Development of an organoid model for lung adenocarcinoma

Inaugural Dissertation

for

the doctoral degree of

Dr. rer. nat.

from the Faculty of Biology

University of Duisburg-Essen

Germany

Submitted by

Antonella Dost

born in Solingen, Germany

February 2021

The experiments underlying the present work were conducted at the Stem Cell Program of Boston Children's Hospital, and the Genetics Department of Harvard Medical School, in the laboratory of Prof. Dr. Carla Kim.

1. Examiner: Prof. Dr. Michael Ehrmann
2. Examiner: Prof. Dr. Carla Kim
3. Examiner: Prof. Dr. Martin Schuler

Chair of the Board of Examiners: Prof. Dr. Hemmo Meyer

Date of the oral examination: May 20, 2021

DOI: 10.17185/duepublico/74560
URN: urn:nbn:de:hbz:465-20240318-120703-5

Parts of this work are included in the following publication:

Dost, A.F.M., Moye, A.L., Vedaie, M., Tran, L.M., Fung, E., Heinze, D., Villacorta-Martin, C., Huang, J., Hekman, R., Kwan, J.H., Blum, B.C., Louie, S.M., Rowbotham, S.P., Sainz de Aja, J., Piper, M.E., Bhetariya, P.J., Bronson, R.T., Emili, A., Mostoslavsky, G., Fishbein, G.A., Wallace, W.D., Krysan, K., Dubinett, S.M., Yanagawa, J., Kotton, D.N., and Kim, C.F. (2020). Organoids Model Transcriptional Hallmarks of Oncogenic KRAS Activation in Lung Epithelial Progenitor Cells. Cell Stem Cell 27, 663-678.e8.

Acknowledgements

First and foremost, I would like to acknowledge and thank Prof. Dr. Carla Kim, who agreed to host me as a PhD student after our interview in a Café in the North End in 2016, even though I was not part of the Harvard PhD program. Carla has been a steady anchor during the ups and downs of my PhD. When I decided to drop my first project, she was supportive and gave me the freedom to come up with my own project. Of course, this kind of freedom came at the cost of great insecurity - Is the project good enough? Am I asking the right questions? Is this even relevant for the field? In these phases of self-doubt, Carla always took my issues seriously and encouraged me to keep finding my way. She pushed me early on to take advantage of any occasion to present my work, from internal seminars to international conferences, which slowly helped me get a better perspective of my own work. Carla always stayed calm and constructive, regardless of stress or frustration, and inspired the same behavior in me. She excels at making connections with collaborators, and taught me by example the value of being open and collaborative. During the coronavirus pandemic, Carla was transparent and managed the lab skillfully, always listening to our concerns, advice, and input. I will always look up to Carla, an incredibly driven woman, managing not only a lab at Harvard, but also her two daughters and family with care and inexhaustible patience.

I would also like to thank Prof. Dr. Michael Ehrmann, my mentor since the second year of my Bachelor's in 2012. Talking to Michael always left me excited about research, but most importantly, about life. His casual and laid-back way of sharing his wisdom, and his philosophy to not take everything so seriously, more than once put things into perspective for me. During my Master's, Michael encouraged me to apply to Harvard labs for my thesis ("They cook with water just like everyone else"), something I would never have considered without him. Only because of my Boston experience during my Master's did I end up in Carla's lab for my PhD.

I will also thank Dr. Aaron Moye, a post-doc in our lab who helped shape my project, and complemented it with his own work. What started as a casual comparison of data and the realization that our projects had many parallels, ended as a close collaboration and a shared first-author publication in Cell Stem Cell. Without any effort, Aaron and I complemented each other in our work, and we both greatly benefited from it. Aaron always had an open ear, and a quick question always ended up in a lengthy and inspiring discussion about science. He is in a perpetual good mood and brightens the lab with his unlimited positivity.

I would like to thank the core facilities at Harvard Medical School, Dana Farber Cancer Insti-

tute, Harvard T.H. Chan School of Public Health, and Boston Children's Hospital. Especially, I want to thank Ronald Mathieu for his expert FACS knowledge, and Mahnaz Paktinat, who helped me run most of my FACS experiments and shared her bread recipes freely with me. A big thanks to the histopathologist Roderick Bronson, whose interesting views on events and people always made me laugh out loud, and who gladly taught me the art of his trade. I want to thank my collaborators at University of California Los Angeles for providing human patient data, especially Dr. Linh Tran, Dr. Jane Yanagawa, and Dr. Steven Dubinett.

Most importantly, I want to thank every lab member of the Kim lab. I asked every single one of them for help at some point during my PhD and it was always given freely. Marghe, Patri, Caroline, and Emery, who made sure that the lab kept running as smoothly as can be expected of a place in which a bunch of scientists share reagents. Sam, for always being willing to help and for giving me valuable feedback. Julio, for being his crazy self and making work way more interesting. Carolina and Sharon, for always having an open ear. Irene and Manav, my fellow PhD sufferers. Preetie, for always being patient with me when I had trouble figuring out python code. A PhD is never easy, but with the right people around it is a journey worth all the set-backs.

Lastly, I want to thank my parents, Blanka and Jürgen Dost, who have supported me my whole life. They are my anchor and hold my roots, and have never complained (too much) that their only child decided to live abroad for most of her 20s.

Table of Contents

List of Figures

List of Tables

Abbreviations

°C	degree celsius
α	alpha
Δ	delta
µl	microliter
µm	micrometer
2D	two dimensional
3D	three dimensional
AAH	atypical adenomateous hyperplasia
Acta2	Smooth muscle actin alpha 2
ADC	adenocarcinoma
AdCMV-Cre	Adenovirus containing Cre driven by CMV promotor
AIS	adenocarcinoma in situ
ALI	air liquid interphase
AMP	*Axin2* myofibrogenic progenitor
AP-1	activator protein-1
APC	allophycocyanin
AT1	alveolar type 1
AT2	alveolar type 2
BADJ	bronchioalveolar duct junction
BASC	bronchioalveolar stem cell
BF	brightfield
BMP	bone morphogenic protein
bp	base pair
C	cluster
CAFs	cancer-associated fibroblasts
Cas9	CRISPR-associated protein 9
CD	cluster of differentiation
cDNA	complementary desoxyribonucleic acid
CF	cystic fibrosis
CMV	cytomegalovirus
CO_2	carbon dioxide
CRC	colorectal cancer
Cre	Cre-recombinase
CRISPR	clustered regularly interspaced short palindromic repeats

