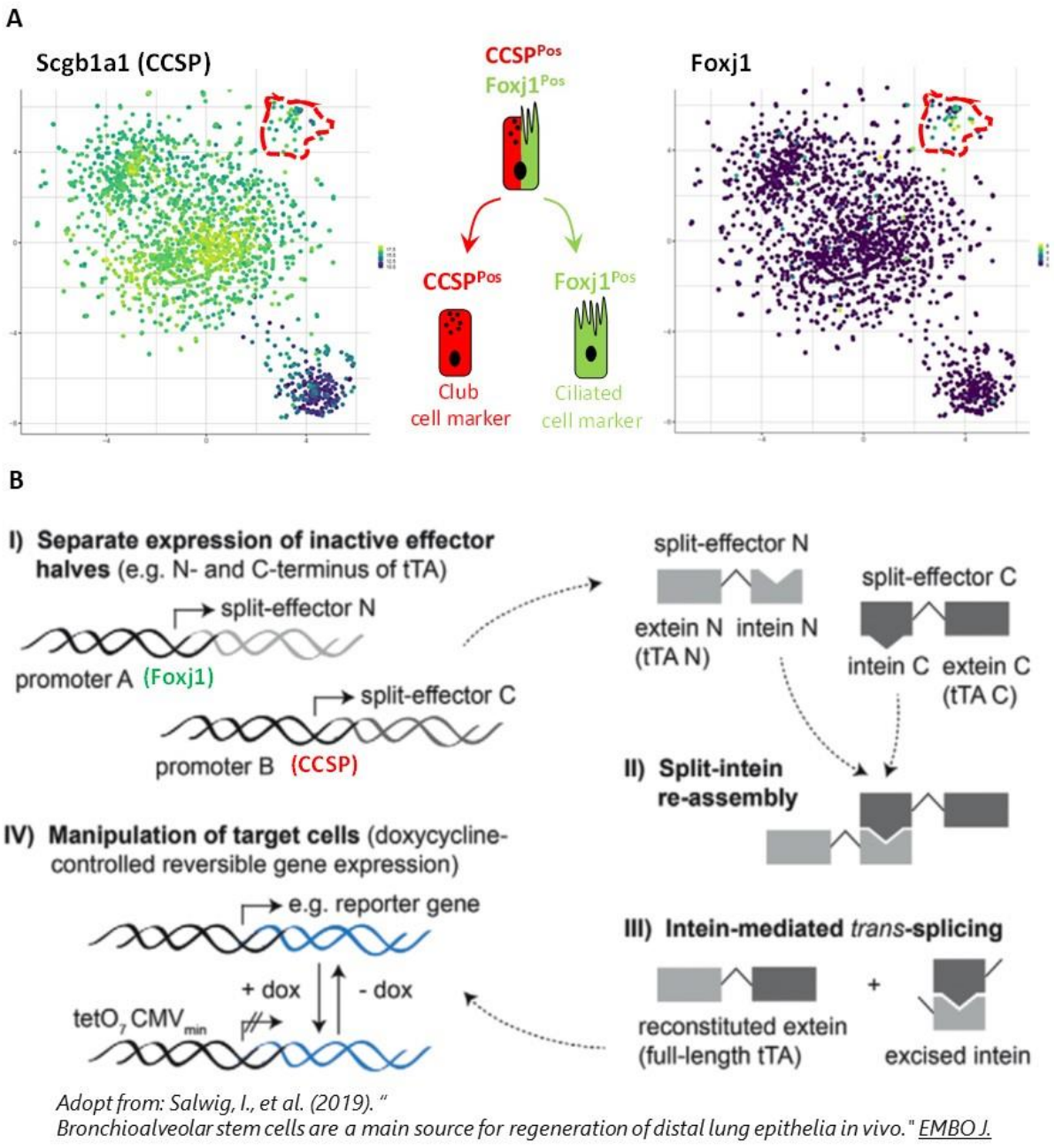
Figure 24: Transcriptional profile of Cyp2f2^{Low}, Cyp2f2^{High}, Ciliated, BASC and AT2 cells.

(**A**) Heatmap of Spearman correlation between all samples and biological replicates. The degree of similarity represented by a colored field ranging from white (0.7) to dark blue (1).

(**B**) Principal component analysis (PCA) of RNA-seq experiments. Percentages represent variance represented by each principal components 1 and 2 in each analysis. Percentages were plotted to visualized similarities among samples. Each epithelial cell population is colored differently and the corresponding replicates of every epithelial cell population are colored the same way.

6.7 Generation of split knock-in mouse strains to target CCSP$^+$ Foxj1$^+$ cells in vivo

To determine the abundance and location of CCSP$^+$ Foxj1$^+$ cells in the lung epithelium and to trace their descendants, a genetic lineage view/tracing system was established to selectively target CCSP$^+$ Foxj1$^+$ cells co-expressing cells *in vivo*. A genetic approach was utilized that was developed previously in our lab by Salwig *et al.*, based on intein-mediated assembly of split effector molecules (in this case tTA) [55] (Figure 25). To use this system, a split *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ knock-in allele was generated by inserting a YFP reporter and N-terminal half of the tTA molecule into the *Foxj1* locus (Figure 26). To confirm expression of the YFP reporter protein under control of the endogenouse *Foxj1* promoter, one of the animals from the first chimeric litter was sacrificed and examined for fluorescent signals in the lung. YFP fluorescence was detected and its expression in ciliated cells was confirmed by staining with Foxj1 and α-tubulin (Figure 27A). Subsequently, the line was crossed to a previously established strain carrying 2A-mCherry fluorescent protein-2A-split-effector C-terminal half inserted into the CCSP locus (CCSP-mCherry). The intein-mediated assembly of split-tTA effector molecules allows selective targeting of CCSP$^+$ Foxj1$^+$ double positive cells *in vivo*. Analysis of the first litter revealed YFP as well as mCherry expression in lung airway epithelium cells (Figure 27).



Adopt from: Salwig, I., et al. (2019). "Bronchioalveolar stem cells are a main source for regeneration of distal lung epithelia in vivo." EMBO J.

Figure 25: t-SNE analysis of single cell RNA sequencing identifies a subpopulation of Club cells, expressing ciliated marker genes.

(**A**) t-SNE plot of CCSP and Foxj1 expression in the second cluster demonstrates co-expression of these two marker genes. The graphical scheme illustrates the existence of an intermediate/stem /progenitor cell type.

(**B**) Schematic overview of strategy for specific *in vivo* manipulation of dual-marker expressing target cells

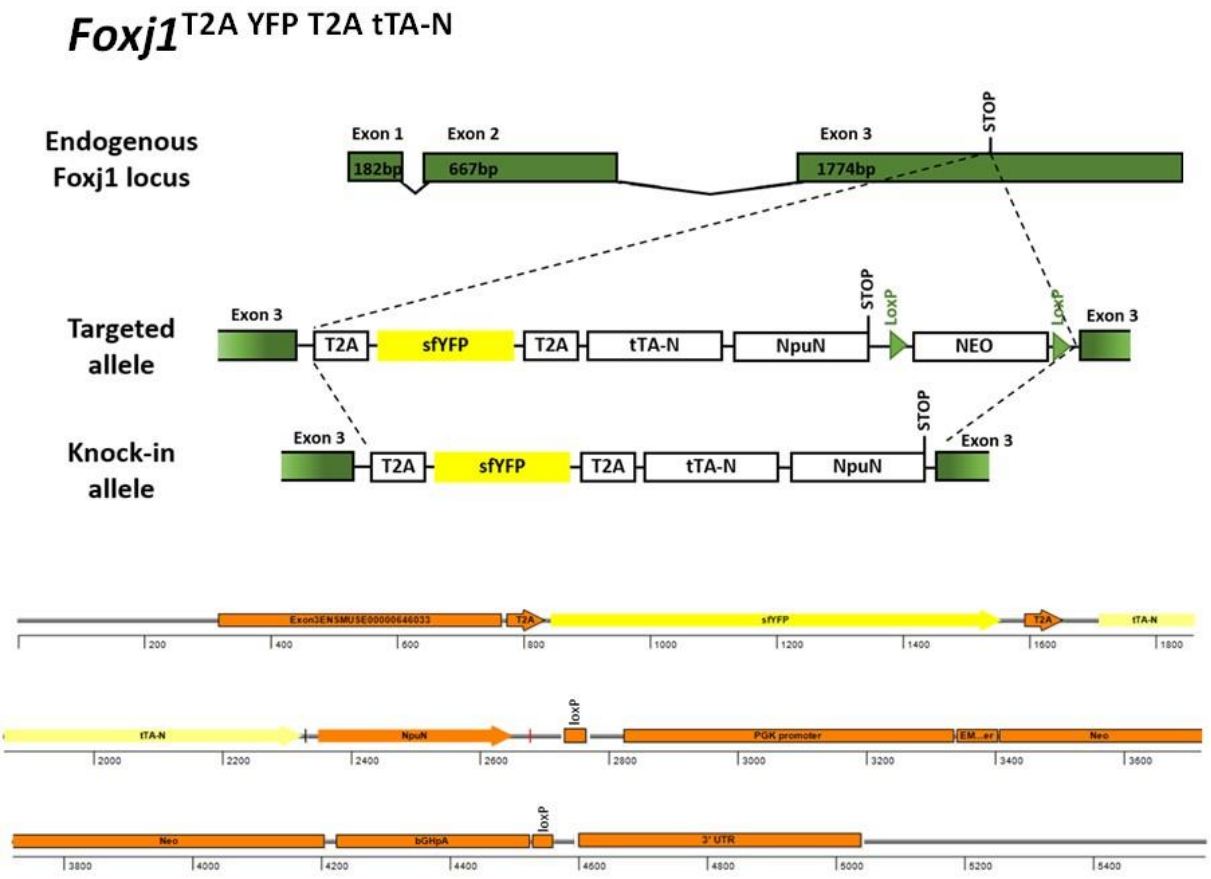


Figure 26: Generation of a split knock-in mouse strain harboring the *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ allele.

A targeting vector was designed, containing a T2A YFP T2A tTA-N and Neomycin (NEO) resistance cassettes, which were inserted into the endogenous stop codon of the *Foxj1* locus. The NEO cassette, flanked by loxP sites, in the targeted allele was excised by crossing F1 progeny with CMV-Cre deleter mouse strain, resulting in the *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ knock-in allele.

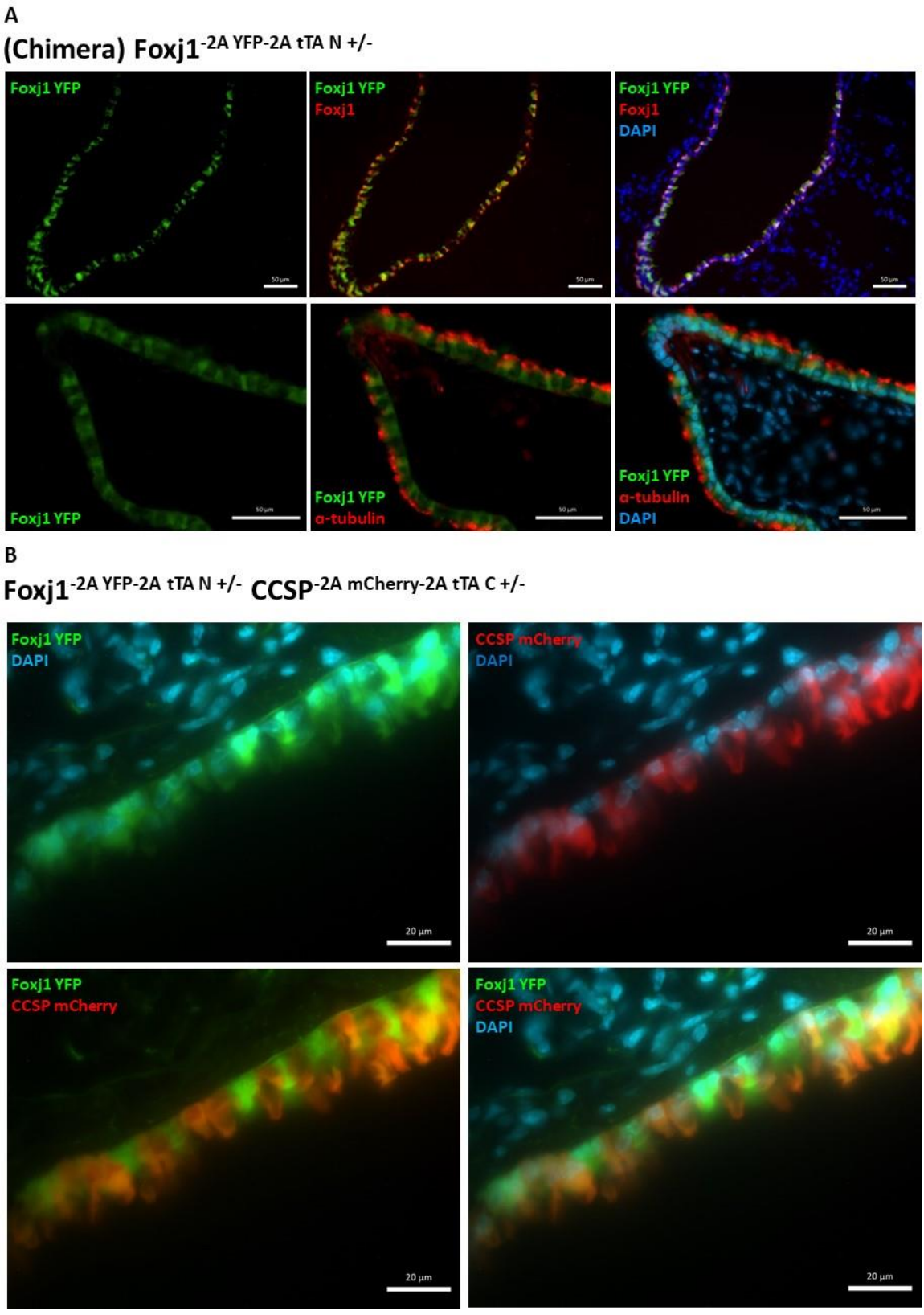


Figure 27: Validation of fluorescence reporter expression in ciliated and Club cells of the adult mouse lung epithelium.

(**A**) Fluorescence images of *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ in adult chimera and expression of the YFP reporter together with Foxj1 (**A** upper panel, red) and α-tubulin (**A** lower panel, red) staining in ciliated cells. Scale bar: 50µm. (**B**) Fluorescence images of *Foxj1*$^{T2A\ YFP\ T2A\ tTA\text{-}N}$ *CCSP*$^{T2A\ mCherry\ T2A\ tTA\text{-}C}$ knock-in strain. Scale bar: 20µm.

6.8 Identification and characterization of CCSP$^+$ Foxj1$^+$ cells and their descendants in healty mouse lungs

To reconstitute the tTA effector molecule and to visualize the anatomical location of CCSP$^+$ Foxj1$^+$ cells *in vivo*, a mouse strain carrying two knock-in allele (*Foxj1$^{-2A\ YFP\text{-}2A\ tTA\ N\ +/-}$, CCSP$^{-2AmCherry\text{-}2A\ tTA\ C\ +/-}$*) was crossed to *tetO$_{bi}^{luc/\ lacZ\ pos}$* reporter mice [82]. CCSP$^+$ Foxj1$^+$ cells (dual marker co-expressing cells) are labeled with LacZ in the resulting mice (CCSP$^+$ Foxj1$^+$ viewer) (Figure 28 (see also Figure 29)). To establish a tool that enables both viewing and permanent labeling of all derivatives of CCSP$^+$ Foxj1$^+$ cells (= CCSP$^+$ Foxj1$^+$ v-race (view + trace)), the mouse strain carrying the two knock-in alleles (*Foxj1$^{-2A\ YFP\text{-}2A\ tTA\ N\ +/-}$* and *CCSP$^{-2AmCherry\text{-}2A\ tTA\ C\ +/-}$*) were crossed to the *tetO$_{bi}^{luc/Cre\ pos}$* [83] and *Rosa26-LacZ* strains [84]. In the resulting mice, the reconstituted split-tTA effector induces expression of Cre recombinase in CCSP$^+$ Foxj1$^+$ cells. Expression of the Cre recombinase mediates permanent excision of the stop codon flanked by loxP sites in the *Rosa26-LacZ* (Figure 28B). Thus, recombination results in permanent labeling of CCSP$^+$ Foxj1$^+$ cells as well as any cells derived from them (Figure 28 and see also Figure 30).