CSC	cancer stem cell
Cspg4	Chondroitin sulfate proteoglycan 4
CT	computed tomography
D	asparagine
d	day
DAPI	4',6-diamidino-2-phenylindole
DATP	damage-associated transient progenitor
DE	differential expression
Des	Desmin
DFCI	Dana-Farber Cancer Institute
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleic acid
dsDNA	double stranded desoxyribonucleic acid
Ecad	E-cadherin
ECM	extracellular matrix
EGF	epithelial growth factor
EGFR	epithelial growth factor receptor
EH	epithelial hyperplasia
ELF5	E74-like factor 5
Emp	empty
EMT	epithelial mesenchymal transition
EPCAM	Epithelial cell adhesion molecule
Epha2	Ephrin type-A receptor 2
ETV5	Ets Variant Transcription Factor 5
FA2	ForceAtlas2
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
Fgf	Fibroblast growth factor
fl	floxed
FMO	fluorophore minus one
G	glycine
GAP	GTPase activating proteins
Gcnt3	Glucosaminyl (N-Acetyl) Transferase 3
GDP	guanosine-diphosphate

GEF	guanine nucleotide exchange factor
GEMM	genetically engineered mouse model
GFR	growth factor reduced
GO	Gene Ontology
Gsto1	Glutathione S-Transferase Omega 1
GTP	guanosine-triphosphate
h	hour
H + E	hematoxylin and eosin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hif1a	hypoxia-inducible factor 1-alpha
HMGA2	High-mobility-group 2
HOPX	Homeodomain-only protein
HPCS	high-plasticity cell state
IF	immunofluorescence
Igfbp2	Insulin Like Growth Factor Binding Protein 2
K	LSL-Kras G12D mouse model
Klf	Krüppel Like Factor
KP	LSL-Kras G12D; p53 fl/fl mouse model
KPY	LSL-Kras G12D; p53 fl/fl; LSL-YFP mouse model
KRAS	Kirsten rat sarcoma viral oncogene homolog
KY	LSL-Kras G12D; LSL-YFP mouse model
L	leucine
Lgr	Leucine-rich repeat-containing G-protein coupled receptor
LSL	Lox Stop Lox
LUAD	lung Adenocarcinoma
Lyz2	LysozymeM
MANC	mesenchymal alveolar niche cell
MAPK	Mitogen-activated protein kinase
MAPKK	Mitogen-activated protein kinase kinase
MAPKKK	Mitogen-activated protein kinase kinase kinase
Mcam	Melanoma cell adhesion molecule
MECCs	mesenchymal cultured cells
mg	miligram
MIA	minimally invasive adenocarcinoma
min	minute

Mio	million
ml	mililiter
mM	milimolar
mRNA	messenger ribonucleic acid
mTOR	mammalian target of rapamycin
Myl9	Myosin light chain 9
NF-κB	Nuclear factor kappa B
ng	nanogram
NKX2-1	NK2 homeobox-1
NSCLC	non-small cell lung cancer
O_2	oxygen
OFE	organoid forming efficiency
P	p53 fl/fl mouse model
PBS	phosphate-buffered saline
PBS-T	phosphate buffered saline with 0.2% triton
Pdgfra	platelet-derived growth factor receptor alpha
Pdgfrb	platelet-derived growth factor receptor beta
PDO	patient derived organoid
PDPN	Podoplanin
PE	phycoerythrin
PF10	phosphate buffered saline with 10% fetal bovine serum
PFU	particles forming units
PI3K	intracellular phosphoinositide 3-kinase
PI3K	phosphoinositide 3-kinase
R	arginine
RNA	ribonucleic acid
ROCK	Rho-associated protein kinase
rpm	rounds per minute
RT-qPCR	quantitative reverse transcription polymerase chain reaction
SCA1	Stem cell antigen 1
scanpy	single-cell analysis in python
scATAC-Seq	single cell assay for transposase-accessible Chromatin-sequencing
SCC	squamous cell carcinoma
scDNA-Seq	single cell DNA-Sequencing
sci-CAR	single-cell combinatorial indexing chromatin accessibility and mRNA

SCLC	small cell lung cancer
scRNA-Seq	single cell ribonucleic acid-Sequencing
SEER	Surveillance, Epidemiology, and End Results
shRNA	small hairpin ribonucleic acid
SNPs	single nucleotide polymorphisms
SOX	SRY-related HMG-box
SPC	Surfactant protein C
Tagln	Transgelin
TCGA	The Cancer Genome Atlas
TF	transcription factor
TP53	Tumor protein p53 (human)
Trp53	Tumor protein p53 (mouse)
tSNE	t-distributed stochastic neighbor embedding
U/kg	units per kilogram
UCLA	University of California Los Angeles
UMAP	Uniform Manifold Approximation and Projection
US	United States
V	valine
Vim	Vimentin
WHO	World Health Organization
Wnt	Wingless-related integration site
wt	wildtype
Y	LSL-YFP mouse model
Ybx	Y-box binding protein
YFP	Yellow Fluorescent Protein

1 Introduction

1.1 Lung Biology

Our lungs process over 7 thousand liters of air on a daily basis. The larynx is our air gatekeeper, forming a passage to our largest airway, the trachea. The trachea is a hollow tube held open by cartilage rings, that bifurcates into the two main bronchi. The right bronchus leads to the three lung lobes that make up the right lung, while the left bronchus leads to the two lung lobes that make up the left lung, leaving space for the heart. From the two main bronchi, the airways branch out in a dichotomous manner into smaller bronchi and bronchioles with decreasing diameters, similar to an upside-down tree. Bronchioles, the narrowest entity of the airways, are only about half a milimeter wide, and terminate in the highly vascularized respiratory unit of the lung, alveoli (Hsia *et al.*, 2016). The hundred millions of alveoli have a surface area of about 130 m^2, important for efficient gas exchange between air and our blood (Gehr *et al.*, 1978; Weibel *et al.*, 1993). Oxygen (O_2) from the air enters our blood stream through diffusion, while carbon dioxide (CO_2) exits the blood to be exhaled. In order to prevent the alveolar sacs from collapsing, alveolar cells produce pulmonary surfactant that decreases the surface tension and helps to keep the bubble-like structures open (Weaver and Whitsett, 1991). In the more proximal lung, cells produce mucins, surfactants, and fluids that line the airways and trap inhaled particles, such as dust, pollutants, or allergens (Rogers, 1994; Rokicki *et al.*, 2016). Synchronously beating cilia then transport the mucus upward and out of the lungs (Braiman and Priel, 2008). This is important to keep the lungs clean of potentially harmful particles and microorganisms that are part of the thousands of liters of air we breath in daily. A very complex epithelium of highly specialized cells, including multiple progenitor and stem cells, and a microenvironment that is only starting to be revealed, are necessary to ensure the proper function of this important organ (Chen and Fine, 2016).

1.1.1 Epithelial cell types of the airways

The airways of the lungs span from the two main bronchi to the terminal bronchioles. In mice, the proximal parts of the airways are lined by a pseudostratified mucociliated epithelium that turns into a simple cuboidal epithelium around the distal airways. The airway epithelium consists of multiple stem and progenitor cells, as well as specialized differentiated cells (figure 1 A). A multitude of studies, using a variety of injury models, have revealed astounding plasticity of the epithelial airway populations.