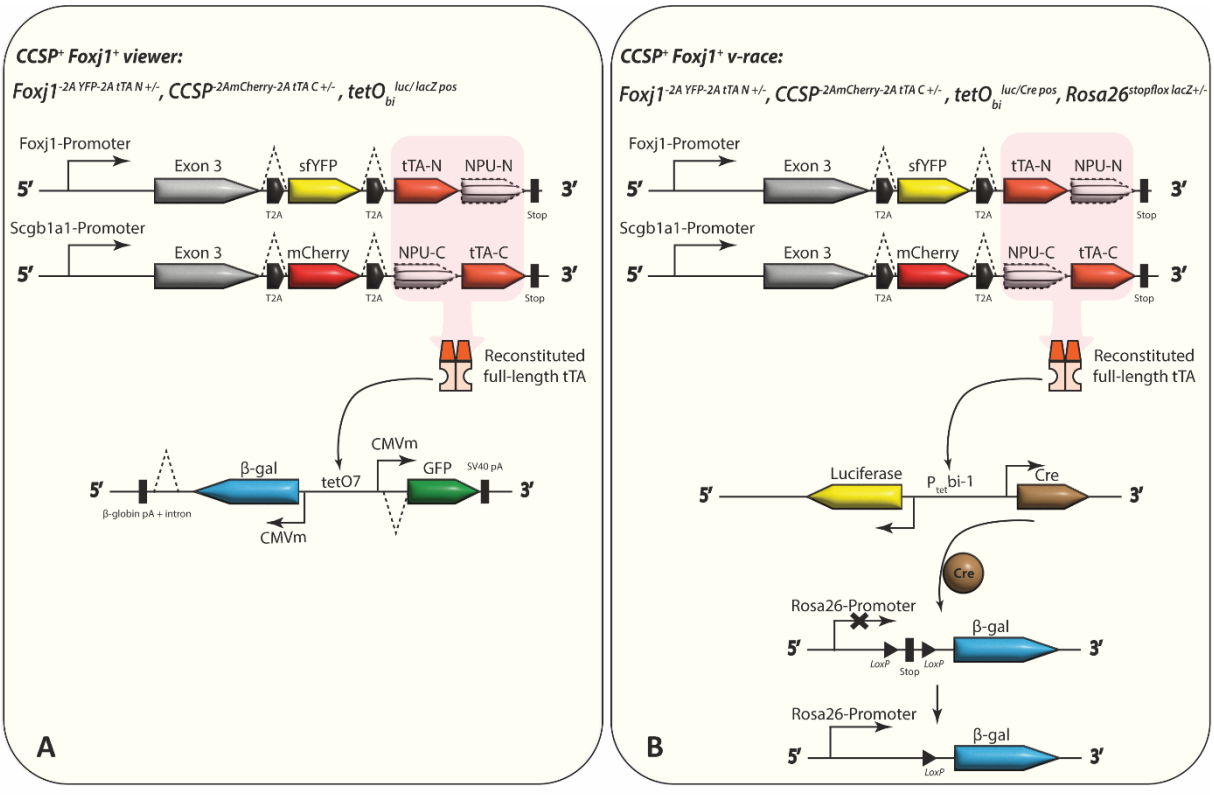


Figure 28: Schematic view of the generation of CCSP⁺ Foxj1⁺ viewer and CCSP⁺ Foxj1⁺ v-race knock-in mouse lines.

(**A**) Schematic outline depicting generation of the CCSP⁺ Foxj1⁺ viewer knock-in mouse line. To generate CCSP⁺ Foxj1⁺ viewer, the mouse strain carrying two knock-in allele (*Foxj1*$^{-2A\ YFP\text{-}2A\ tTA\ N\ +/-}$, *CCSP*$^{-2AmCherry\text{-}2A\ tTA\ C\ +/-}$) was crossed to *tetO*$_{bi}^{luc/\ lacZ\ pos}$ reporter mice, resulting in LacZ-labelling of CCSP⁺ Foxj1⁺ dual marker co-expressing cells.

(**B**) Schematic outline depicting generation of the CCSP⁺ Foxj1⁺ v-race knock-in mouse line. To produce this quadruple knock-in mouse line, mice carrying two knock-in alleles (*Foxj1*$^{-2A\ YFP\text{-}2A\ tTA\ N\ +/-}$and *CCSP*$^{-2AmCherry\text{-}2A\ tTA\ C\ +/-}$) were crossed to *tetO*$_{bi}^{luc/Cre\ pos}$ mice. Reconstituted split-tTA effectors in CCSP⁺ Foxj1⁺ v-race animals induce expression of Cre recombinase, which mediates excision of the stop codon flanked by loxP sites in *Rosa26-LacZ* mice. The recombination results in permanent labeling of CCSP⁺ Foxj1⁺ cells as well as any cells derived from them.

Analysis of the CCSP⁺ Foxj1⁺ viewer strain revealed that the majority of CCSP and Foxj1 co-expressing cells are located at airway branch points (Figure 29), consistent with the proposed location of ᵛClub cells [40, 52, 79]. CCSP⁺ Foxj1⁺ cells were named Branch point-associated Double Positive Cells (BDPCs), due to their characteristic localization. Accordingly, the CCSP⁺ Foxj1⁺ viewer strain was named 'BDPC viewer' and the CCSP⁺ Foxj1⁺ v-race strain 'BDPC v-race'. The frequency of labeled cells that appeared at branch points in BDPC viewer animals is on average 0.31 cell per branch point (n=4) (Figure 29 iv).

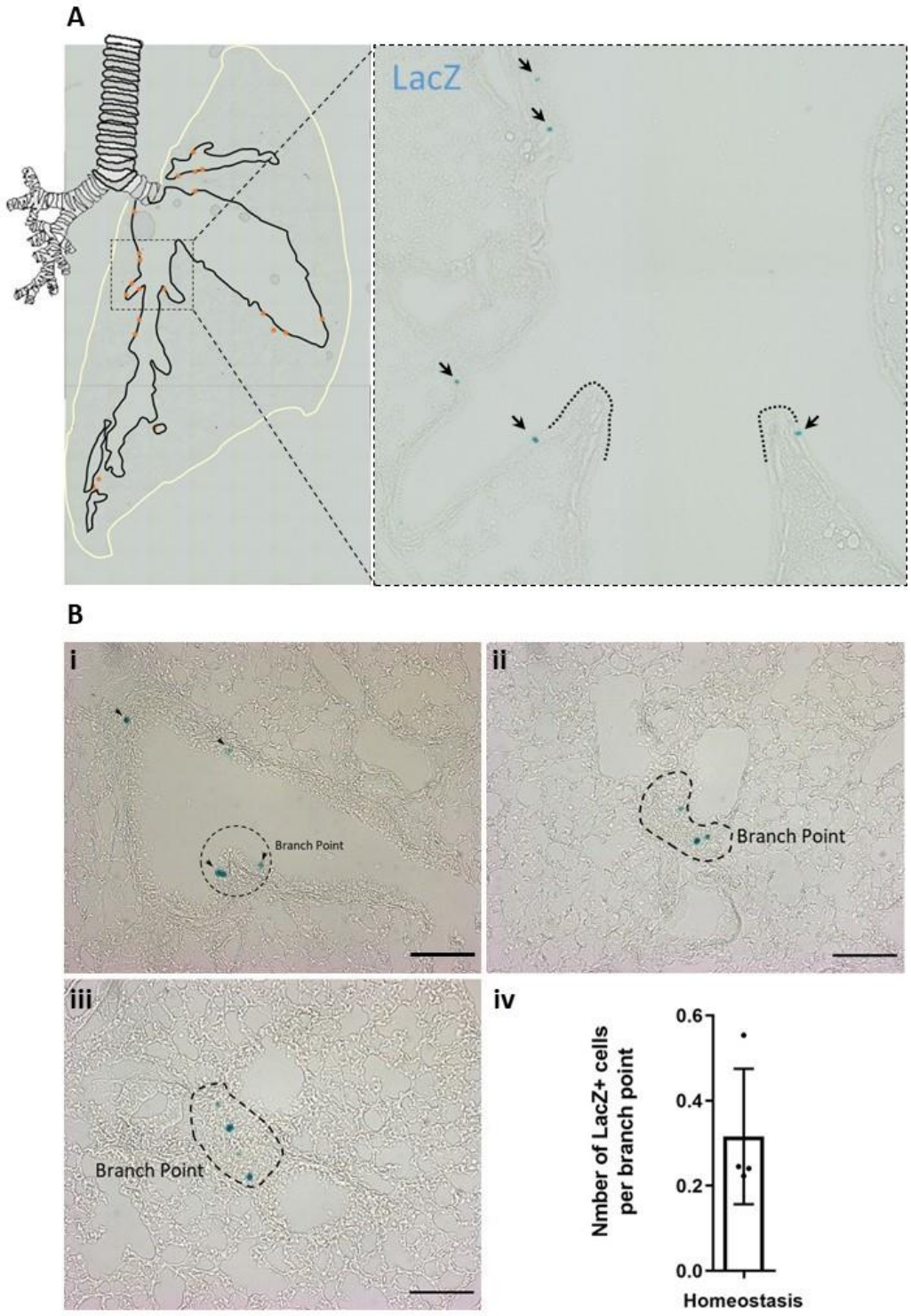


Figure 29: Identification of CCSP$^+$ Foxj1$^+$ co-expressing cells in normal lungs.

β-galactosidase staining of lung sections from BDPC viewer animals at postnatal week 8. Visualization of split-effectors expression from Foxj1 and Scgb1a1 loci identifying BDPCs (arrows) at branch points (dashed lines **A** and dashed circles **B i-iii**).

(**A**) Images of the whole left pulmonary lobe of adult mice. Every red filled circle represents LacZ positive cell. Arrows in the magnified panel on the left and **B** (**i-iii**) indicates that CCSP$^+$ Foxj1$^+$ cells are mainly associated with branch points. (**B**) **vi** Quantification of the number of LacZ$^+$ cells per branch point in BDPC viewer knock-in animals under homeostatic conditions (n=4). Scale bars: 100 µm

Next, the cell fate of BDPCs during lung homeostasis was traced by harvesting lungs from adult mice of the BDPC v-race strain. LacZ signals were identified in the large majority of ciliated cells (Figure 30- 31). Quantification of 14842 FACS purified ciliated and 13847 Club cell populations revealed that 60% of ciliated cells and 0.2% of Club cells are derived from BDPCs (Figure 31C), which is consistent with the histological analysis of LacZ activity in ciliated cells of the airway epithelium.

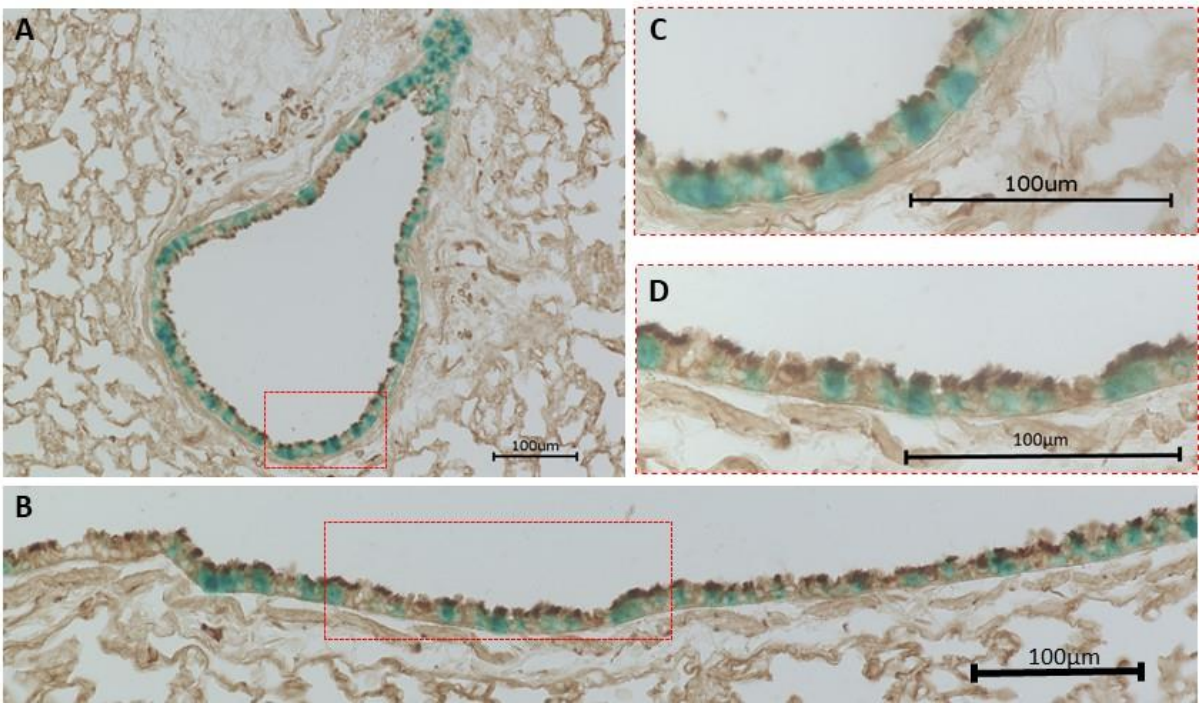


Figure 30: Fate mapping of BDPCs during lung homeostasis.

(**A,B**) Adult mouse lung sections (thickness 10µm) of BDPC v-race mice were stained for both β-galactosidase (blue) and α-tubulin (brown). α-tubulin was stained using standard immunohistochemical techniques and DAB to detect antigen-antibody complexes. For each panel, the red boxed areas are zoomed in **C** and **D** to magnify and highlight co-stained regions. Scale bar: 100µm.

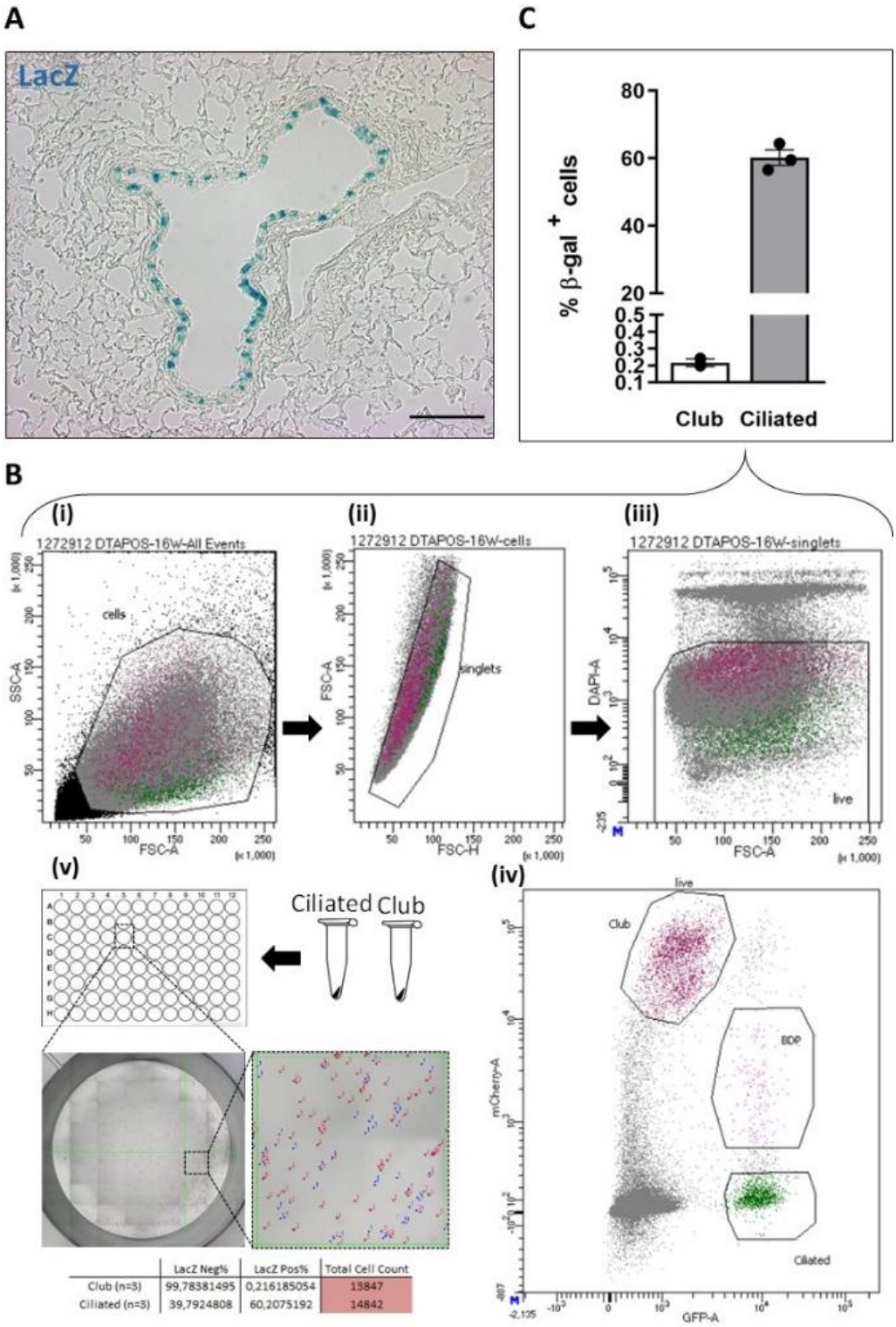


Figure 31: FACS based quantification of BDPCs derivatives.

(**A**) β-galactosidase staining of lineage-traced cells in v-race adult mouse lungs and frequency of blue cells present in the epithelium of lung airway. Scale bar: 100 µm. (**B**) FACS isolation of Club and ciliated cells from the BDPC v-race strain. After live/dead discrimination by DAPI **(iii)**, cells were separated based on fluorescent reporter gene activity in each specific population, distinguishing mCherry in Club cells and YFP in ciliated cells **(iv)**. FACS-purified cells were pelleted, seeded in 96 well plate followed by overnight LacZ staining. A total number of 13847 Club and 14842 ciliated cells were counted and the ratio of LacZ positive cells was calculated relative to non-labeled cells (n=3) **(v)**. (**C**) The graph shows the percentage of traced Club (0.2%) and ciliated cells (60%) from BDPC v-race strain under homeostatic conditions.