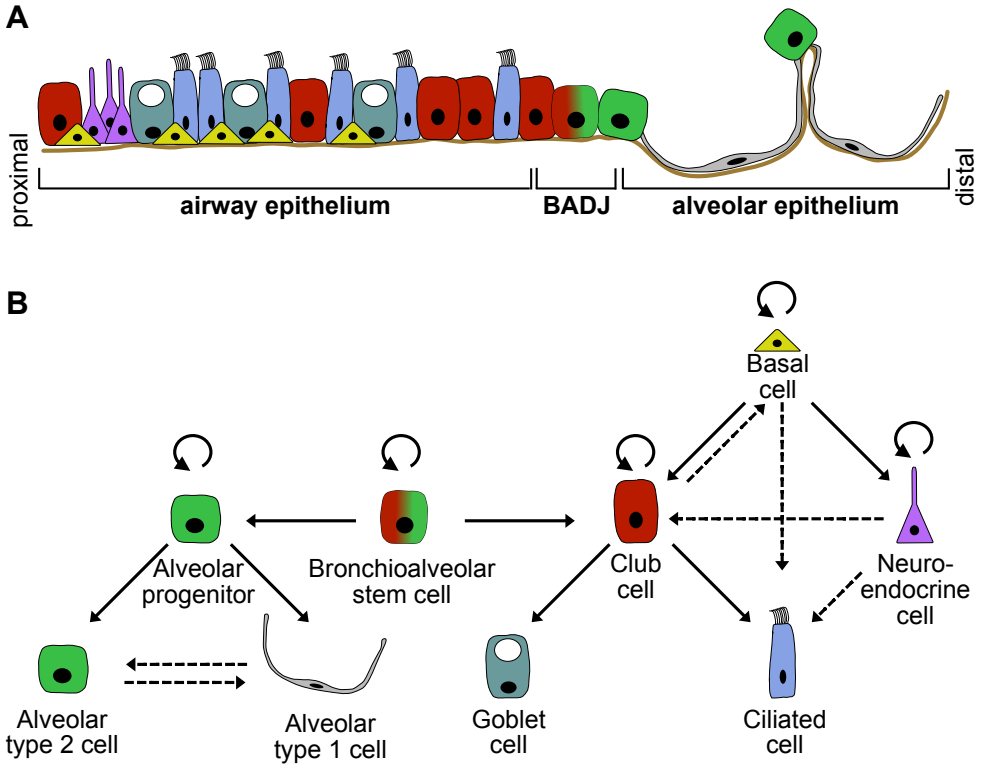

Figure 1: Overview of epithelial cells in the adult murine lung
(A) Schematic overview of epithelial cells lining the airways, bronchioalveolar duct junction (BADJ), and the alveolar region. Brown line represents basal lamina. **(B)** Hierarchy of epithelial cells in the lung. Rare cell types such as tuft cells and ionocytes were not included. Circular line: can self-renew. Straight line: gives rise to. Straight dotted line: gives rise to after severe injury.

The most important self-renewing stem cell population of the proximal airways are basal cells. They are in close contact with the basal lamina and have been shown to directly or indirectly give rise to every major epithelial cell population of the airways, including club cells, goblet cells, ciliated cells, and neuroendocrine cells (figure 1 B) (Rock *et al.*, 2011b, 2009; Rock and Hogan, 2011; Rawlins and Hogan, 2008; Rawlins *et al.*, 2009b). Club cells are also called secretory cells and are the main self-renewing stem cell population in the distal bronchi. These cells play a role in detoxifying the lungs and secrete surfactants and other fluids that line the airways. Club cells can give rise to goblet cells and ciliated cells (Chen *et al.*, 2009; Rawlins *et al.*, 2009b). Furthermore, club cells have been shown to dedifferentiate to basal cells following basal cell ablation (Rawlins *et al.*, 2009b; Tata *et al.*, 2013). Goblet cells are the primary producer of mucins, the main component of the airways-lining mucus that traps particles and microorganisms (Rogers, 1994). This mucus is transported out of the lungs by the synchronous upward beating of the cilia on ciliated cells (Braiman and Priel, 2008). Ciliated cells have

been shown to be terminally differentiated (Rawlins *et al.*, 2009b). Neuroendocrine cells are sensory cells that can respond to conditions such as hypoxia, hypercarbia, or mechanical stretch by releasing neurotransmitters, neuropeptides, or immune response activating factors (Boers *et al.*, 1996; Cutz *et al.*, 2013; Branchfield *et al.*, 2016). They contain a small stem cell population that can self-renew and give rise to club and ciliated cells after severe injury (Ouadah *et al.*, 2019). Additionally, there are multiple specialized rare cell populations such as tuft cells and ionocytes, the latter having only been described recently (Krasteva and Kummer, 2012; Plasschaert *et al.*, 2018; Montoro *et al.*, 2018). Another rare cell population resides in the bronchioalveolar duct junction (BADJ), the area where bronchioles terminate in the alveoli. These bronchioalveolar stem cells (BASCs) have only been identified in murine lungs thus far, and can give rise to club cells and alveolar cells (figure 1 B) (Kim *et al.*, 2005; Liu *et al.*, 2019a; Salwig *et al.*, 2019).

1.1.2 Epithelial cell types of the alveoli

The epithelium of the alveoli is a simple, thin-walled, squamous epithelium consisting of two major epithelial cell types (figure 1 A). Alveolar type 1 (AT1) cells are thin, stretched out, and comprise nearly 95% of the alveolar surface area. They are part of the air-blood barrier through which gas exchange takes place as a result of diffusion (Weibel, 2015). Whether AT1 cells are terminally differentiated or contain a progenitor cell sub-population remains controversial (Jain *et al.*, 2015; Wang *et al.*, 2018b). Alveolar type 2 (AT2) cells are the stem cell of the alveoli and maintain cell homeostasis. They are cuboidal in shape and produce and secrete pulmonary surfactant, which decreases the surface tension in the alveoli and prevents collapsing of the alveolar sacs. AT2 cells can self-renew and give rise to AT1 cells (figure 1 B) (Rock *et al.*, 2011a; Barkauskas *et al.*, 2013; Desai *et al.*, 2014). Multiple studies have identified subpopulations within the bulk AT2 population that only arise after severe injury and possible represent transient cell states (Chapman *et al.*, 2011; Vaughan *et al.*, 2015; Zuo *et al.*, 2015). However, more recently evidence emerged that a subset of AT2 cells displays increased progenitor function both in injury and homeostasis (Nabhan *et al.*, 2018; Zacharias *et al.*, 2018). This bipotent cell expresses similar levels of AT2 marker genes such as surfactant protein C (SPC), but displays higher wingless-related integration site (Wnt)-pathway activity, indicated by high expression of the Wnt target gene *Axin2*. In the human lung, single cell RNA-Sequencing (scRNA-Seq) also recently revealed two distinct AT2 subpopulations (Travaglini *et al.*, 2020). The bulk of AT2 cells has high expression of canonical AT2 markers and is

quiescent, while a 10 times less abundant AT2 subpopulation has lower expression of AT2 markers and high Wnt activity. Although this finding is reminiscent of the Wnt-high AT2 population identified in mice, there are transcriptional differences between the murine and human populations.