To better understand the distribution of lableled ciliated cells through the entire airway tree, along the proximal-distal axis in the adult lung of the BDPC v-race strain, a combined tubulin and LacZ staining was performed (Figure 32). To facilitate comparison, four different zones along the proximal-distal axis were analyzed. Enlarged images of each zone are presented in Figure 32 (1-4). Zone number represents the most proximal area and is closest to the first bifurcation of the main stem bronchi. Area number two and three represent mid zone bronchioles, and number four represent terminal bronchioles. Ciliated cell surface density decreases in a proximal to distal direction [85], which is reflected by the density and number of LacZ-labeled cells.

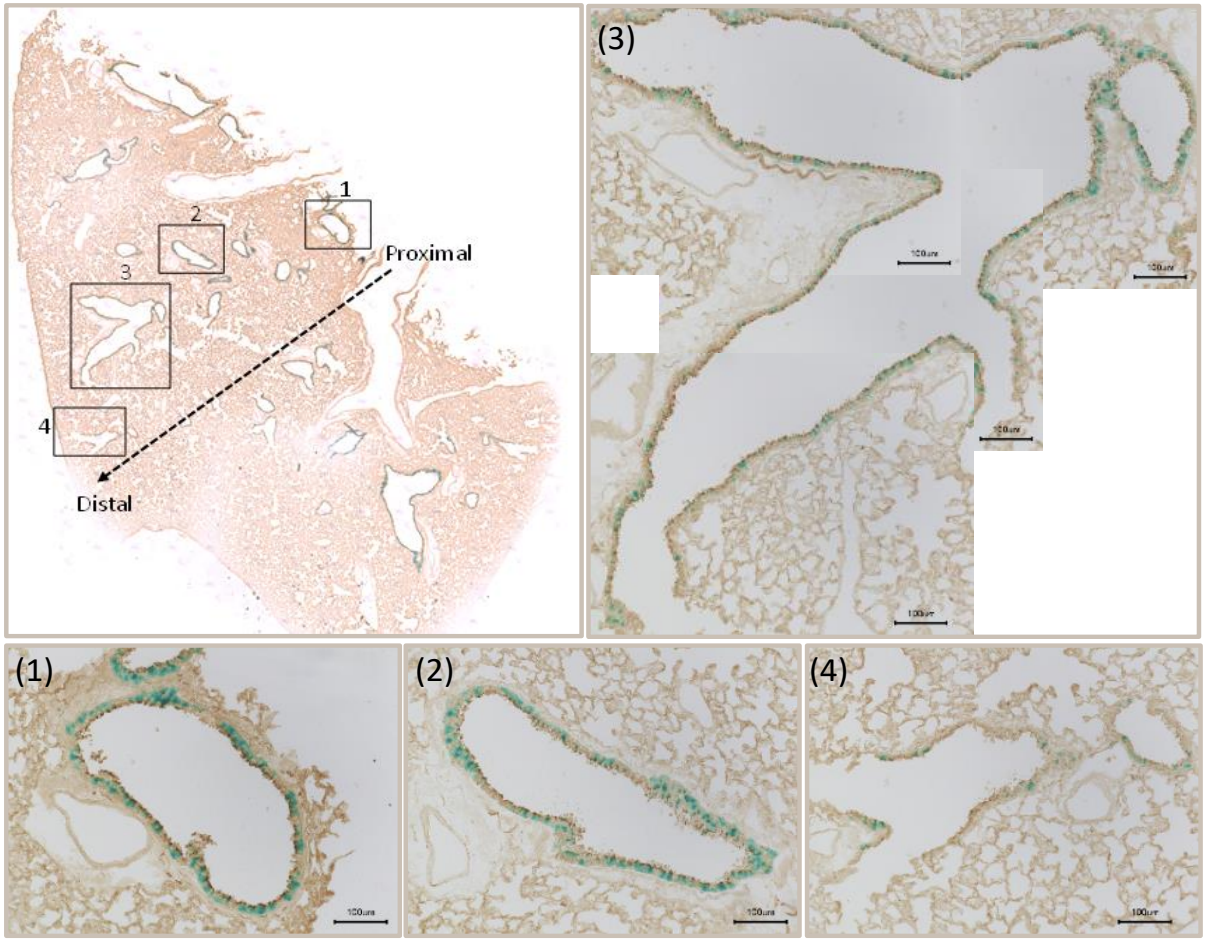


Figure 32: The density and distribution of the BDPC desendants changes along the proximal to distal axis.

Combined β-galactosidase (blue) and α-tubulin (brown) staining on sections of the right lobe from adult BDPC v-race mice under homeostatic conditions. Frequencies of LacZ+ cells at specific positions relative to the proximal distal axis (dashed black arrow) are compared in four different zones (**1**-**4**). **1**-**4** Zoomed images of magnified boxed regions. Zone (**1**) is the most proximal and (**4**) is the most diastal zone. Scale bar: 100µm.

6.9 Contribution of BDPCs to regeneration after naphthalene and influenza injuries.

To understand the roles of BDPCs in lung repair and regeneration after injuries, the BDPC-specific reporter strains were exposed to naphthalene, which specifically depletes Club cells in bronchioles. Lung samples of BDPC viewer mice were harvested 3 and 7 days after naphthalene treatment and LacZ positive cells were quantified at each time point. The number of labeled cells ($CCSP^+$ and $Foxj1^+$ co-expressing cells) at 3dpn and the number of labeled $LacZ^+$ cells at 7dpn increased 10-fold (Figure 33).

The same experiment was done with adult BDPC v-race mice but lungs were harvested 21 days after naphthalene treatment to examine the role of BDPCs for regeneration (Figure 34). At 21 dpn clusters of newly formed Club cells have largely expanded, although the bronchiolar epithelial lining continues to regenerate [86]. By performing lacZ staining on lung sections of injured BDPC v-race and overlaying the LacZ positive patches with fluorescence images, small clusters of BDP-derived Club cells were observed at bifurcations (Figure 34C). In addition, a 10x increase of traced cells in the Club cell fraction was apparent based on FACS-based quantification, confirming results obtained by microscopic analysis (Figure 34B).

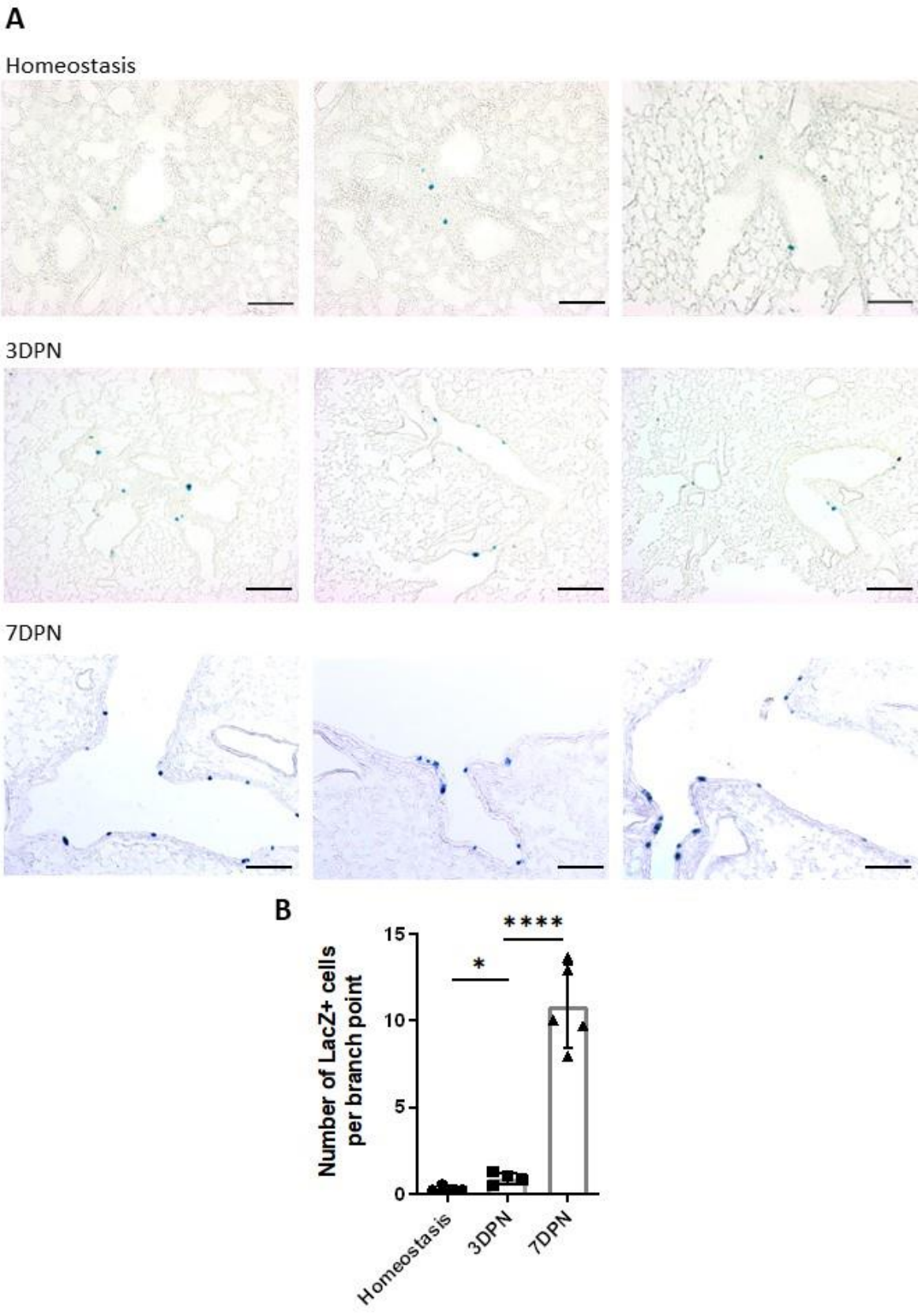


Figure 33: BDPCs are naphthalene resistant and expand during epithelial repair.

(**A**) β-galactosidase staining of lung tissue derived from control and naphthalene-treated BDPC viewer mice at 3 and 7 dpn. Scale bar: 100µm (**B**) Quantification of $LacZ^{+}$ cells 3dpn and 7dpn in comparison to homeostasis. The data are presented as mean ± SD, n = 5, *P<0.05 and ****P<0.0001 compared with homeostasis.

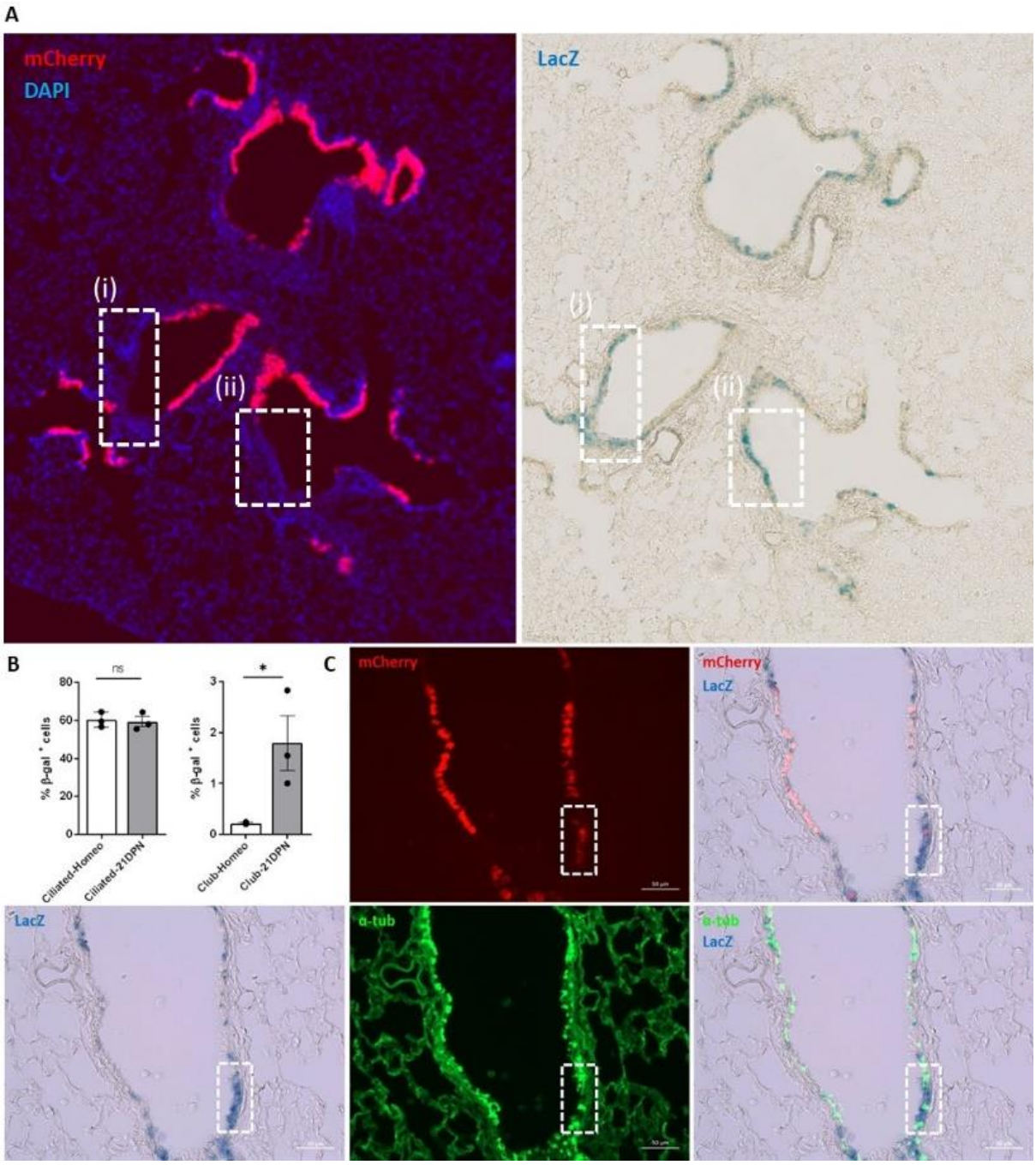


Figure 34: BDPCs give rise to newly formed Club cells after naphthalene-mediated injury.

(**A**) Sequential imaging of newly formed mCherry^{+} Club cells and LacZ^{+} signal on lung sections, 21 days after naphthalene treatment. Boxed regions ((i), (ii)) show the same region of the lung epithelium after either fluorescence imaging or LacZ-staining. (**B**) Quantification of BDPC derived ciliated and Club cells 21 days post naphthalene. Data represented as mean ± SD, *P<0.05. (**C**) Sequential imaging of lung section (fluorescence and bright field) 21 days post naphthalene treatment. Boxed region shows cluster of mCherry^{+} Club cells derived from BDPCs.

Naphthalene is toxic for Club cells but affects ciliated cells only mildly (see below Figure 35). Similarly, BDPCs were virtually unaffected by naphthalene (Figure 33) [87], and still gave rise to ciliated cells (Figure 30- 31). To explore whether BDPCs are damaged by the influenza virus, which damages many different cell types in the respiratory epithelium, (Figure 36) [88, 89],

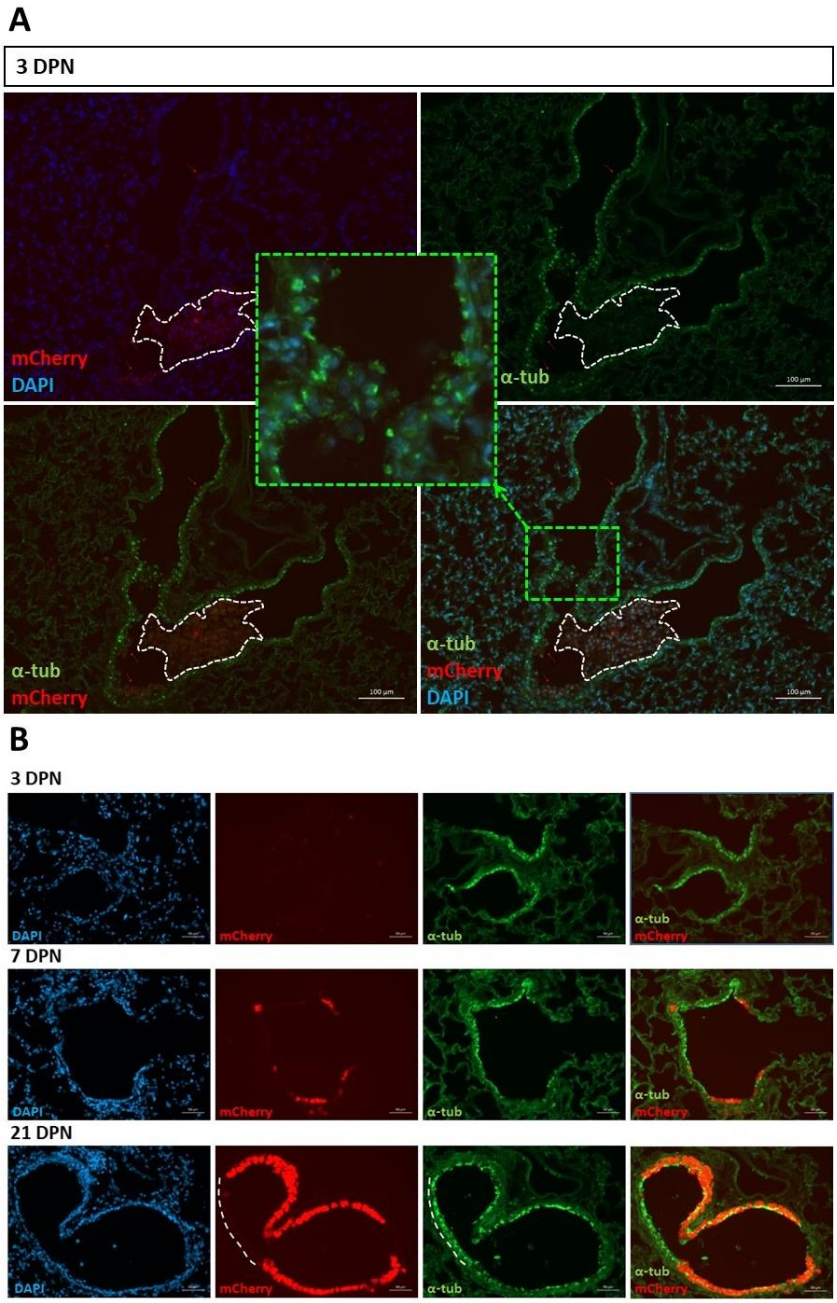


Figure 35: Impact of naphthalene induced injury on Club cells and ciliated cells.

(**A**) mCherry^{+} Club cells and immunostaining for α-tubulin (ciliated cells) on adult mouse lung tissue 3 days post naphthalene. Nearly all Club cells are exfoliated and shed into the airway lumen (white dashed closed shape). Ciliated cells (green) are almost not affected except small patches that also become exfoliated and shed into airway lumen (green dashed box normal and zoom-up view). Scale bar: 100µm. (**B**) Dynamics of bronchiolar repair 3, 7 and 21 days post naphthalene-induced injury. mCherry expression (Club cells) and immune staining for α-tubulin (ciliated cells) at 3,7 and 21 dpn. Scale bar: 50µm.

BDPC v-race mice were infected with influenza virus and lung tissue was analyzed 9 weeks after infection.

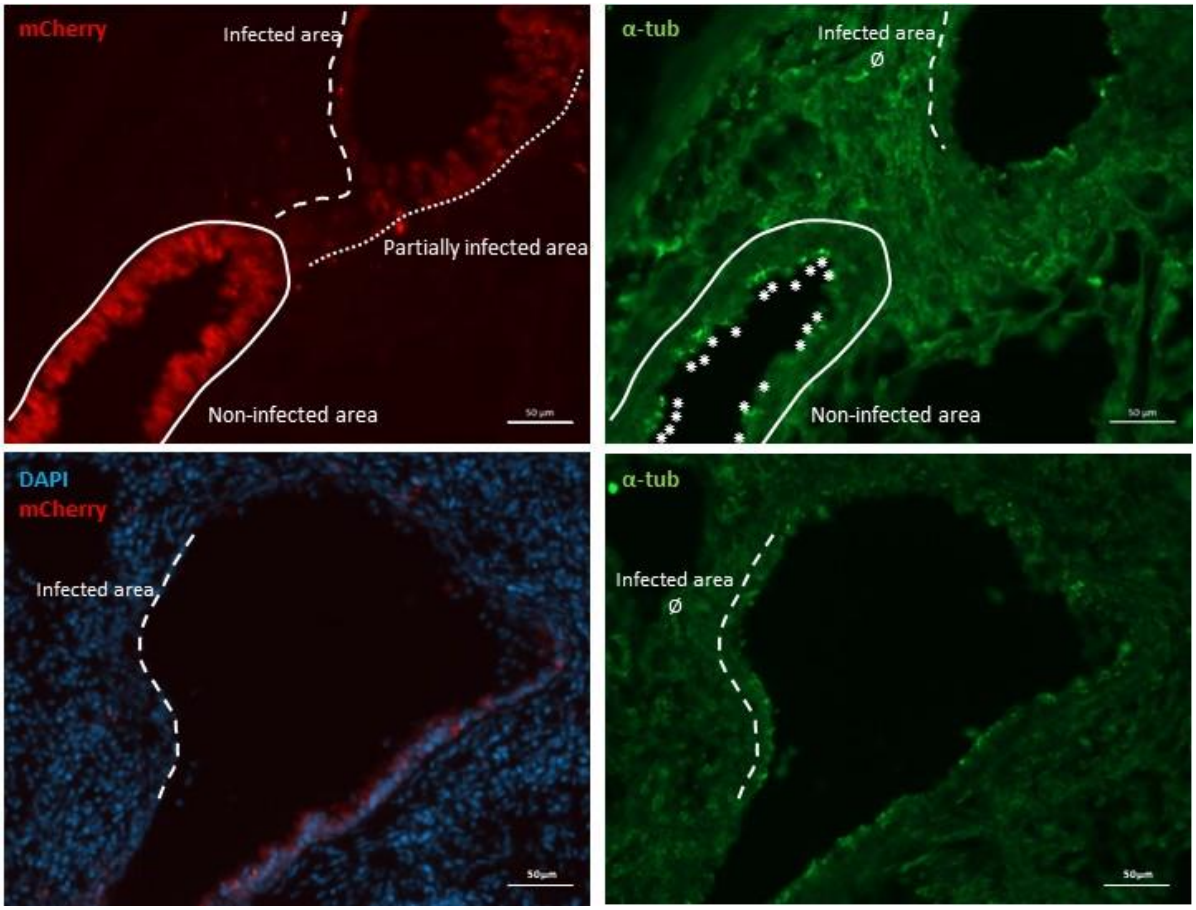


Figure 36: Impact of H1N1 induced injury on Club cells and ciliated cells.

Endogenous fluorescence (CCSP-mCherry^{+}, Club cells) and immunostaining for α-tubulin (ciliated cells) on adult mouse lung tissue, one week post influenza injury. Areas marked by closed lines represent a noninfected area with the presence of both Club cells (mCherry expression) and ciliated cells (asterisks mark α-tubulin green). Dashed lines indicate heavily infected areas and strong tissue damage. This area is almost devoid of any mCherry expression (lack of Club cells) and ciliated cells. Dotted lines indicate partially infected areas represented by faint mCherry expression and presence of some ciliated cells. Scale bar: 50mμ.

Influenza virus infection does not cause a uniform injury in the lung. Club cells and ciliated cells survive in some areas but are almost completely lost in others. BDPC-derived cells were rare in infected areas, while non-infected areas essentially showed the same pattern for BDPC-derived cells as during homeostasis (Figure 37), suggesting that BDPCs are susceptible to influenza infection.

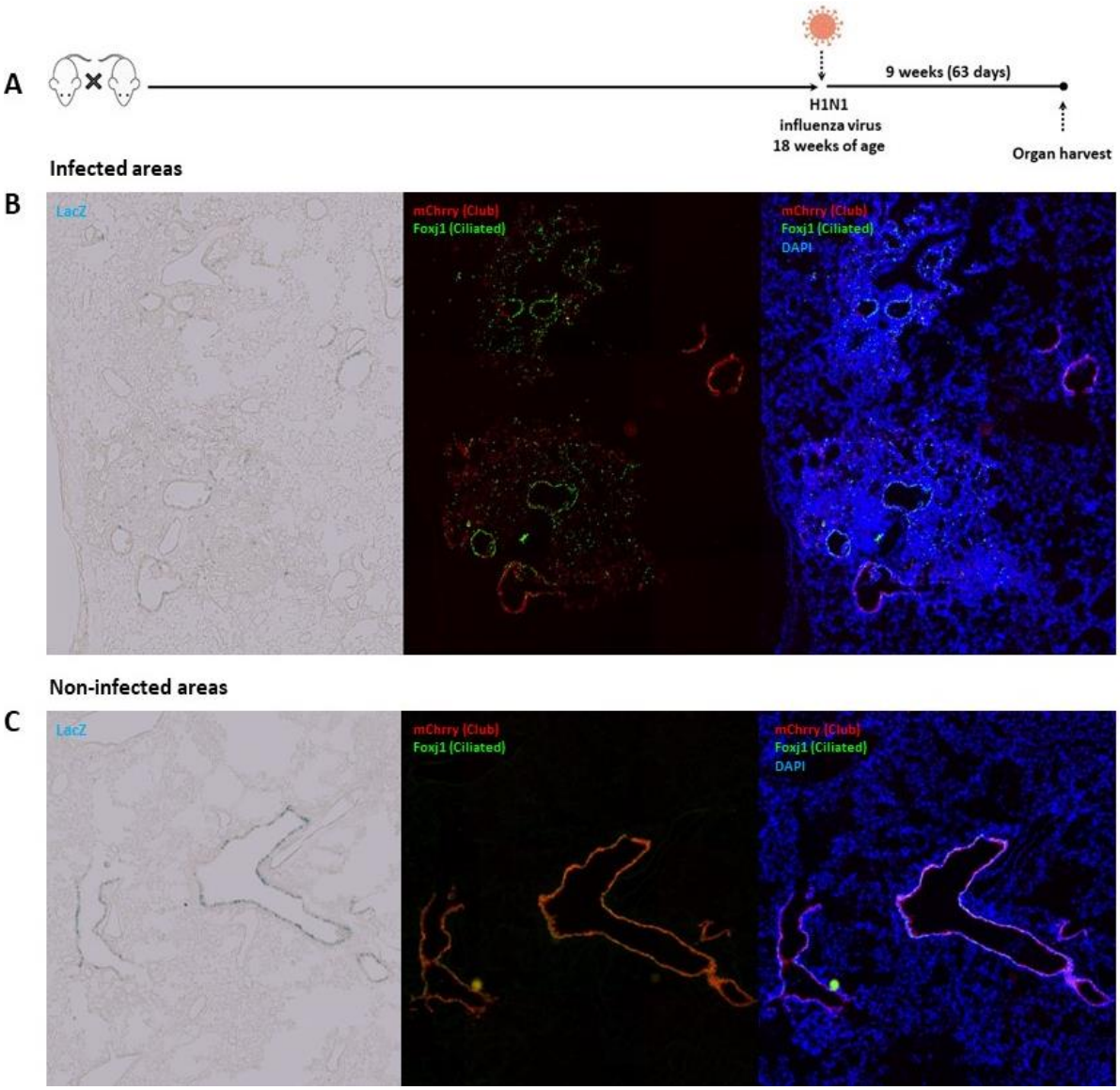


Figure 37: Contribution of BDPCs to airway epithelium repair 9 weeks post H1N1 infection-induced injury.

(**A**) Schematic outline of the experimental design. Adult BDPC v-race mice (18 weeks of age) were infected intratracheally with influenza A/H1N1 (2000 pfu). Samples were analyzed 9 weeks after influenza infection.

(**B,C**) β-galactosidase staining (bright field) and fluorescence microscopy of endogenous reporter expression (CCSP-mCherry) and immune staining for Foxj1 (green) of the same region within infected lungs. Infected areas are shown in **B**, non-infected areas in **C**.

6.10 Clonal analysis of BDPCs during homeostasis and after lung injury

It is not trivial to trace formation of BDPC-derived patches of ciliated cells when pre-existing ciliated cells that form during organ growth are also traced. Therefore, pilot experiments were performed to identify conditions, allowing exclusive labelling of BDPCs in adult animals after naphthalene treatment and influenza infection, avoiding widespread tracing of ciliated cells during organ growth (Figure 38).

The system used to label BDPCs depends on doxycycline exposure (Figure 25B), offering the opportunity to determine the time window of gene expression at will. In the presence of doxycycline, the system is off and in the absence of doxycycline it is on. When doxycycline is absent, the tTA molecule binds to its cognate promoter, leading to activation of gene expression. Administration of doxycycline prevents binding of tTA to the promoter and gene expression is switched off (Figure 38A). Labeling of cells will only occur in the time period when doxycycline is withdrawn (Figure 38 C,D). Since BDPC v-race animals were used for all pulse-chase experiments, in which tTA activates Cre-dependent recombination, labeling is permanent and will not be reverted upon re-initiation of doxycycline administration but only prevent *de novo* labeling of newly forming BDPCs. With other words, only derivatives of BDPCs that are present during the time of doxycycline withdrawal will be visualized.

First, the exact time window was defined, which results in permanent labeling of the majority of BDPCs but minimizes tracing of BDPC-derivatives emerging during organ growth. Definition of this time window is a critical step for all next pulse-chase and clonal analysis of BDPCs (Figure 38C).

Starting at weaning age, doxycycline exposure was suspended for different time periods to allow labeling of BDPCs, and lungs were harvested at 12 weeks of age (Figure 38). As expected, prolonged doxycycline withdrawal correlated with increased presence of lineage-traced cells as shown by whole mount β-galactosidase staining of accessory lobe preparations (Figure 38D). Quantification of labeled cells on histological sections revealed that already a short withdrawal of doxycycline resulted in efficient labeling of BDPCs (Figure 39). Since the difference between the 1- and 2-weeks labeling period was rather minor, the 2-weeks doxycycline withdrawal regimen was selected.

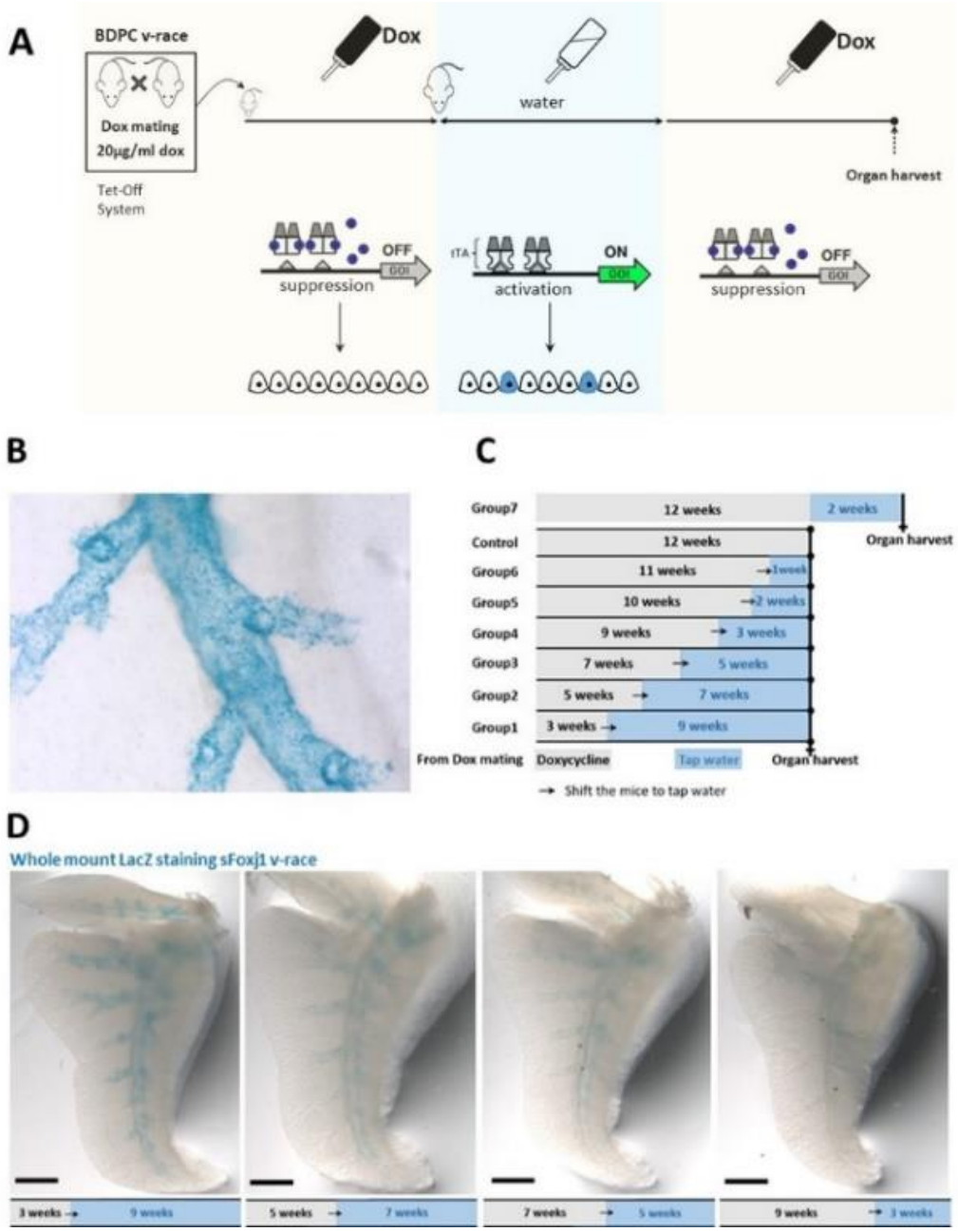


Figure 38: Schematic view of the inducible manipulation of gene expression in the Tet-Off system and pilot experiments to determine the time window for doxycycline administration.