Undoubtedly, the Wnt-pathway is important for AT2 identity maintenance and stem cell function, as it prevents differentiation to AT1 cells. With the help of organoids, the signals necessary for AT2-AT1 differentiation have been revealed in recent years. Wnt pathway activation, addition of interleukin-1 , and inhibition of bone morphogenetic protein (BMP) signaling prevents AT2 cells from differentiating and keeps them in a stem cell state; in contrast, inhibition of Wnt and activation of BMP causes their differentiation to AT1 cells (Chung *et al.*, 2018; Katsura *et al.*, 2019; Weiner *et al.*, 2019; Katsura *et al.*, 2020; Choi *et al.*, 2020). Distinct differentiation states between AT2 and AT1 cells are starting to emerge but are still poorly understood (Choi *et al.*, 2020). For instance, it is generally accepted that the transcription factor (TF) Homeodomain-only protein (HOPX) is a marker of AT1 cells, a cell type that was generally thought to be terminally differentiated (Barkauskas *et al.*, 2013). However, it has been reported that HOPX+ AT1 cells are able to proliferate and recover AT2 cells after severe lung injury (figure 1 B) (Jain *et al.*, 2015). Another study found that HOPX+ AT1 cells can be split into Insulin Like Growth Factor Binding Protein 2 (Igfbp2) negative and positive cells, representing populations with progenitor cell function, and terminally differentiated cells, respectively (Wang *et al.*, 2018b). This highlights the stunning plasticity of lung stem and progenitor cells, and our gap in understanding all possible transient or persisting cell states.

1.1.3 Lung development

The development of the trachea and lungs is orchestrated by a complex interplay of TFs, signaling pathways, and epigenetic regulators. The exact molecular signals underlying this process are not completely understood; however, the major players have been identified in the past decades (figure 2). During embryonic mouse development, the lungs and trachea arise from the anterior forgut endoderm around embryonic day (E) 9.0 (Herriges and Morrisey, 2014; Whitsett *et al.*, 2019). Lung specification is determined by the expression of the TFs NK2 homeobox-1 (NKX2-1) and GATA6 in the endodermal cells on the ventral side of the anterior foregut (Zhang *et al.*, 2007). These cells become the trachea and the two lung buds that form the left and right lungs. Studies have shown that the activity of Wnt and BMP pathways play

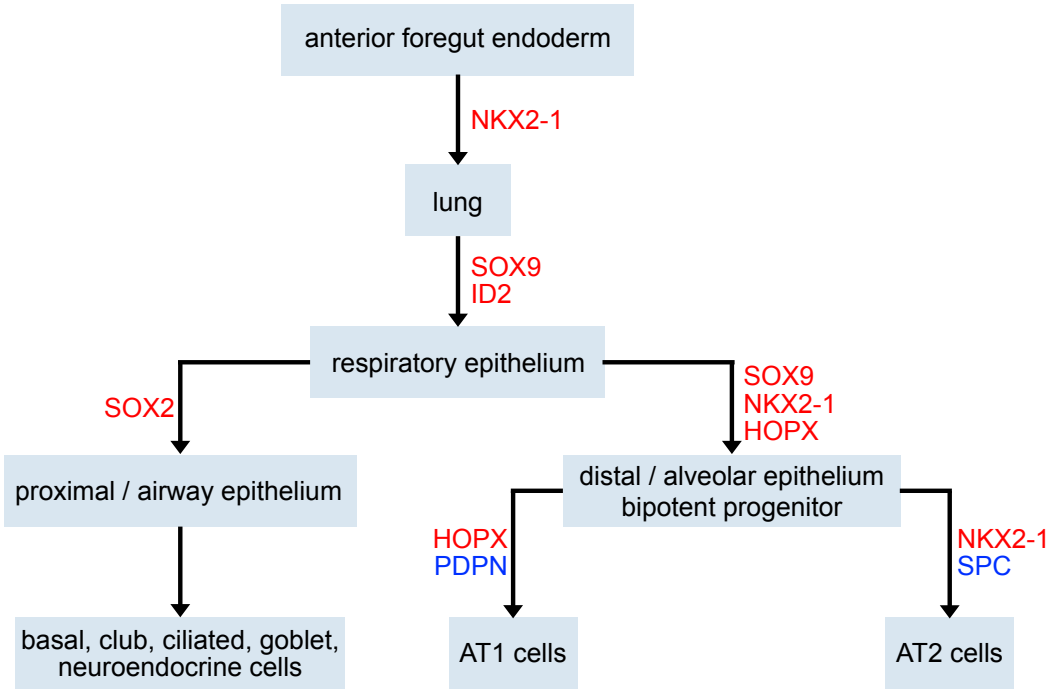

Figure 2: Overview of the development of epithelial lineages in the murine lung
Schematic shows how the epithelial lung lineages develop from the anterior forgut endoderm. Lung identity is specified by expression of NKX2-1. Respiratory epithelial progenitor cells are SOX9 and ID2 positive, then split into the proximal airway lineage and the distal lineage. The proximal lineage is defined by expression of SOX2 and includes basal, club, ciliated, goblet, and neuroendocrine cells. SOX9 expression, together with NKX2-1 and HOPX, marks a bipotent progenitor cell that eventually gives rise to the two distal alveolar lineages, AT1 and AT2 cells. Transcription factors are marked in red, non-TF proteins are marked in blue.

essential roles in specifying these respiratory progenitor cells (Harris-Johnson *et al.*, 2009; Goss *et al.*, 2009). During the next developmental stages, the two lung buds undergo branching morphogenesis and build the tree-like airway network of the lungs. A complex interplay of fibroblast growth factor (Fgf), BMP, and sonic hedgehog signaling both in the endoderm and mesoderm are thought to be critical for the proper development of the branching airways (Sekine *et al.*, 1999; Ohuchi *et al.*, 2000; Hyatt *et al.*, 2002; Pepicelli *et al.*, 1998; Bellusci *et al.*, 1997). During this process, cell identities are acquired along the proximal-distal axis. Sex determining region Y-box (SOX) proteins play an important role in this patterning process. SOX9 and Inhibitor of differentiation 2 (ID2) expressing cells, located at the distal tip, are multipotent progenitor cells that give first rise to airway, then to alveolar cell lineages (Rawlins *et al.*, 2009a; Rockich *et al.*, 2013; Chang *et al.*, 2013). Cells that exit the tip on the proximal side around E10 - 15 turn off SOX9 and upregulate SOX2 expression. These cells eventually

give rise to basal, club, ciliated, neuroendocrine, and goblet cells (Tompkins *et al.*, 2011; Que *et al.*, 2009). In contrast, cells that exit the tip around E16 - 18 turn off SOX9 and start co-expressing genes of the two alveolar epithelial cell types, such as HOPX, NKX2-1, and SPC, becoming bipotent progenitor cells (Desai *et al.*, 2014; Treutlein *et al.*, 2014; Jain *et al.*, 2015). Interestingly, one study showed that these cells shortly upregulate the TF high-mobility-group 2 (HMGA2), and that a knock out of *Hmga2* leads to an increase in SOX9 progenitor cells and a decrease in alveolar cells (Singh *et al.*, 2014). Furthermore, studies found both Wnt and BMP signaling, and members of the Ets family of TFs, namely Ets Variant Transcription Factor 5 (ETV5) and E74-like factor 5 (ELF5), to be important drivers of distal lineage identities and the alveolarization process (Weaver and Whitsett, 1991; Lu *et al.*, 2001; Mucenski *et al.*, 2003; Liu *et al.*, 2003; Metzger *et al.*, 2008). Eventually, the bipotent progenitor cells give rise to AT1 cells, expressing HOPX, Podoplanin (PDPN) and other AT1 markers, and to AT2 cells, expressing AT2 markers such as NKX2-1 and SPC (figure 2).

1.2 Lung Cancer

Lung cancer is the leading cause of cancer-related death world-wide. The World Health Organization (WHO) estimates that there were over 2 million new lung cancer cases and 1.76 million lung cancer deaths in 2018 (World Health Organization, 2018). The five-year survival rate for patients that were diagnosed with lung cancer between 1995–2009 was below 20% in most geographical areas in the developing and developed world (Allemani *et al.*, 2015). For people who were diagnosed with lung cancer in the United States (US) between 2010-2016, the 5-year survival rate was 21%, placing lung cancer among the cancer types with the worst prognosis (Howlader *et al.*, 2020). Over 2 million years of life were lost to lung cancer in the US in 2017, a result of a combination of high incidence and low survival rates (figure 3).

1.2.1 Subtypes and clinicopathological progression of lung cancer

Like most cancers, lung cancer can be categorized into localized, regional, and distant stages. Localized lung cancer is a single tumor confined to one lung. Regional cancer has direct extensions into regional tissues, including blood vessels, or extensions into regional, ipsilateral lymph nodes. Distant stage cancer has separate tumor nodules in different lung lobes, in the contralateral lung, or has metastasized to distal lymph nodes or other organs (Young

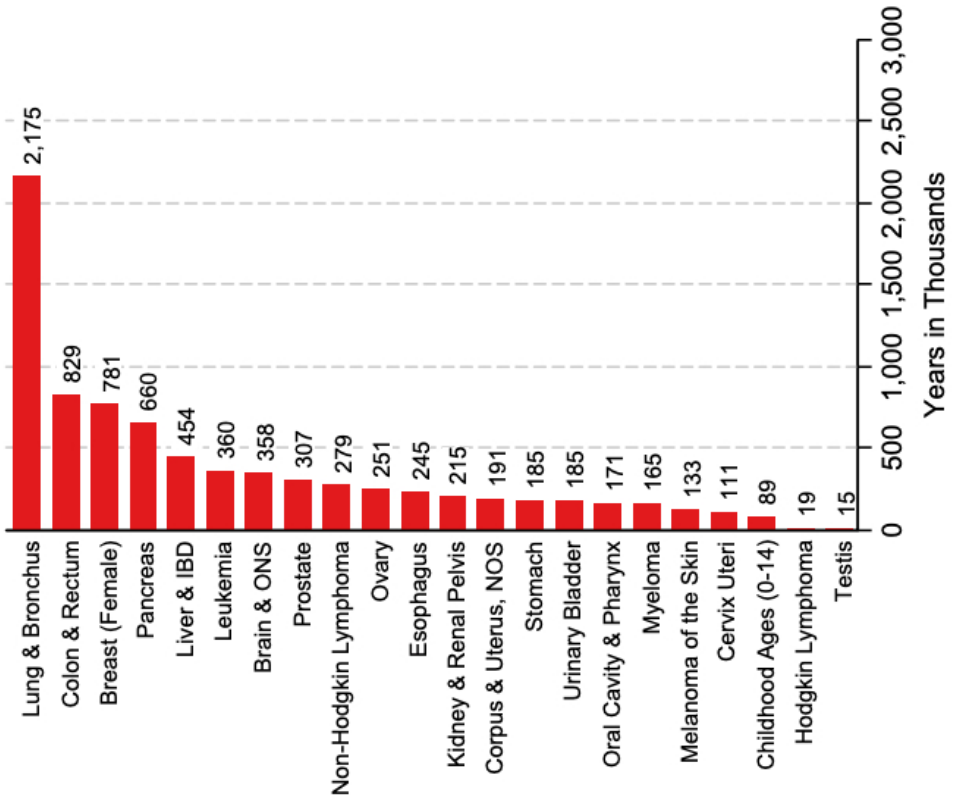

Figure 3: Statistics for years of life lost due to cancer
Graph depicting the number of person-years lost due to different cancer types in the year 2007 in the USA. Graph taken from the Surveillance, Epidemiology, and End Results (SEER) cancer statistics (Howlader *et al.*, 2020).

et al., 2001). While 5-year survival rates in the US are as high as 59% when lung cancer is diagnosed at a localized stage, it drops to under 6% when it is diagnosed at a distant stage. Unfortunately, the majority of patients (57%) is diagnosed with distant disease (Howlader *et al.*, 2020). However, new imaging methods such as low-dose computed tomography (CT) now allow for earlier detection and are implemented as screenings in risk groups, resulting in a trend towards earlier detection (figure 4) (Heuvelmans *et al.*, 2017; De Koning *et al.*, 2020).