(**A**) Tet-off system allows silencing of gene expression by administration of tetracycline or tetracycline-derivatives like doxycycline (dox). Gene expression can be switched off by initiating mating under doxycycline treatment and keeping the mother and the litter on doxycycline up to adulthood. Withdrawal of doxycycline for 2 weeks will activate expression of the reporter gene. Reapplication of doxycycline prevents *de novo* labeling and allows clonal analysis of cells that were marked during the pulse period when doxycycline was absent. (**B**) Whole mount β-galactosidase staining of cleared accessory lobe of BDPC v-race, reflecting constant active labeling. (**C**) Schematic, showing experimental groups to manipulate and limit number of BDPC-derived cells close to the number of actual co-expressing cells in BDPC viewer mice. Group1 represents the longest pulse time window while Group6 represents the shortest. All mice were sacrificed at 12 weeks of age. (**D**) Whole mount β-galactosidase staining of cleared accessory lobes of Group1 to Group4. Reduction in the number of labeled ($LacZ^+$) cells is reflected by the lower density of blue precipitates in group1 vs Group4. Scale bar: 2mm.

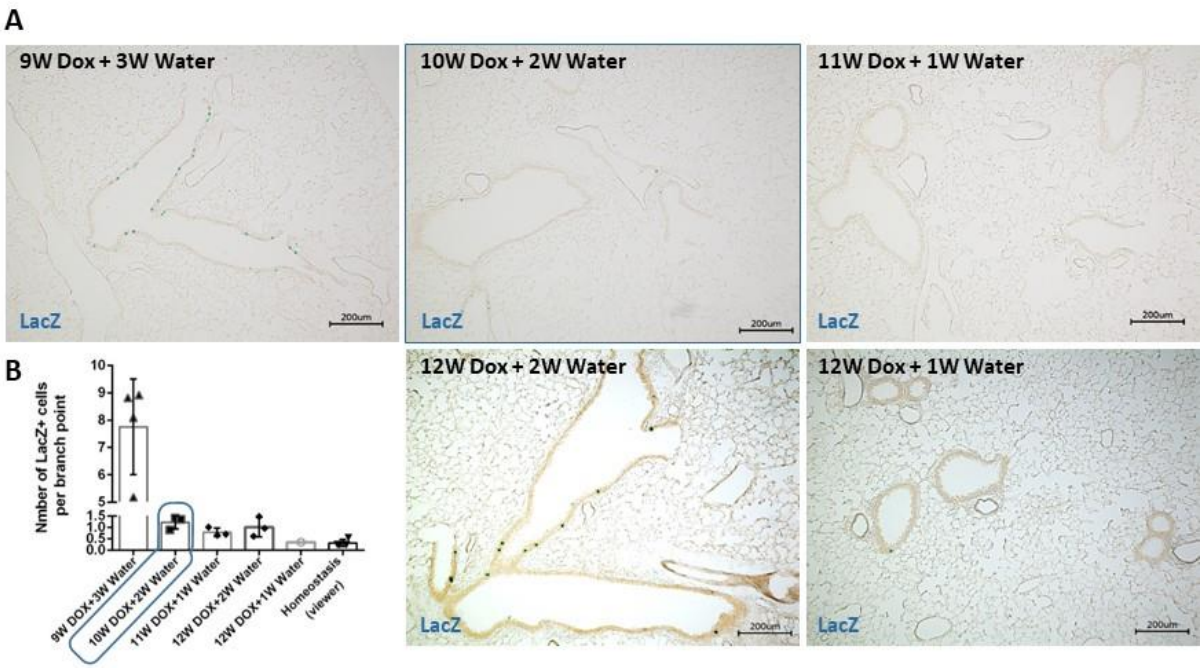


Figure 39: Withdrawal of doxycycline for 1 week activates β-galactosidase labeling in BDPC v-race.

(**A**) β-galactosidase staining of lung sections from BDPC v-race animals after indicated labeling periods (see also Figure 38**:**). Scale bar: 200µm. (**B**) Quantification of the number of $LacZ^+$ cells per branch point in Group4 to Group7. Group6 has the shortest pulse period of 1 week, resulting in number of $LacZ^+$ cells that are similar to BDPC viewer mice. To maximize the number of labelled BDPCs, a 2-weeks labeling period was chosen (blue box in the graph).

After determining the time for optimal labeling of BDPCs, expansion and differentiation of individual BDPCs during homeostasis and lung regeneration was analyzed. Two weeks labeling followed by a 6 months chase period revealed a relatively low contribution of $CCSP^+$ $Foxj1^+$ co-expressing cells to homeostatic maintenance of the lung epithelium (Figure 41). The number of BDPC-derivatives was only marginally higher than the number of actual BDPCs, which is most likely the reflection of a very slow turnover of ciliated cells under homeostatic conditions (Figure 41B).

Since cellular turnover is very low in the undamaged healthy lung, which prevents efficient analysis of the capacity of BDPCs to generated progenies such as cilitated and Club cells, BDPC v-race mice were infected with the H1N1 virus. BDPC v-race mice were kept on doxycycline the whole time except for a short pulse of doxycycline withdrawal for 2 weeks, immediately before introducing the virus, which suppresses any *de novo* labeling during infection and regeneration (Figure 41). Surprisingly, the damage inflicted by influenza virus infection did not result in a major increase of BDPC-derivatives. At 63 dpi, histological analysis of the lacZ-positive cells of lung from infected mice revealed only minor differences compared to the

homeostatic condition, suggesting that BDPCs are either destroyed by the influenza virus (which is reminiscent of the situation with BASCs) and therefore cannot contribute to regeneration, that damage of ciliated cells due to the influenza infection was minimal or that BDPCs only contribute to formation of new cilitated and Club cells under different disease conditions.

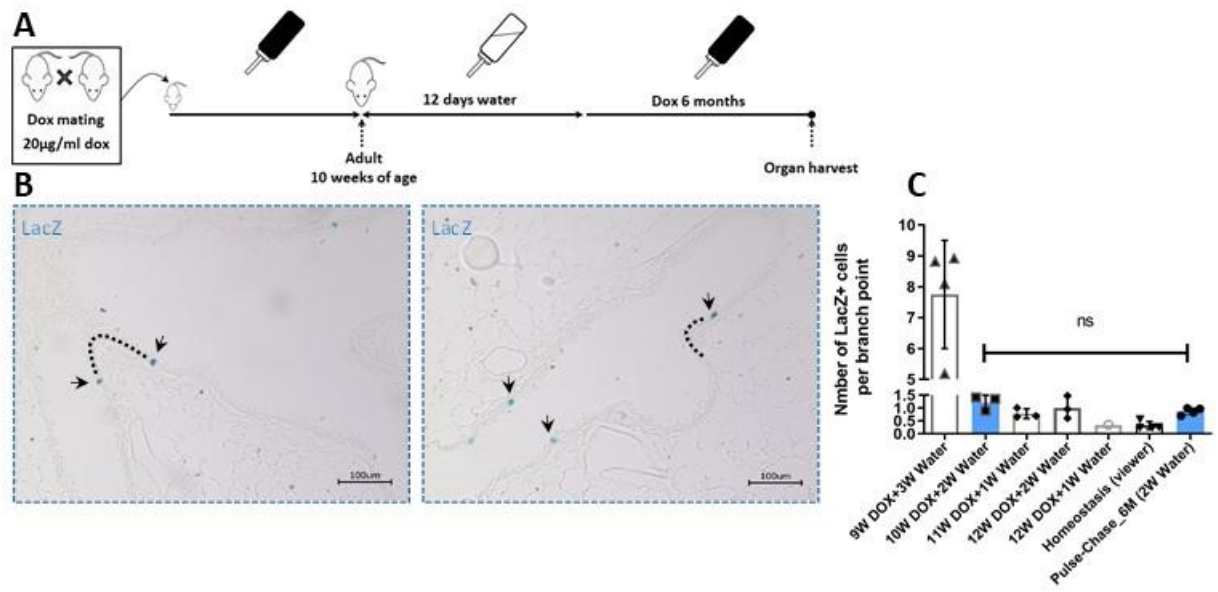


Figure 41: Genetic clonal analysis of single BDPCs in homeostasis after 6 months.

(**A**) Schematic depiction of the experiment design. Animals were maintained on doxycycline until adulthood. After 2-weeks of labeling by doxycycline withdrawal (tap water), the animals were kept on doxycycline and the lungs were harvested for assessment after 6 months. (**B**) β-galactosidase staining of pulsed-chased samples 6 months later. Arrows indicate LacZ signals and the black dashed line represents airway bifurcations. Scale bar: 100µm. (**C**) Quantification of the number of $LacZ^{+}$ cells following the 6 months pulse-chase period and comparison to Group5 from the doxycycline pilot experiment (n=4).

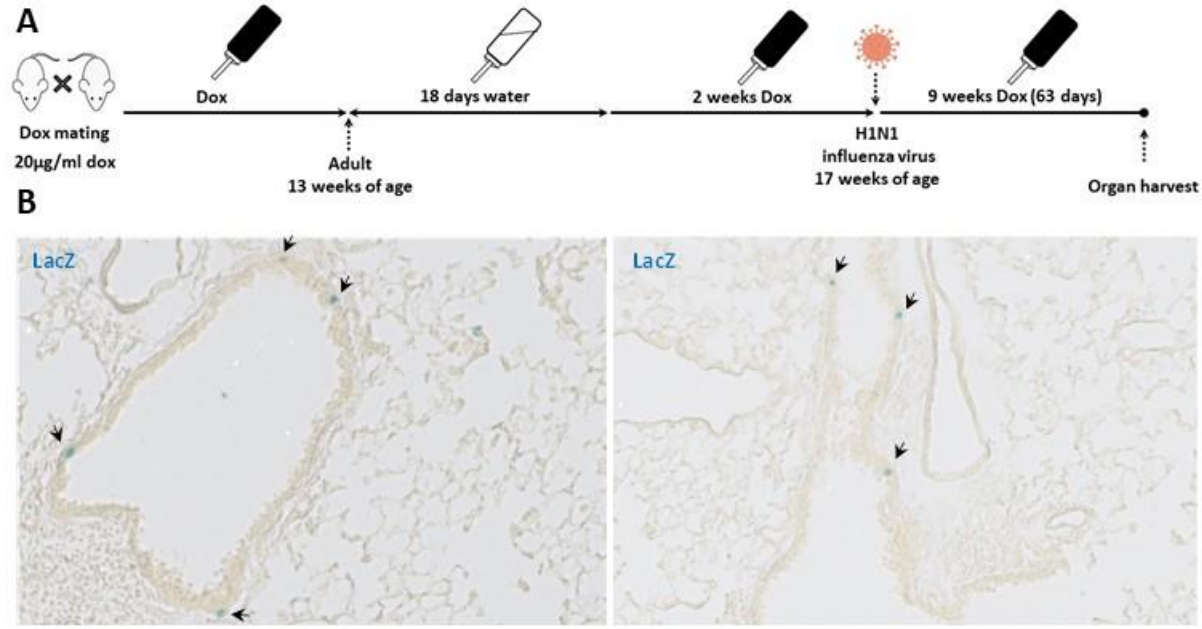


Figure 41: Pulse-chase lineage tracing of BDPCs 9 weeks following viral infection.

(**A**) Schematic depiction of the lineage tracing experiment based on two weeks of BDPC-labeling and H1N1 infection (BDPC v-race strain). (**B**) At 63 dpi, only few cells expressing β-galactosidase are present in bronchioles, mainly in noninfected areas (black arrows).

6.11 Selective ablation of BDPCs results in reduced cellularity of the bronchiolar epithelium

To better understand the physiological role of BDPCs for lung homeostasis and regeneration, BDPCs were genetically ablated due to Cre-mediated expression of the diphtheria toxin fragment A (DTA) (Figure 42). To establish this mouse line, the BDPC v-race were crossed with the *Rosa26*$^{stopfloxed\ DTA+/-}$ mouse stain [90]. Since Cre-mediated recombination between loxP sites only happens in CCSP^{+} Foxj1^{+} co-expressing cells, BDPCs will get selectively depleted.

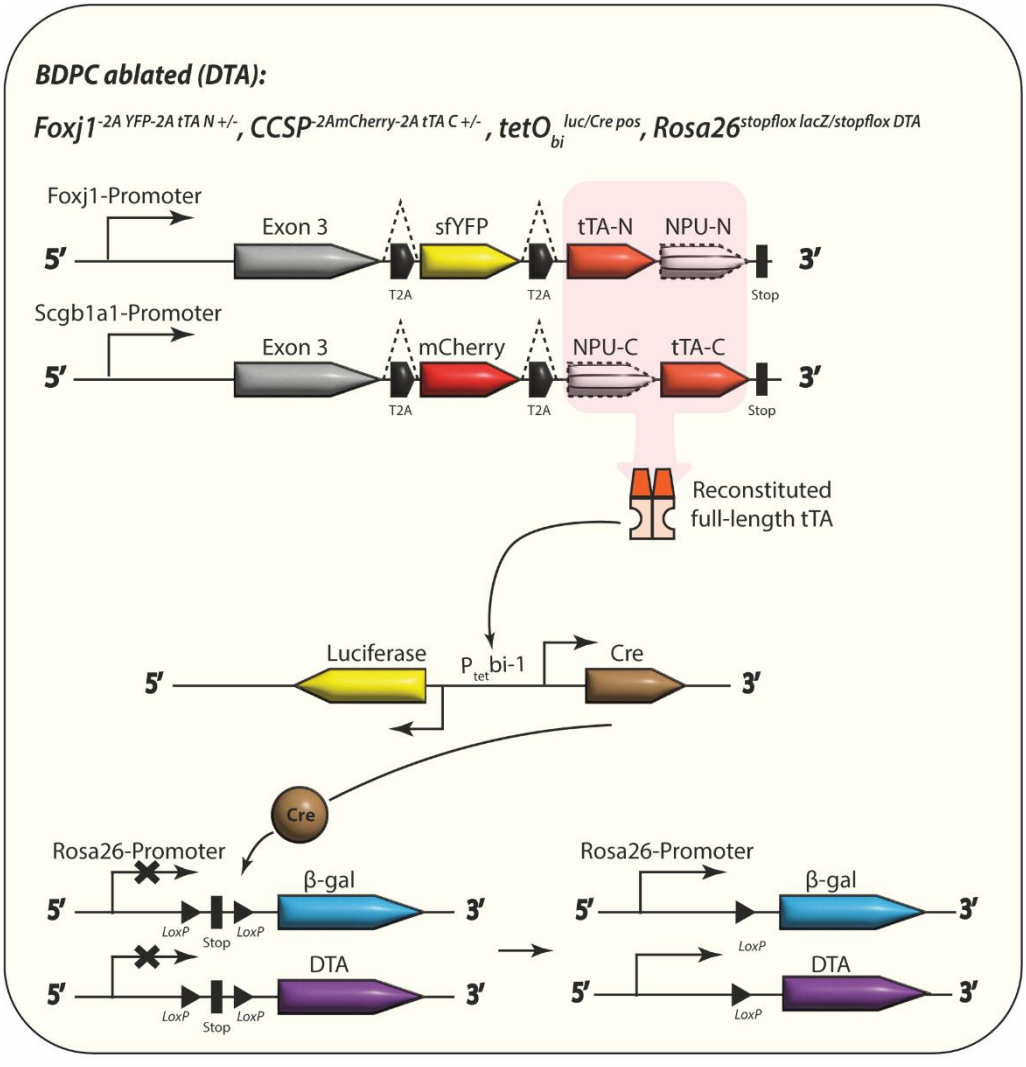


Figure 42: Schematic overview of generation of BDPC-ablated mice (DTA).

Lung samples were collected for histological examinations at the age of four weeks. H&E staining of lung sections revealed a reduced diameter and reduced cellularity of the epithelium in DTA-positive compared to control animals (Figure 43A).

LacZ staining of DTA and control animals confirmed that depletion of labelled BDPCs worked efficiently. Only very few residual $LacZ^{+}$ cells were observed, which have escaped DTA-mediated cell killing due to unknown reasons (Figure 43B). Quantification of Club, ciliated and BDPCs by FACS indicated a substantial decline of the number of ciliated cells, which was expected since previous findings demonstrated that BDPCs act as progenitors for ciliated cells. In addition, an approximately 60% reduction of the Club cell population was observed (Figure 43C), which was surprising. The decline of Club cells suggest that ablation of BDPCs stresses and depletes Club cells, probably due to differentiation to ciliated cells, a job normally done

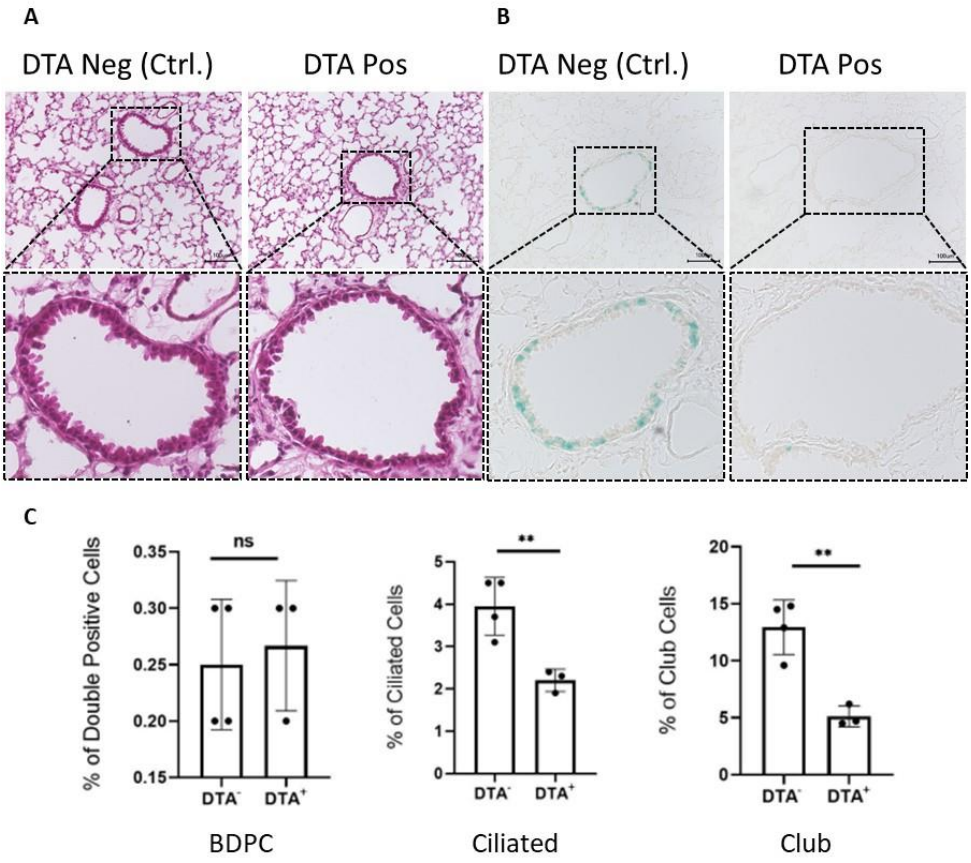


Figure 43: DTA-mediated selective ablation of BDPCs leads to reduced cellularity of the bronchiolar epithelium.