Lung cancer can be categorized into multiple histological subtypes. The two major subtypes are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), comprising 84% and 13% of all lung cancers in the US, respectively. NSCLC can be further subdivided into lung adenocarcinoma (LUAD), making up 50% of all cases, followed by squamous cell carcinoma (SCC), making up 23% of cases (Howlader *et al.*, 2020). All histological subtypes of lung cancer have a significant association with tobacco smoking. However, the association is

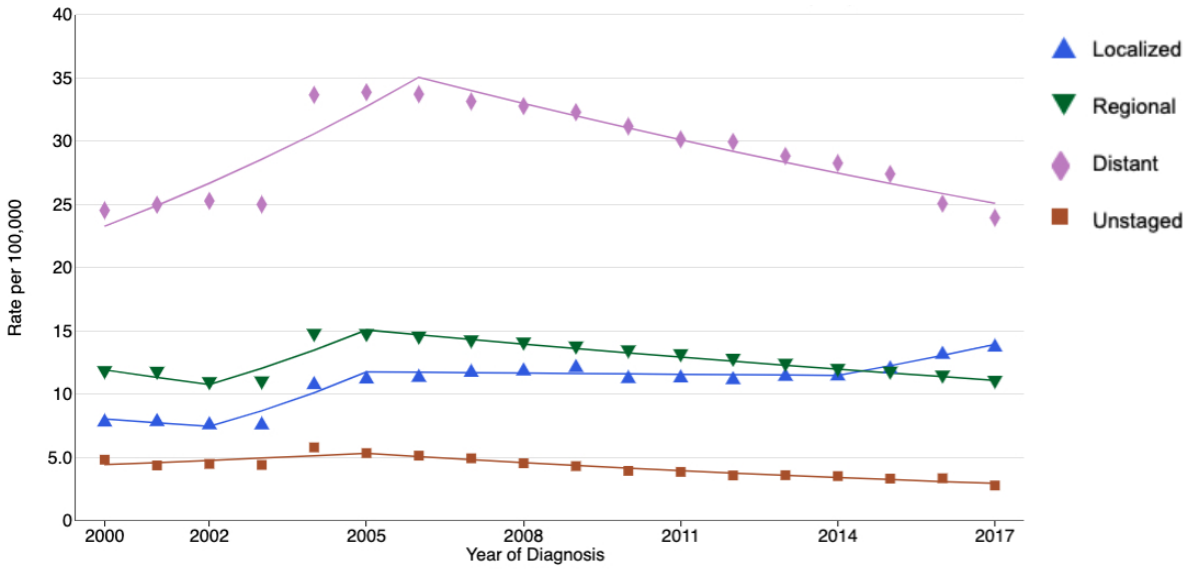

Figure 4: Trend of lung adenocarcinoma stages at diagnosis
Graph depicting the rate of people diagnosed with lung cancer at localized stage, regional stage, distant stage, or unstaged between the years 2000 and 2017 in the USA. Since 2006, diagnosis at distant stage has been decreasing, while diagnosis at localized stage is increasing. Graph taken from Surveillance, Epidemiology, and End Results (SEER) cancer statistics (Howlader *et al.*, 2020).

stronger with SCLC and SCC compared to LUAD, which is also frequently found in non-smokers (Khuder, 2001).

LUAD is the most common type of lung cancer. It progresses stepwise starting with atypical adenomatous hyperplasia (AAH), turning into adenocarcinoma in situ (AIS), then into minimally invasive adenocarcinoma (MIA), and finally into invasive LUAD (Inamura, 2018). AAH and AIS are pre-invasive lesions that usually occur in the distal lung. They are characterized by proliferation of abnormal epithelial cells along the alveolar walls without structural disruption, a pattern referred to as lepidic growth. AIS lesions are larger than AAH and display a higher degree of cellular atypia. Both lesions are difficult to detect with current imaging technology and their detection is often coincidental. MIA is a stage I tumor with predominantly lepidic growth pattern and a small invasive component that is $\leq$0.5 cm. Tumors with invasive components of $>$0.5 - 3 cm are referred to as stage I, $>$3 - 5 cm stage II, $>$5 - 7 cm stage III, and $>$7 cm stage IV (Rami-Porta *et al.*, 2017).

1.2.2 Mutational landscape of adenocarcinoma patients

Whole-exome sequencing of 412 samples of LUAD tumors and matched healthy tissue from The Cancer Genome Atlas (TCGA) database revealed the mutational landscape of LUAD (Collisson *et al.*, 2014). The most commonly mutated genes in this dataset were Tumor protein p53 (*TP53*) and Kirsten rat sarcoma viral oncogene homolog (*KRAS*), found in 46% and 33% of the samples, respectively. *TP53* is a tumor suppressor gene with important roles in cell cycle control, while *KRAS* is the most common oncogene in epithelial cancers, involved in proliferation and survival mediated through the Ras/Mitogen-activated protein kinase (MAPK)-pathway (Haigis, 2017). Notably, Epidermal growthfactor receptor (*EGFR*), another oncogene that is part this pathway, was mutated in 14% of the samples and *KRAS* and *EGFR* mutations were mutually exclusive, indicating redundant functions of the two genes. Other oncogenes that were mutated were *BRAF* (10%), *PIK3CA* (7%), *MET* (7%), and *RIT1* (2%). Mutations in tumor suppressor genes were observed in *STK11* (17%), *KEAP1* (17%), *NF1* (11%), *RB1* (4%), and *CDKN2A* (4%). Other genes that were mutated were *SETD2* (9%), *ARID1A* (7%), *MGA* (8%), *RBM10* (8%), *SMARCA4* (6%), and *U2AF1* (3%). Besides somatic mutations, LUAD patients also harbored notable gene amplifications such as *NKX2-1, TERT, MDM2, KRAS, EGFR, MET, CCNE1, CCND1, TERC, MECOM*, and *CCND3*, and gene deletions including the locus for *CDKN2A* (Collisson *et al.*, 2014; Weir *et al.*, 2007).

Smaller studies found that *KRAS* or *EGFR* mutations were usually present in pre-invasive LUAD lesions, indicating that these oncogenic mutations are initiating events in LUAD. In contrast, *TP53* mutations were significantly enriched in invasive lesions, indicating that *TP53* alterations play an important role in invasiveness during tumor progression (Chen *et al.*, 2019; Nakanishi *et al.*, 2009).

The vast majority of *KRAS* mutations in LUAD patients occurs in codon 12, followed by codon 13 and 61. In codon 12, most mutations are tobacco smoking induced transversions, leading to G12C (GGT to TGT) and G12V (GGT to GTT) alterations (Haigis, 2017). In contrast, G12D (GGT to GAT) is the most common transition in never smokers (Dogan *et al.*, 2012). All of the mentioned mutations in codon 12 result in constitutively activated KRAS protein (Hunter *et al.*, 2015).

1.2.3 Genetically engineered mouse models to study adenocarcinoma

There are multiple notable genetically engineered mouse models (GEMMs) that are used by the research community to study LUAD. Two of the most used GEMMs were developed almost two decades ago, and harbor mutations in *Kras* and *Trp53* (p53), the two most commonly mutated genes in patients (Jackson *et al.*, 2001, 2005; Collisson *et al.*, 2014). The Lox-Stop-Lox (LSL)-*Kras* G12D (K) mouse model has a Cre-recombinase (Cre)-inducible oncogenic KRAS G12D version that is knocked into the endogenous locus of one of the *Kras* alleles. Cre induction by intranasal or intratracheal delivery of lentivirus or adenovirus containing Cre driven by the ubiquitous cytomegalovirus (CMV) promoter (AdCMV-Cre) leads to excision of the stop codon in front of the oncogenic KRAS G12D sequence and to its expression driven by the endogenous KRAS promotor (DuPage *et al.*, 2009; Jackson *et al.*, 2001). Upon KRAS G12D expression, the mice develop epithelial hyperplasia (EH) in the bronchiolar region, AAH in the alveolar space, and adenomas that rarely progress to LUADs (figure 5). Thus, a single oncogenic *Kras* allele is sufficient to induce lung cancer in GEMMs. Because *TP53* is the most commonly mutated tumor suppressor gene in LUAD patient, the same group crossed the K mouse with the *Trp53*$^{floxed(fl)/wildtype(wt)}$ GEMM to obtain K; *Trp53*$^{fl/fl}$ (P) mice (Jackson *et al.*, 2005). In KP mice, KRAS G12D is expressed and both copies of *Trp53* are knocked out upon Cre expression. KP mice develop AAH, adenomas, and LUADs at a high frequency and in a shorter time frame compared to the K model, and additionally exhibit advanced tumors that metastasize to regional and distant sites (figure 5).

There are other GEMMs with oncogenic KRAS versions to induce LUAD. Overexpression of KRAS G12V under the beta-actin promotor results in lung adenomas and LUAD with short latencies (Meuwissen *et al.*, 2001). Systemic tamoxifen induced expression of KRAS G12V from its endogenous promotor leads only to lung lesions, despite frequent recombination events in other tissue sites (Guerra *et al.*, 2003). In contrast, sporadic, somatic activation of KRAS G12D in a wt or *Tp53* deficient background causes spontaneous early onset lung cancer and a variety of tumors in other organs (Johnson *et al.*, 2001). Doxycycline induced overexpression of KRAS G12D in the lung results in numerous lung tumors (Fisher *et al.*, 2001). The combination of KRAS G12D with concomitant homozygous knock-outs of tumor suppressors such as *Pten* or *Stk11* strongly augments tumor formation and metastasis, and loss of *Stk11* leads to the development of lung SCC (Iwanaga *et al.*, 2008; Zhang *et al.*, 2017a).

Moreover, mouse models with activating mutations in other components of the Ras/MAPK

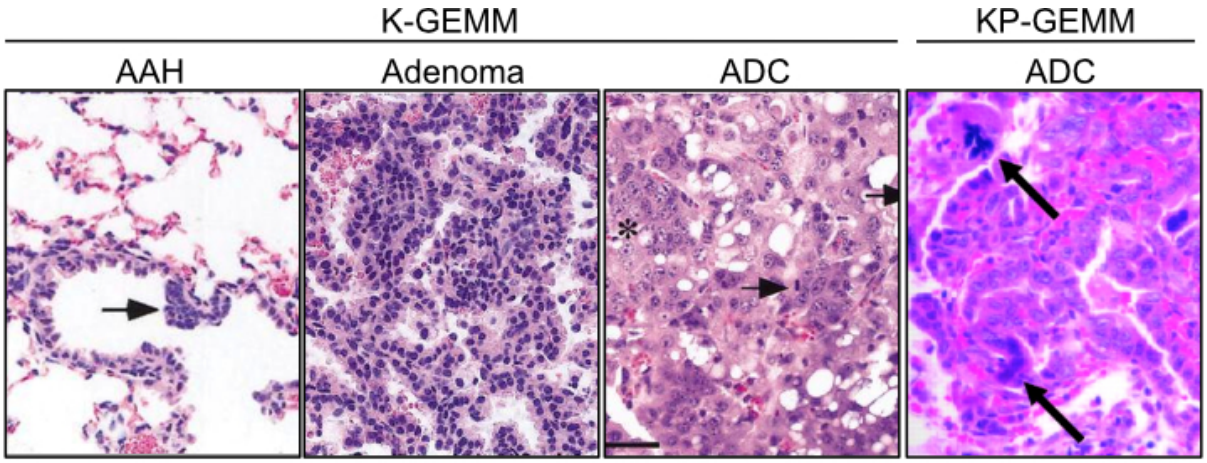

Figure 5: Lung adenocarcinoma progression in the K and KP mouse models
In the K genetically engineered mouse model (GEMM), adenocarcinoma (ADC) progression starts with atypical adenotamous hyperplasia (AAH; arrow), that turns into adenoma, and finally into ADC containing pleomorphic (star) and dividing (arrow) cells. In the KP model, giant, multinucleated cancer cells (thick arrows) appear in some ADC lesions. Images adapted from Jackson *et al.* (2001, 2005)

pathway have been established. Two doxycycline inducible EGFR mutants, EGFRL858R and EGFR ΔL747–S752, cause formation of LUAD with features of bronchioalveolar carcinoma (Politi *et al.*, 2006). Mice with an activating Braf V600E mutation develop adenomas that only rarely progressed to an advanced stage (Dankort *et al.*, 2007). Doxycycline induction of oncogenic C-RAF in lung cells leads to development of macroscopic lung tumors (Ceteci *et al.*, 2011).

There are other LUAD mouse models that use different combinations of mutations commonly found in oncogenes or tumor suppressor genes. Notably, most of them contain alterations in a Ras/MAPK pathway component, emphasizing the importance of this pathway for the formation of LUAD. Two of the most widely used GEMMs to model LUAD to date are the K and KP mouse models (Jackson *et al.*, 2001, 2005). Because oncogenic KRAS is expressed from its endogenous promotor, these models closely resemble physiological levels of KRAS expression. Furthermore, Cre induction through virus administration leads to sporadic activation of the gene alterations in few adult lung cells, rather than embryonic activation or continuous systemic activation through prolonged administration of tamoxifen or doxycycline. These conditions lead to tumor progression that recapitulates human LUAD, making the K and KP GEMMs valuable models of the lung cancer field.

1.2.4 Signaling pathways in adenocarcinoma

There are many signaling pathways that play a role in LUAD development. However, by far the most frequently activated is the MAPK-pathway. Using data from the TCGA cohort, it was concluded that 76% of all studied LUAD cases had MAPK-pathway activation (Collisson *et al.*, 2014). The MAPK signaling pathway communicates an external signal to the cell nucleus, which leads to increased survival and proliferation. An external ligand, such as epidermal growth factor (EGF), binds to a membrane-bound receptor, such as EGFR, which activates the cytoplasmatic tyrosine kinase activity of the receptor. The receptor then becomes phosphorylated, which leads to binding of docking proteins and a guanine nucleotide exchange factor (GEF). The GEF stimulates the exchange of guanosine-diphosphate (GDP) to guanosine-triphosphate (GTP) on a Ras-family protein, such as KRAS. Ras proteins are binary molecular switches that are in an "on"-state when GTP is bound, and in an "off"-state when GDP is bound. As small GTPases, they hydrolyze GTP to GDP, a process catalyzed by GTPase activating proteins (GAPs). In the "on"-state, Ras proteins activate the kinase activity of RAF protein. Through phosphorylation, MAPK kinase kinase (MAPKKK) activates MAPK kinase (MAPKK), which activates MAPK, a process known as kinase cascade. MAPK then phosphorylates downstream effectors in the nucleus which leads to the expression of target genes that promote growth, proliferation, and survival (Roberts and Der, 2007). Oncogenic mutations in Ras proteins cause a constitutive "on"-state by preventing the hydrolysis of GTP, often by rendering them insensitive to GAPs (Riely *et al.*, 2009).