(**A**) H&E staining of BDPC-ablated mice (4 weeks old) vs Ctrl. Zoom-up view of the boxed regions highlights reduced cellularity and diameter of the epithelium in DTA^{Pos} vs DTA^{Neg} (Ctrl.). (**B**) Efficiency of ablation was assessed by performing β-galactosidase staining on ablated vs non-ablated sample. Scale bar: 100µm. (**C**) FACS-based quantification of BDPCs, ciliated and Club cells in ablated vs ctrl. (13 weeks old). **P-value of 0.009 and 0.003 ciliated and Club respectively.

by BDPCs. The absence of BDPCs seems to force Club cells into a compensating pathway, which still allows formation of ciliated cells but at the expense of regular Club cell numbers.

Next, the DTA^{pos} strain was exposed to influenza infection to study the effect of BDPCs ablation on lung regeneration. Ablation of BDPCs and its direct and indirect effect on ciliated and Club cells, respectively, caused a significant delay in bronchiolar regeneration. Moreover, much larger Krt5^{+} pods appeared 2 weeks after influenza infection in the BDPC-ablated strain compared to control animals , which indicates activation of basal cells, which only contribute to distal lung regeneration after very severe damage or when regular progenitor cells are not available (Figure 44).

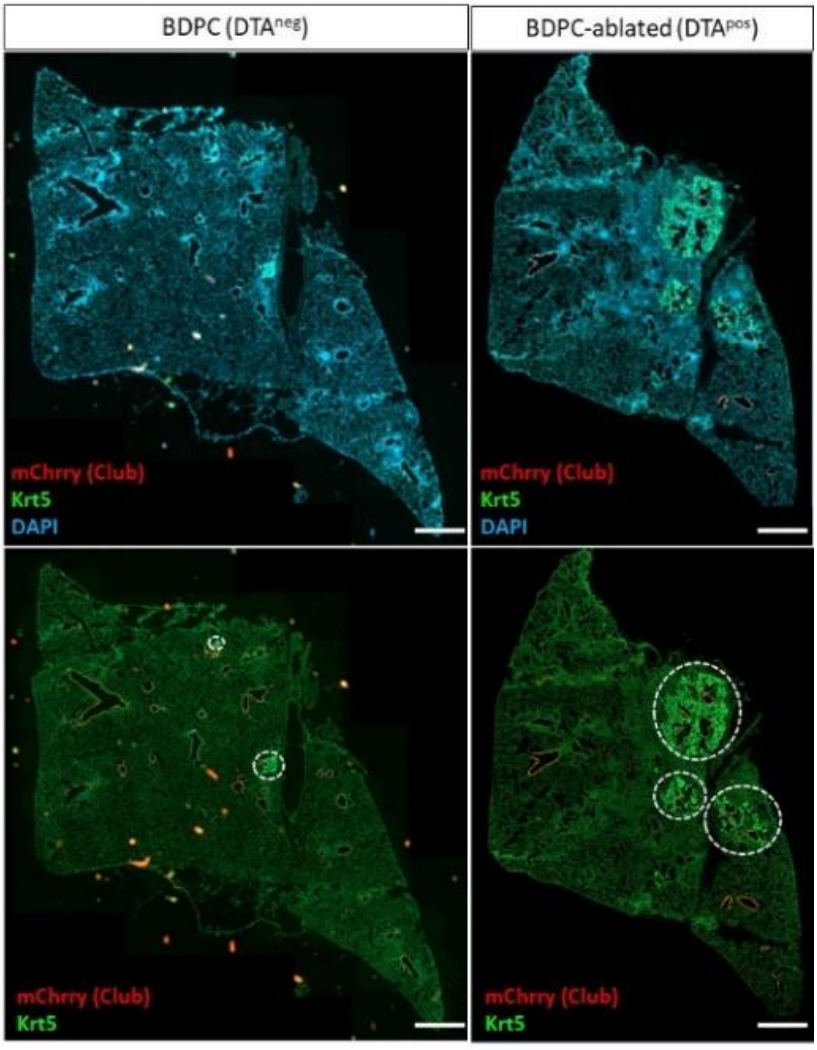


Figure 44: Presence of large Krt5^{+} patches two weeks after influenza infection in BDPC-ablated (DTA^{pos}) mouse.

BDPC-ablated (DTA^{pos}) (left panel) and BDPC v-race (DTA^{neg}) (panel right) mouse lung 2 weeks after viral infection. Immunofluorescence images of infected lungs of anti-Krt5 (green), endogenous CCSP^{mCherry} (red) and DAPI (blue, upper panel). Dashed circles marked Krt5^{+} patches in BDPC-ablated (DTA^{pos}) and BDPC (DTA^{neg}). *n*=3 mice (12 weeks of age). Scale bars, 1mm.

7 DISCUSSION

The study identified a novel epithelial cell population in adult mouse lung that concomitantly expresses Club and ciliated cell marker. The cell population can give rise to either Club or ciliated cells and is positioned at airway branch points, which is why the cells were named Branch Point associated Double Positive Cells (BDPCs) (sub-chapter 6.2). Knock-in mouse strains were generated to target and lineage trace BDPCs *in vivo*, using intein-mediated assembly of split-tTA [55] (sub-chapter 6.7). Reconstitution of the two inactive halves of the tTA molecule in CCSP$^+$ Foxj1$^+$ co-expressing cells (BDPC viewer strain) allowed localization of BDPCs at airway branching points. To lineage trace progeny of BDPCs, a knock-in strain was established (BDPC v-race), in which reconstituted split-tTA induces expression of Cre recombinase, marking both CCSP$^+$ Foxj1$^+$ co-expressing cells and their descendants. Analysis of BDPC v-race mice revealed that BDPCs generate almost 60% of ciliated cells and 0.2% of Club cells in homeostatic conditions (sub-chapter 6.8). Finally, the contribution of BDPCs for regeneration of the pulmonary epithelium after bronchiolar and influenza injury was assessed, with and without DTA-mediated ablation of BDPCs (6.9-6.11).

7.1 BDPCs are airway branch point associated and give rise to Club and the majority of ciliated cells in adult mouse lungs

Single-cell RNA sequencing (scRNA-seq) data revealed that a subset of CCSP$^+$ cells also expresses the ciliated cell marker Foxj1. This novel gene signature of epithelial cell states have not been described previously. To specifically track CCSP$^+$ Foxj1$^+$ BDPCs *in vivo*, a genetic lineage-tracing system was established, taking advantage of split-intein mediated split-tTA reconstitution, specific labeling CCSP$^+$ Foxj1$^+$ BDPCs. Visualizing of co-expressing cells revealed that they are associated with airway branch points (Figure 29). It is well known that distinct anatomical niches, including airway branch points in adult mouse lungs, harbor stem/progenitor cells [9]. In 1995, *in situ* localization of CCSP mRNA after naphthalene treatment, revealed two distinct niches as starting points for epithelial recovery, one at airway bifurcations and the other at distal terminal bronchioles (BADJ) (Figure 45) [79]. The CCSP mRNA was detected without prior proliferation of a progenitor cells, suggesting that cells

residing at these locations before naphthalene treatment differ from other Club cells and are capable to survive naphthalene toxicity. Since both airway bifurcation- and BADJ-associated cells are resistant to naphthalene and show low Cyp2f2 expression, they were initially collectively variant Club cells (^{v}Club). The localization of ^{v}Club cells at airway bifurcations is in close proximity to neuroepithelial bodies, which might indicate a potential role of neuroepithelial cells in the regulation of ^{v}Club cells, although clear evidence for such a hypothesis is lacking. Previous lineage tracing studies using a *Cgrp-CreER* allele suggested that differentiated neuroendocrine cells start to proliferate after naphthalene-induced loss of Club cells and give rise to lineage labeled Club and ciliated cells [53]. However, these conclusions have to be viewed with caution, since Club cells may upregulate *Cgrp* during regeneration and thereby disguise their origin [3]. It seems more likely that NE cells do not directly contribute to epithelial renewal, but are critical for the maintenance of a reservoir of ^{v}Club cells, providing the necessary microenvironment or niche at branch points [41].

The identification of BASCs (green ovals in Figure 45) at BADJs [43] and development of genetic tools to target dual marker expressing cells [54, 55] , revealed that BASCs contribute to the regeneration of bronchiolar Club cells after naphthalene injury and of alveolar epithelial cells after bleomycin injury [14]. BASCs are not the only source for newly formed Club cells after naphthalene treatment. Salwig *et al.* reported that up to 43% of newly formed Club cells are derived from other sources and speculated that another source of newly forming Club cells are branch point-residing ^{v}Club cells [55] (also see the sub-chapter 7.2).

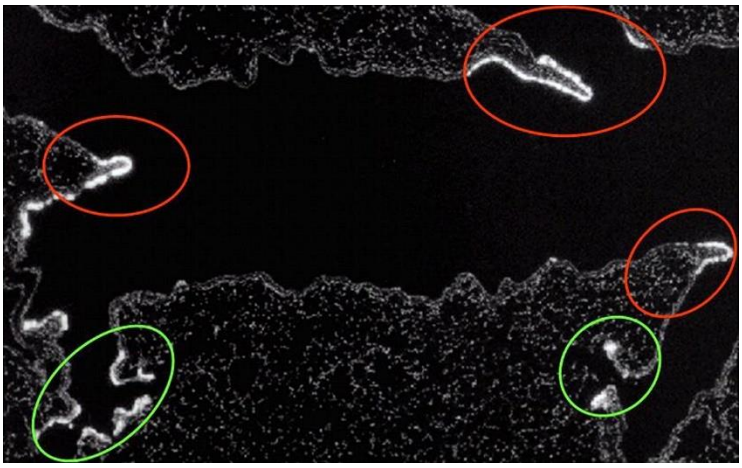


Figure 45: In situ localization of mRNAs for CC10 after naphthalene treatment.

In situ hybridization for CCSP mRNA (white autoradiographic grains) 72h post naphthalene treatment. Regenerative zones were identified at branch points in the airway (red ovals) and at bronchioalveolar duct junction (green ovals). Adopted from Stripp, B. R., K. Maxson, R. Mera, and G. Singh. Am J Physiol 269, no. 6 Pt 1 (Dec 1995): L791-9.

We found that BDPCs give rise to the majority of ciliated cells (60%) and only contributed to a minor degree to Club cells (0,2%) under homeostatic conditions. Apparently, renewal of $CCSP^{+}$ cells under homeostatic conditions is primarily achieved by regular Club cells, which are able to self-renew, but also generate ciliated cells during postnatal growth and adult homeostasis. However, it is currently unknown how differentiation of Club cells to ciliated cells is regulated and which cellular state(s) are involved. The distribution of labeled ciliated cells in BDPCs v-race differs along the proximal-distal axis, being more abundant and prominent in trachea and lobar bronchi, but more dispersed towards more distal terminal bronchioles (Figure 32). This finding is in agreement with reports in the literature that ciliogenesis differs spatially during airway generation and that ciliated cell surface density decreases in a proximal to distal direction [85]. In addition, BDPCs v-race pulse-chase experiments indicated a rather long life span of ciliated cells, which is in good agreement with previous investigations, reporting an estimated ciliated cell life span from 6 months in the trachea and over 17 months in terminal bronchioles [85, 91]. In this study the fate of BDPCs descendants was monitored for 6 months after a labeling pulse of two weeks, revealing nearly identical numbers of labeled cells 2 weeks and 5 months after the pulse. These findings clearly demonstrate that the turnover of ciliated cell is longer that 6 months and ciliogenesis mainly happen shortly after birth rather than in adult mice.

7.2 BDPCs and ^{v}Club cells share similar features

The scRNA-seq data set also showed that $CCSP^{+}$ $Foxj1^{+}$ cells express relatively low levels of Cyp2f2 in comparison to regular Club cells. Since $Cyp2f2^{Low}$ cells share a very similar anatomical localization with ^{v}Club cells, BDPC viewer mice were challenged with naphthalene to compare their survival rate with ^{v}Club cells. The airway epithelium of BDPCs viewer mice was assessed 3 days and 7 days post naphthalene, revealing that BDPCs are resistant to naphthalene. Furthermore, a 3- and 10-fold increase was observed at 3dpn and 7dpn, respectively. Newly formed BDPCs appeared mainly in association with branch points but also randomly in separate locations along the airway epithelium (Figure 33).

Lineage tracing experiments with BDPC v-race mice after naphthalene-induced injury also showed a strong 10-fold increase in the number of labeled Club cells 21 days post injury. No

significant change was observed in the number of ciliated cells, which is in agreement with the well-known fact that naphthalene is a Club cell specific cytotoxicant and ciliated cells are only mildly or no responsive at all to naphthalene [87]. In conclusion, BDPCs might represent ^{v}Club cells or a closely related population, based on naphthalene resistance, location at branch points, expression of CCSP, and the ability to give rise to Club cells after naphthalene treatment [79] (see also sub-chapter 7.1). However, some differences between BDPCs and ^{v}Club cells are apparent. The size and distribution of BDPC-derived, newly formed CCSP^{+} patches is smaller and more dispersed, respectively, compared to ^{v}Club cell-derived cells [39]. In addition, the majority of ciliated cells formed during postnatal lung development is derived from BDPCs, which has not been reported for ^{v}Club cells. However, it has to be pointed out that no lineage tracing study has been performed so far. Thus, some of the observed differences between BDPCs and ^{v}Club cells might be caused by different technical approaches and/or insufficient analysis of ^{v}Club cells.

7.3 BDPCs appear to be locked in a 'frozen' state with little contribution to regeneration following influenza infection

Damage caused by naphthalene treatment is relatively specific to Club cells and is rapidly repaired. In contrast, infection of mouse airways with a sublethal dose of influenza A (H1N1) virus results in profound and nonselective damage to the epithelium [89, 92-94]. Therefore, we expected that influenza infection of BDPCs v-race mice would provide more insights into the contribution of BDPCs for regeneration. Surprisingly, only small numbers of BDPC-derived, labelled cells were found in BDPCs v-race mice after influenza virus infection, both 14 and 63 days after infection. We expected a higher contribution of BDPCs for lung regeneration, since the influenza virus affects ciliated cells. Similar results were obtained from the clonal analysis of BDPCs post influenza, which also showed a minor contribution of BDPCs for regeneration in the most severely damaged areas.

Clonal analysis after influenza virus infection was done in a pulse-chase experiment, e.g. BDPCs were labeled by omitting doxycycline treatment, followed by doxycycline administration during and after influenza infection, thereby preventing any de novo labelling during the regeneration phase. The experimental setting may lead to an underestimation of

labeled cells, since cells that transiently express CCSP and Foxj1 during regeneration will not be labeled. Thus, it is reasonable to conduct additional experiments, in which labeling is also allowed after injury and during regeneration. This way, newly generated BDPCs will be labeled and the contribution of $CCSP^+$ $Foxj1^+$ derivatives can be scored. Despite these considerations, the differences in BDPC behavior after influenza virus and naphthalene-induced injury are difficult to understand. The number of BDPCs increases strongly 7 days after naphthalene treatment in BDPC viewer mice and BDPC-derived Club increases 10-fold, 21 days after naphthalene in BDPC v-race animals. In contrast, no strong increase of BDPCs and BDPC-derived Club cells happen after influenza virus infection. Three potential scenarios may be envisioned to explain these observations: (i) BDPCs are destroyed by virus and therefore cannot contribute to regeneration; (ii) BDPCs are locked in an arrested state, which cannot be overcome by damage inflicted by the influenza virus; (iii) another stem or progenitor cells population is activated by the influenza virus (but not by naphthalene injury), which outcompetes BDPCs, rendering their contribution minimal.

7.4 DTA-mediated ablation of BDPCs reveals the role of BDPCs during regeneration and identifies BDPCs as arrested in a transitory state

The lineage tracing approach revealed the contribution of BDPCs for regeneration, but did not answer the question whether BDPCs are instrumental for efficient lung regeneration. To address this problem, BDPCs were ablated by specific activation of the DTA gene, either during homeostasis and under injury condition, Surprisingly, DTA-mediated ablation of BDPCs resulted in the loss of 60% of Club cells and only 44% loss of ciliated cells under homeostatic conditions, although the contribution of BPPCs to Club cells, assessed by lineage tracing is minor. The depletion of Club and ciliated cells was reflected by changes in epithelial and airway morphology, which was less cellular and did not show the typical appearance of a pseudostratified epithelium.

It was concluded that ablation of BDPCs puts Club cells under strong stress, resulting in their exhaustion. In BDPC-DTA mice, Club cells are unable to enter the transitory state, characterized by co-expression of Foxj1 and CCSP, since such cells will get depleted immediately but need to differentiate directly into ciliated cells when replacement is

necessary. The constant ablation of any Club cells entering the transitory CCSP^{+} Foxj1^{+} state exhausts Club cells, leading to depletion. This hypothesis emphasizes the critical role of a transitory BDPC-state, which Club cells need to acquire before differentiating into ciliated cells. Club cells can bypass the transition state and directly convert to ciliated cells, but this apparently comes with a price and strongly stresses Club cells, which is avoided by the majority of Club cells (Figure 46). The relatively moderate decline of ciliated cells after BDPC-ablation supports this assumption, indicating direct replacement of ciliated cells by regular Club cells. Direct conversion of Club to ciliated cells without proliferation also explains why the loss of Club cells is so severe. Each individual Club cell can give rise only to a single ciliated cell, dramatically changing the stoichiometry of the lung epithelium, although the lining of the lung epithelium is to some extent maintained. A principal component analysis (PCA) of RNAseq data also demonstrated that Cyp2f2^{Low} expressing Club cells are similar to both Club (Cyp2f2^{High}) and ciliated cells. Moreover, the Cyp2f2^{Low} population showed high expression of Foxj1, which identifies BDPCs as an intermediate between Club (Cyp2f2^{High}) and ciliated cells.

The huge loss of Club cells in BDPC-ablated lungs, negatively influences regeneration after influenza injury, which was obscured during the first analysis, which was only based on scoring of LacZ labeled cells in infected areas. Analysis of BDPC-ablated lungs at 14dpi and 63dpi, uncovered a delay in regeneration and the presence of larger Krt5^{+} epithelial ‘‘pods’’, which are typical for severely damaged lungs following severe H1N1 influenza virus infections [95-97]. The appearance of Krt5^{+} patches after influenza infection of DTA^{+} mice, indicates that the loss of BDPCs enforces expansion of a different cell population, which is activated to compensate for regeneration defects in BDPC-ablated lungs [92, 96, 97].

The results support a model in which Club cells serve as the parental population for new Club cells, BDPCs, and ciliated cells. Currently, it is not entirely clear why acquisition of a transitory state is mandatory to generate ciliated and leave the Club cell compartment unharmed. Our model proposes that constant ablation of BDPCs, force the parental population (Club cells) to circumvent the roadblock by direct conversion into ciliated cells, which leads to exhaustion and depletion of parental population itself. It seems likely that the increase in BDPCs 7 days post naphthalene reflects higher transition from Club to ciliated cells and not enhanced proliferation of BDPCs. However, additional experiments are necessary to corroborate this

hypothesis. At present, it is difficult to understand why generation of ciliated cells sometimes happen trough a transition state and sometimes by direct conversion.

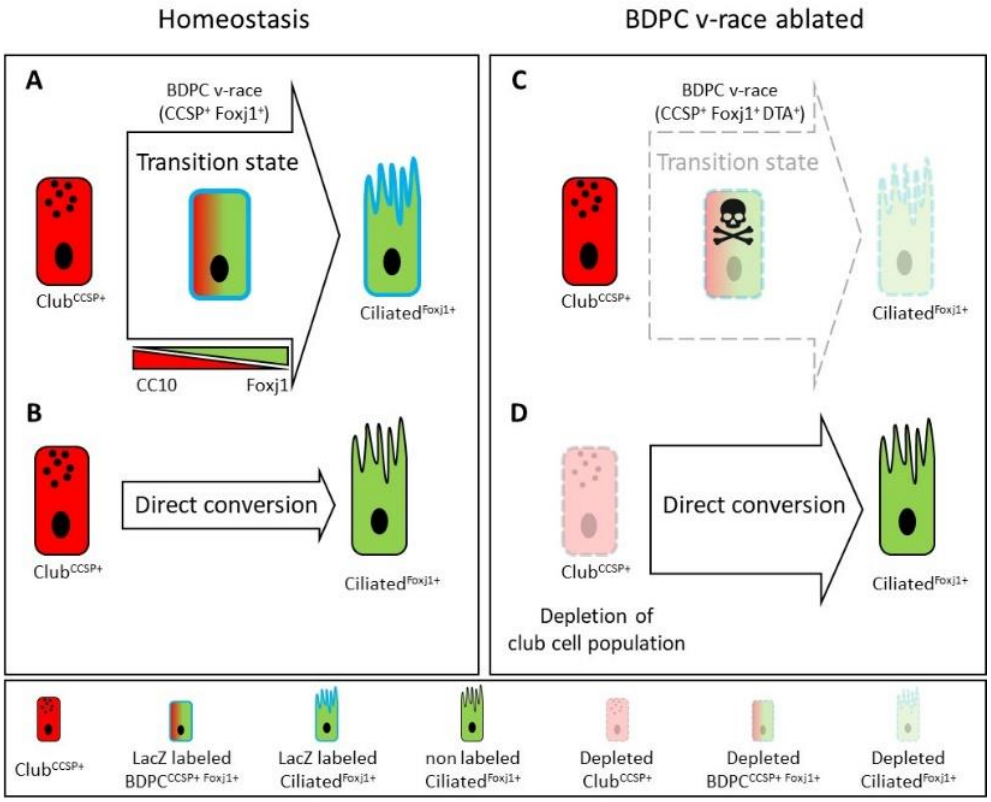


Figure 46: **Transitory state of BDPCs is critical for maintaining normal numbers of ciliated and regular Club cells.**

The schematic depiction shows labeling of BDPCs and their derivatives in BDPC v-race mice under homeostatic (left panel) and ablation condition (right panel). Homeostasis panel shows that BDPCs are labeled and majority of Club cells pass through the transitory state to give rise to ciliated cells (**A**) and minority take the direct path (**B**). In the ablated condition, passing through the transitory state is blocked since Club cells will get killed (**C**), therefore Club cells are force to pass a direct conversion to compensate absence of ciliated cells at the expense of depleting their own population (**D**).

Transition states are usually considered as short-lived (hence, transitory state) [60]. BDPCs appear to follow a different path and are “frozen” in a transition/intermediate state for a rather long time. The transition/intermediate state is a versatile condition when functioning properly allowing timely exit, which is critical for normal lung function. Accordingly, failure to exit the transition state will compromise repair processes. For instance, during the transdifferentiation process from AEC2 to AEC1, failure of AEC1 maturation can result in the accumulation of cells "stuck" in the transitional state. [98]. Inactivation of Notch signaling results in the accumulation of an intermediate cell population between AT1 and AT2 cells after *Pseudomonas aeruginosa* infection [99]. Similarly, inactivation of *Cdc42* results in the

accumulation of AT1-AT2 transitional cells after pneumonectomy. In both cases, accumulation of AT1-AT2 transitional cells leads to accelerated fibrosis. There may be overlap between cell states; therefore it is difficult distinguish between different cellular states. Cells in intermediate states corresponding to the *Ctgf* pre-alveolar type-1 transitional cell (*Ctgf* PATs), *Krt8*$^{+}$ alveolar differentiation intermediate (*Krt8*$^{+}$ ADI) and damage-associated transient progenitors (DATPs) in the mouse have also been found in fibrotic regions of human lungs from IPF patients. However, some genes are differently expressed between men and mice, suggesting some differences in the mouse and human response to injury and disease [63, 64].

It has been speculated that genomic damage and senescence are common traits of transition state stem cells, which cause stalling within the transition state that may happen under disease conditions [62]. However, it is not clear what is the cause and what is the consequence, since extended arrest in a transitory state under disease conditions may lead to senescence and not vice versa. Another critical question is the stability of the transition state in healthy tissues. How long does a transition state last? Do cells continuously move through such a state at a relatively rapid pace? The time period within a transitory state may also depend on the developmental stage (embryonic, postnatal or adulthood) and/or the physiological condition (homeostasis or disease).

7.5 Concluding remarks and future directions

Taken all results together, a novel discrete subpopulation of Club cells expressing both Club and ciliated cell defining marker genes (CCSP^{+} Foxj1^{+}) was discovered that reside at branch points and give rise to the majority of ciliated cells and to a minor fraction of Club cells. BDPCs share many similarities with ^{v}Club cells such as a location at branch points and resistance to xenobiotic compounds such as naphthalene. Ablation of BDPCs diminished the ciliated cell population but resulted in a more severe depletion of Club cells, which is most probably caused by a compensatory mechanism, in which regular Club cells directly convert into ciliated cells without proliferation, leading to their depletion. The findings underscore the high cellular plasticity of lung epithelium cells, which generates enormous regenerative capabilities. On the hand, they make clear that compensatory lung regeneration mechanisms have severe limitations, which lead to various pathologies when activated for extended time periods.

BDPCs were identified as the main cell population that generates ciliated cells. The bulk of Club cells also has the potential to generate ciliated cells, but direct conversion of Club to ciliated cells without going through the BDPC state is rare. Comparisons of the *in vitro* differentiation potential of BDPCs versus regular Club cells towards ciliated cells may provide further insights. It is also attractive to compare the ability of BDPCs versus regular Club cells for formation of ciliated cells by performing pulse-chase labeling at earlier time points during postnatal development, when massive formation of ciliated cells occurs. Advanced transcriptional profiling of Club cells, in which BDPCs are ablated, will be informative and uncover whether proposed stress state of Club cells under these conditions is reflected by corresponding transcriptional changes.

At present, it is not clear whether BDPCs are identical to ^{v}Club cells, represent a different entity or are partially overlapping. A previous study identified Upk3a1 as a potential marker for ^{v}Club cells (U-CCs). U-CCs survive naphthalene and give rise to CCSP^{+} patches at branch points, suggesting that Upk3a1-expressing cells may be related to ^{v}Club cells too [100]. To unequivocally define the identity and role of all these branch point associated cellular composition, generation of an Upk3a1/CCSP split mouse or tracing of Club cells with specific markers that are not present in ^{v}Club cells, would be truly helpful and also clarify whether BDPCs differ from vClub cells and/or U-CCs. We reason that BDPCs are not completely identical to ^{v}Club cells but rather represent an advanced state of ^{v}Club cells, which are about to differentiate to ciliated cells. To confirm this hypothesis, the increased contribution of BDPCs to Club cells after naphthalene-induced injury needs to be assessed in more details. Do BDPCs first convert to ^{v}Club cells and then to Club cells? If true, an increase in the number of Upk3a1^{+} cells should happen after naphthalene treatment but not when BDPCs are ablated. Again, creation of an Upk3a1/CCSP split mouse will provide insights, since ablation of ^{v}Club cells in such as mouse line will eradicate BDPCs as well.

Although several arguments support the claim that BDPCs are locked in a transition state, others support a different view. BDPCs are mostly located at airway branch points and are not intermingled amongst Club and ciliated cells, which seems as a requirement for transitory cells. However, BDPCs may actively migrate to bifurcations when going through transition, seeking support from neighboring cells, such as NEBs. Alternatively, the local niche at airway branches may support conversion of Club cells into BDPCs. Under pathological conditions,

such a conversion may also happen at other places, followed by secondary migration of ectopically born BDPCs to bifurcations.

The question remains whether BDPCs are “just” in a transition state between differentiated Club and ciliated cells or represent a stem/progenitor cell population in its own right. Better characterization of the transcriptional landscape of BDPCs and studies to understand the role of the branch point-associated niche, will provide new insights and aid our understanding of the role of BDPCs in lung regeneration and repair.

Transitory states have been mostly considered as short-lived. However, evidence is accumulating, suggesting prolonged existence of transitory/intermediates states that serve biological functions [98] but the underlying reasons are unclear. Are transition state cells able to deliberately alter their fate responding to local needs? For example, intestinal progenitors maintain an open chromatin configuration at enterocyte loci, allowing them to toggle between secretory and enterocyte progenitors [101, 102]. Acquisition of a transition state may also accelerate differentiation, as in the case of BDPCs that readily differentiate to ciliated cells, because stem/progenitor cells have already advanced towards differentiation and therefore need less time to complete differentiation. Acquisition of a transitory state might also involve adoption of a „safe mode“, enabling transition state cells to withstand adverse cues (e.g. BDPCs resistance to naphthalene).

Many more potential reasons may exist why acquisition of a stable transition state is beneficial, which need to distinguished from conditions, in which cells rapidly pass from one state to the other. We expect that other cells will be identified that are locked in transition either under homeostatic or diseased conditions. Further discoveries lay ahead that hold great promise to unravel the full implications of transitional states in homeostasis and pathogenesis.

8 REFERENCES

1. Ciechanowicz, A., *Stem Cells in Lungs.* Adv Exp Med Biol, 2019. **1201**: p. 261-274.
2. Hogan, B. and P.R. Tata, *Cellular organization and biology of the respiratory system.* Nat Cell Biol, 2019.
3. Hogan, B.L., et al., *Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function.* Cell Stem Cell, 2014. **15**(2): p. 123-38.
4. Volckaert, T. and S. De Langhe, *Lung epithelial stem cells and their niches: Fgf10 takes center stage.* Fibrogenesis Tissue Repair, 2014. **7**: p. 8.
5. Sivakumar, A. and D.B. Frank, *Paradigms that define lung epithelial progenitor cell fate in development and regeneration.* Curr Stem Cell Rep, 2019. **5**(4): p. 133-144.
6. Bartman, C.M., A. Matveyenko, and Y.S. Prakash, *It's about time: clocks in the developing lung.* J Clin Invest, 2020. **130**(1): p. 39-50.
7. Herriges, M. and E.E. Morrisey, *Lung development: orchestrating the generation and regeneration of a complex organ.* Development, 2014. **141**(3): p. 502-13.
8. Magnusson, M.K., O. Baldursson, and T. Gudjonsson, *Lung Epithelial Stem Cells*, in *Stem Cells & Regenerative Medicine: From Molecular Embryology to Tissue Engineering*, K. Appasani and R.K. Appasani, Editors. 2011, Humana Press: Totowa, NJ. p. 227-241.
9. Wang, Y. and N. Tang, *The diversity of adult lung epithelial stem cells and their niche in homeostasis and regeneration.* Sci China Life Sci, 2021.
10. Harkema, J.R., K.J. Nikula, and W.M. Haschek, *Chapter 14 - Respiratory System*, in *Fundamentals of Toxicologic Pathology (Third Edition)*, M.A. Wallig, et al., Editors. 2018, Academic Press. p. 351-393.
11. Lee, R.E., S.M. Miller, and S.H. Randell, *Adult Pulmonary Epithelial Stem Cells and Their Niches*, in *Encyclopedia of Tissue Engineering and Regenerative Medicine*, R.L. Reis, Editor. 2019, Academic Press: Oxford. p. 319-336.
12. Linnoila, R.I., *Functional facets of the pulmonary neuroendocrine system.* Lab Invest, 2006. **86**(5): p. 425-44.
13. Pan, J., et al., *Immunohistochemical characterization of the chemosensory pulmonary neuroepithelial bodies in the naked mole-rat reveals a unique adaptive phenotype.* PLoS One, 2014. **9**(11): p. e112623.
14. Jones-Freeman, B. and M.R. Starkey, *Bronchioalveolar stem cells in lung repair, regeneration and disease.* J Pathol, 2020.
15. Beers, M.F. and Y. Moodley, *When Is an Alveolar Type 2 Cell an Alveolar Type 2 Cell? A Conundrum for Lung Stem Cell Biology and Regenerative Medicine.* Am J Respir Cell Mol Biol, 2017. **57**(1): p. 18-27.
16. Fehrenbach, H., *Alveolar epithelial type II cell: defender of the alveolus revisited.* Respir Res, 2001. **2**(1): p. 33-46.
17. Bishop, A.E., *Pulmonary epithelial stem cells.* Cell Prolif, 2004. **37**(1): p. 89-96.
18. Barkauskas, C.E., et al., *Type 2 alveolar cells are stem cells in adult lung.* J Clin Invest, 2013. **123**(7): p. 3025-36.
19. Desai, T.J., D.G. Brownfield, and M.A. Krasnow, *Alveolar progenitor and stem cells in lung development, renewal and cancer.* Nature, 2014. **507**(7491): p. 190-4.
20. Nabhan, A.N., et al., *Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells.* Science, 2018. **359**(6380): p. 1118-1123.

21. Treutlein, B., et al., *Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq.* Nature, 2014. **509**(7500): p. 371-5.
22. Zacharias, W.J., et al., *Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor.* Nature, 2018.
23. McQualter, J.L., *Endogenous lung stem cells for lung regeneration.* Expert Opin Biol Ther, 2019. **19**(6): p. 539-546.
24. Tata, P.R. and J. Rajagopal, *Plasticity in the lung: making and breaking cell identity.* Development, 2017. **144**(5): p. 755-766.
25. Rawlins, E.L. and B.L. Hogan, *Epithelial stem cells of the lung: privileged few or opportunities for many?* Development, 2006. **133**(13): p. 2455-65.
26. Tata, P.R., et al., *Dedifferentiation of committed epithelial cells into stem cells in vivo.* Nature, 2013. **503**(7475): p. 218-23.
27. Jain, R., et al., *Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung.* Nat Commun, 2015. **6**: p. 6727.
28. Schilders, K.A., et al., *Regeneration of the lung: Lung stem cells and the development of lung mimicking devices.* Respir Res, 2016. **17**: p. 44.
29. Zheng, D., L. Yin, and J. Chen, *Evidence for Scgb1a1(+) cells in the generation of p63(+) cells in the damaged lung parenchyma.* Am J Respir Cell Mol Biol, 2014. **50**(3): p. 595-604.
30. Kapere Ochieng, J., et al., *Differentiated type II pneumocytes can be reprogrammed by ectopic Sox2 expression.* PLoS One, 2014. **9**(9): p. e107248.
31. Das, D., R.B. Fletcher, and J. Ngai, *Cellular mechanisms of epithelial stem cell self-renewal and differentiation during homeostasis and repair.* Wiley Interdiscip Rev Dev Biol, 2020. **9**(1): p. e361.
32. Kotton, D.N. and A. Fine, *Lung stem cells.* Cell Tissue Res, 2008. **331**(1): p. 145-56.
33. Liu, X. and J.F. Engelhardt, *The glandular stem/progenitor cell niche in airway development and repair.* Proc Am Thorac Soc, 2008. **5**(6): p. 682-8.
34. Borthwick, D.W., et al., *Evidence for stem-cell niches in the tracheal epithelium.* Am J Respir Cell Mol Biol, 2001. **24**(6): p. 662-70.
35. Rock, J.R., S.H. Randell, and B.L. Hogan, *Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling.* Dis Model Mech, 2010. **3**(9-10): p. 545-56.
36. Rock, J.R., et al., *Basal cells as stem cells of the mouse trachea and human airway epithelium.* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12771-5.
37. Hong, K.U., et al., *In vivo differentiation potential of tracheal basal cells: evidence for multipotent and unipotent subpopulations.* Am J Physiol Lung Cell Mol Physiol, 2004. **286**(4): p. L643-9.
38. Reynolds, S.D., et al., *Airway injury in lung disease pathophysiology: selective depletion of airway stem and progenitor cell pools potentiates lung inflammation and alveolar dysfunction.* Am J Physiol Lung Cell Mol Physiol, 2004. **287**(6): p. L1256-65.
39. Griffiths, M.J., D. Bonnet, and S.M. Janes, *Stem cells of the alveolar epithelium.* Lancet, 2005. **366**(9481): p. 249-60.
40. Hong, K.U., et al., *Clara cell secretory protein-expressing cells of the airway neuroepithelial body microenvironment include a label-retaining subset and are critical for epithelial renewal after progenitor cell depletion.* Am J Respir Cell Mol Biol, 2001. **24**(6): p. 671-81.

41. Guha, A., et al., *Neuroepithelial body microenvironment is a niche for a distinct subset of Clara-like precursors in the developing airways.* Proc Natl Acad Sci U S A, 2012. **109**(31): p. 12592-7.
42. Reddy, R., et al., *Isolation of a putative progenitor subpopulation of alveolar epithelial type 2 cells.* Am J Physiol Lung Cell Mol Physiol, 2004. **286**(4): p. L658-67.
43. Kim, C.F., et al., *Identification of bronchioalveolar stem cells in normal lung and lung cancer.* Cell, 2005. **121**(6): p. 823-35.
44. Giangreco, A., S.D. Reynolds, and B.R. Stripp, *Terminal bronchioles harbor a unique airway stem cell population that localizes to the bronchoalveolar duct junction.* Am J Pathol, 2002. **161**(1): p. 173-82.
45. Hegab, A.E., et al., *Isolation and characterization of murine multipotent lung stem cells.* Stem Cells Dev, 2010. **19**(4): p. 523-36.
46. McQualter, J.L., et al., *Endogenous fibroblastic progenitor cells in the adult mouse lung are highly enriched in the sca-1 positive cell fraction.* Stem Cells, 2009. **27**(3): p. 623-33.
47. Teisanu, R.M., et al., *Prospective isolation of bronchiolar stem cells based upon immunophenotypic and autofluorescence characteristics.* Stem Cells, 2009. **27**(3): p. 612-22.
48. Zacharek, S.J., et al., *Lung stem cell self-renewal relies on BMI1-dependent control of expression at imprinted loci.* Cell Stem Cell, 2011. **9**(3): p. 272-81.
49. Li, F., et al., *Diversity of epithelial stem cell types in adult lung.* Stem Cells Int, 2015. **2015**: p. 728307.
50. Whitsett, J.A., et al., *Building and Regenerating the Lung Cell by Cell.* Physiological Reviews, 2019. **99**(1): p. 513-554.
51. McCrimmon, D.R., G.F. Alheid, and E.J. Zuperku, *REFLEXES FROM THE LUNGS AND CHEST WALL*, in *Encyclopedia of Respiratory Medicine*, G.J. Laurent and S.D. Shapiro, Editors. 2006, Academic Press: Oxford. p. 618-626.
52. Reynolds, S.D., et al., *Neuroepithelial bodies of pulmonary airways serve as a reservoir of progenitor cells capable of epithelial regeneration.* Am J Pathol, 2000. **156**(1): p. 269-78.
53. Song, H., et al., *Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17531-6.
54. Liu, Q., et al., *Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction.* Nat Genet, 2019.
55. Salwig, I., et al., *Bronchioalveolar stem cells are a main source for regeneration of distal lung epithelia in vivo.* EMBO J, 2019.
56. Kalderon, D., *Investigating Adult Stem Cells Through Lineage analyses.* Stem Cell Rev Rep, 2022. **18**(1): p. 2-22.
57. Rawlins, E.L. and A.K. Perl, *The a"MAZE"ing world of lung-specific transgenic mice.* Am J Respir Cell Mol Biol, 2012. **46**(3): p. 269-82.
58. Hirrlinger, J., et al., *Split-CreERT2: temporal control of DNA recombination mediated by split-Cre protein fragment complementation.* PLoS One, 2009. **4**(12): p. e8354.
59. MacLean, A.L., T. Hong, and Q. Nie, *Exploring intermediate cell states through the lens of single cells.* Curr Opin Syst Biol, 2018. **9**: p. 32-41.
60. Moris, N., C. Pina, and A.M. Arias, *Transition states and cell fate decisions in epigenetic landscapes.* Nat Rev Genet, 2016. **17**(11): p. 693-703.

61. Mulas, C., et al., *Cell state transitions: definitions and challenges.* Development, 2021. **148**(20).
62. Verheyden, J.M. and X. Sun, *A transitional stem cell state in the lung.* Nat Cell Biol, 2020. **22**(9): p. 1025-1026.
63. Strunz, M., et al., *Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis.* Nat Commun, 2020. **11**(1): p. 3559.
64. Kobayashi, Y., et al., *Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis.* Nat Cell Biol, 2020. **22**(8): p. 934-946.
65. Choi, J., et al., *Inflammatory Signals Induce AT2 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration.* Cell Stem Cell, 2020. **27**(3): p. 366-382 e7.
66. Reid, A. and B. Tursun, *Transdifferentiation: do transition states lie on the path of development?* Curr Opin Syst Biol, 2018. **11**: p. 18-23.
67. Liu, Z., et al., *Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte.* Nature, 2017. **551**(7678): p. 100-104.
68. Treutlein, B., et al., *Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq.* Nature, 2016. **534**(7607): p. 391-5.
69. Liu, P., N.A. Jenkins, and N.G. Copeland, *A highly efficient recombineering-based method for generating conditional knockout mutations.* Genome Res, 2003. **13**(3): p. 476-84.
70. Scheitz, C.J. and T. Tumbar, *New insights into the role of Runx1 in epithelial stem cell biology and pathology.* J Cell Biochem, 2013. **114**(5): p. 985-93.
71. Hoi, C.S., et al., *Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin.* Mol Cell Biol, 2010. **30**(10): p. 2518-36.
72. Xu, Q., M. Tam, and S.A. Anderson, *Fate mapping Nkx2.1-lineage cells in the mouse telencephalon.* J Comp Neurol, 2008. **506**(1): p. 16-29.
73. Chen, M.J., et al., *Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter.* Nature, 2009. **457**(7231): p. 887-91.
74. Mangan, J.K. and N.A. Speck, *RUNX1 mutations in clonal myeloid disorders: from conventional cytogenetics to next generation sequencing, a story 40 years in the making.* Crit Rev Oncog, 2011. **16**(1-2): p. 77-91.
75. Goldstein, L.D., et al., *Massively parallel nanowell-based single-cell gene expression profiling.* BMC Genomics, 2017. **18**(1): p. 519.
76. Reynolds, S.D., et al., *Secretoglobins SCGB3A1 and SCGB3A2 define secretory cell subsets in mouse and human airways.* Am J Respir Crit Care Med, 2002. **166**(11): p. 1498-509.
77. Hoh, R.A., et al., *Transcriptional program of ciliated epithelial cells reveals new cilium and centrosome components and links to human disease.* PLoS One, 2012. **7**(12): p. e52166.
78. Nagendran, M., et al., *Automated cell-type classification in intact tissues by single-cell molecular profiling.* Elife, 2018. **7**.
79. Stripp, B.R., et al., *Plasticity of airway cell proliferation and gene expression after acute naphthalene injury.* Am J Physiol, 1995. **269**(6 Pt 1): p. L791-9.
80. Zheng, D., et al., *Differentiation of Club Cells to Alveolar Epithelial Cells In Vitro.* Sci Rep, 2017. **7**: p. 41661.
81. Muthusamy, N., et al., *A knock-in Foxj1(CreERT2::GFP) mouse for recombination in epithelial cells with motile cilia.* Genesis, 2014. **52**(4): p. 350-8.

82. Krestel, H.E., et al., *A GFP-equipped bidirectional expression module well suited for monitoring tetracycline-regulated gene expression in mouse.* Nucleic Acids Res, 2001. **29**(7): p. E39.
83. Schonig, K., et al., *Stringent doxycycline dependent control of CRE recombinase in vivo.* Nucleic Acids Res, 2002. **30**(23): p. e134.
84. Soriano, P., *Generalized lacZ expression with the ROSA26 Cre reporter strain.* Nat Genet, 1999. **21**(1): p. 70-1.
85. Toskala, E., et al., *Temporal and spatial distribution of ciliogenesis in the tracheobronchial airways of mice.* Am J Physiol Lung Cell Mol Physiol, 2005. **289**(3): p. L454-9.
86. Salwig, I., et al., *Imaging lung regeneration by light sheet microscopy.* Histochem Cell Biol, 2020.
87. Park, K.S., et al., *Transdifferentiation of ciliated cells during repair of the respiratory epithelium.* Am J Respir Cell Mol Biol, 2006. **34**(2): p. 151-7.
88. Wu, N.H., et al., *The differentiated airway epithelium infected by influenza viruses maintains the barrier function despite a dramatic loss of ciliated cells.* Sci Rep, 2016. **6**: p. 39668.
89. Matrosovich, M.N., et al., *Human and avian influenza viruses target different cell types in cultures of human airway epithelium.* Proc Natl Acad Sci U S A, 2004. **101**(13): p. 4620-4.
90. Voehringer, D., H.E. Liang, and R.M. Locksley, *Homeostasis and effector function of lymphopenia-induced "memory-like" T cells in constitutively T cell-depleted mice.* J Immunol, 2008. **180**(7): p. 4742-53.
91. Rawlins, E.L. and B.L. Hogan, *Ciliated epithelial cell lifespan in the mouse trachea and lung.* Am J Physiol Lung Cell Mol Physiol, 2008. **295**(1): p. L231-4.
92. Kumar, P.A., et al., *Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection.* Cell, 2011. **147**(3): p. 525-38.
93. Heaton, N.S., et al., *Long-term survival of influenza virus infected club cells drives immunopathology.* J Exp Med, 2014. **211**(9): p. 1707-14.
94. Quantius, J., et al., *Influenza Virus Infects Epithelial Stem/Progenitor Cells of the Distal Lung: Impact on Fgfr2b-Driven Epithelial Repair.* PLoS Pathog, 2016. **12**(6): p. e1005544.
95. Ray, S., et al., *Rare SOX2(+) Airway Progenitor Cells Generate KRT5(+) Cells that Repopulate Damaged Alveolar Parenchyma following Influenza Virus Infection.* Stem Cell Reports, 2016. **7**(5): p. 817-825.
96. Zuo, W., et al., *p63(+)Krt5(+) distal airway stem cells are essential for lung regeneration.* Nature, 2015. **517**(7536): p. 616-20.
97. Vaughan, A.E., et al., *Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury.* Nature, 2015. **517**(7536): p. 621-5.
98. Shen, M., Z. Luo, and Y. Zhou, *Regeneration-Associated Transitional State Cells in Pulmonary Fibrosis.* Int J Mol Sci, 2022. **23**(12).
99. Finn, J., et al., *Dlk1-Mediated Temporal Regulation of Notch Signaling Is Required for Differentiation of Alveolar Type II to Type I Cells during Repair.* Cell Rep, 2019. **26**(11): p. 2942-2954 e5.
100. Guha, A., et al., *Uroplakin 3a+ Cells Are a Distinctive Population of Epithelial Progenitors that Contribute to Airway Maintenance and Post-injury Repair.* Cell Rep, 2017. **19**(2): p. 246-254.

101. Paksa, A. and J. Rajagopal, *The epigenetic basis of cellular plasticity.* Curr Opin Cell Biol, 2017. **49**: p. 116-122.
102. Kim, T.H., et al., *Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity.* Nature, 2014. **506**(7489): p. 511-5.

9 APPENDIX

9.1 Abbreviations

µl	Microliter
µm	Micrometer
AT1	Alveolar type I
AT2	Alveolar type II
BADJ	Bronchoalveolar duct junction
BASC	Bronchioalveolar stem cell
BDPCs	Branch point associated Double Positive Cells
bp	Base pair
BSA	Bovine Serum Albumin
C	Centigrade
CCSP	Club cell secretory protein
Cyp2f2	Cytochrome P450, family 2, subfamily f, polypeptide 2
DEGs	Differentially expressed genes
DMEM	Dulbecco's Modified Eagle Medium
DTA	Diphtheria toxin fragment A
ETOH	Ethanol
FACS	Fluorescence activated cell sorting
FBS	Fetal Bovine Serum
IDT	Integrated dna technologies
mg	Milligram
ml	Milliliter
mM	Millimolar
NE	Neuroendocrine
NEBs	Neuroepithelial bodies
NEO	Neomycin
PBS	Phosphate-buffered saline
PFA	Paraformaldehydes
PFU	Plaque-forming units
RT	Room temperature

Scgb1a1	Secretoglobin, family 1A, member 1
t-SNE	t-distributed stochastic neighbor embedding
^{v}Club	Variant Club cell

9.2 List of Figures

9.3 Acknowledgements

Words cannot express how grateful I am for the support that I received during the last seven years, in particular for the support during the preparation of this thesis.

My deep gratitude to Prof. Thomas Braun for his invaluable patience, kindness, supports and scientific feedbacks throughout the whole time. I would like to express my immense gratitude and pleasure for having the opportunity to work on this project, under his supervision and be his student. It is hard to put in words how grateful I am. I learned from him when even things doesn't happen in the way they should, to remain patient and always treat people with kindness, sympathy, support, respect and deep understanding.

I would like to thank Dr. Isabelle Salwig, not only a colleague but also a very nice and respectful friend, whom if it was not because of her approach and kind attitude, I wouldn't have such a pleasant start in my PhD and my move to Germany. I am truly thankful for all her precious, untiring and comprehensive support which I will never forget. Her passion for science and deep rooted sense of responsibility and kindness, was truly an asset to have by my side. Also I am so grateful to her for sharing all her knowledge and expertise with me to run the project as strong as possible.

My grateful thanks are also addressed to Dr. Andre Snidre, Dr. Jense Preussner and Dr. Stefan Guenther for their invaluable support in designing knock-in mouse lines, processing and analysis of scRNA-seq, to Prof. Dr. Thomas Böttger and Sylvia Thomas for their processing and analysis of microarray data, to Dr. Pumree Kanrai for generation of H1N1 influenza virus and kindly sharing with me and her help for first influenza experiment. I would like to thank Birgit Spitznagel, Marion Winsent, Sonja Krueger and Susanne Kreutzer for their help in animal experiments, documentation and blastocyst injection for the knock-in strains. I would like to thank the FACS facility especially Ann Atzberger, Kerstin Richter and Kikhi Khrievono for supporting and setting up complicated sorter schemes for me. I would like to thank Dr. Alexandros Sountoulidis and Prof. Dr. Christos Samakovlis for their support in performing in situ hybridization for me.

I was so lucky and blessed to work in a department full of people with big, open hearts and always willing to help. I will never forget this. During the most challenging part of my life, they help my back so that I can pass all of those though challenges and find myself again.

The last but not the least, much love and hugs goes to my family and friends whom I am very much grateful for their beautiful supports, unconditional love and companionship every single day. Thank you and God bless you all.