TP53 is the most mutated gene in human cancer, including LUAD. TP53 pathway alterations were observed in 63% of LUAD patients (Collisson *et al.*, 2014). While TP53 has many different roles, it is best known for its ability to promote cell cycle arrest and apoptosis. This is crucial to ensure DNA damage repair, prevent mutations, and maintain genome stability. Therefore, cancers with mutations in *TP53* are often aneuploid and the genomic instability promotes the accumulation of additional mutations, potentially accelerating tumor progression and metastasis (Ciriello *et al.*, 2013; McGranahan and Swanton, 2017). As a TF, TP53 regulates expression of numerous downstream genes by directly binding to DNA (Sullivan *et al.*, 2018).

Another frequently altered pathway is the intracellular phosphoinositide 3-kinase (PI3K)-mammalian target of rapamycin (mTOR) pathway, that plays important roles in cell cycle regulation. Although PI3K-mTOR pathway members, such as *PIK3CA*, were mutated in only

7% of LUAD patients, the pathway was found to be activated in 25% of the cases (Collisson *et al.*, 2014). One possible explanation for this observed discrepancy is that PI3K can be activated by Ras proteins. Indeed, there seems to be an evident synergy between PI3K-mTOR activation and mutant KRAS (Engelman *et al.*, 2008; Gupta *et al.*, 2007; Yang and Weinberg, 2008). In short, when PI3K is active, it activates Akt through phosphorylation, which has several downstream effects, such as the activation of mTOR.

A pathway that has recently received a lot of attention in the lung field is the Wnt-pathway. Although this pathway is a well-described oncogenic pathway in other cancers such as colon cancer, components of the Wnt-pathway are only rarely mutated in lung cancer (Fodde *et al.*, 2001; Collisson *et al.*, 2014; Ding *et al.*, 2008). However, studies have found that the Wnt pathway is upregulated epigenetically in lung cancer, potentially through chronic exposure to cigarette smoke (Mazieres *et al.*, 2004; Vaz *et al.*, 2017). Furthermore, Wnt signaling mediates LUAD metastases and leads to more aggressive phenotypes when activated together with KRAS G12D in the distal lung (Nguyen *et al.*, 2009; Pacheco-Pinedo *et al.*, 2011). Indeed, there is accumulating evidence from LUAD cell lines and GEMMs that points out the importance of Wnt signaling in lung cancer progression (Stewart, 2014). Recently, a Wnt signal producing niche and a Wnt signal receiving cancer cell population were identified in the KP mouse model and connected to LUAD progression (Tammela *et al.*, 2017). Similar to the MAPK pathway, the Wnt pathway communicates an external signal to the nucleus. When a Wnt ligand binds to a receptor on the cell membrane, the complex that degrades the Wnt mediator beta-catenin becomes disrupted. This allows accumulation of beta-catenin in the cytoplasm and its translocation to the nucleus, where it recruits a transcriptional complex that binds to DNA and alters gene expression of its downstream target genes (Duchartre *et al.*, 2016).

Nuclear factor kappa B (NF-κB) is an important pathway in cancer biology because of its roles in cell survival and proliferation. Although its pathway constituents are not frequently mutated in LUAD (Collisson *et al.*, 2014), NF-κB has been found to be active in lung cancer patients (Tang *et al.*, 2006). The requirement for NF-κB signaling in KRAS driven LUAD was confirmed with cell lines and mouse models (Barbie *et al.*, 2009; Bassères *et al.*, 2010; Meylan *et al.*, 2009; Xue *et al.*, 2011). NF-κB is a dimeric TF with different subunit combinations, most notably RelA/p65. In its inactive state, the complex is sequestered in the cytoplasm by inhibiting proteins. Once activated, the complex relocates into the nucleus and binds DNA directly to promote gene expression (Mitchell *et al.*, 2016).

1.2.5 The cell of origin of adenocarcinoma

The cell of origin of lung cancer is still a topic of investigation. While this question is difficult to study in humans, there is emerging evidence from GEMMs that multiple cell types can progress to LUAD given the right genetic drivers and conditions. In humans, several biomarkers have been identified to distinguish LUAD from other lung cancer types. Many of these proteins, such as Napsin-A, NKX2-1, SP-A1, or SP-H are also highly expressed by AT2 cells, the epithelial stem cell in the alveoli (Jin *et al.*, 2018; Ye *et al.*, 2011; Zhan *et al.*, 2015). Moreover, AAH, the likely precursor lesions for LUAD, is usually located in the alveolar region. These observations in human patients lead to the hypothesis that AT2 cells are a likely cell of origin for LUAD. Recently, we analyzed two KRAS-driven patient stage IA tumors and matched healthy tissue using scRNA-Seq. Although multiple epithelial lung cell types can be detected with this method, only AT2 cells form a transcriptionally distinct cluster that is exclusively made up of tumor cells, indicating that AT2 cells are the likely cell of origin (Dost *et al.*, 2020).

Investigation of the cell of origin of LUAD was also conducted in mouse studies. Expression of KRAS G12D from the AT2 specific SPC promotor leads to the formation of adenomas and LUAD in the alveolar region (Sutherland *et al.*, 2014; Xu *et al.*, 2012). Similar results are achieved when Cre is expressed from the LysozymeM (*Lyz2*) locus, a gene only expressed in mature AT2 cells (Desai *et al.*, 2014). In another study, the researchers showed that KRAS G12V induction leads to the formation of LUAD with high AT2 marker expression in the alveolar region (Mainardi *et al.*, 2014). The expression of oncogenic KRAS from cell specific promotors of other lung epithelial cell types, or from a ubiquitous promotor, leads to cell proliferation and pre-cancerous lesions throughout the lung; however, only alveolar lesions progress to LUAD (Mainardi *et al.*, 2014; Sutherland *et al.*, 2014; Xu *et al.*, 2012; Desai *et al.*, 2014). Finally, in an unbiased approach, we induced KRAS G12D using an AdCMV-Cre virus and analyzed early lesions with scRNA-Seq. Similar to the observations in the patient samples, only AT2 cells form a transcriptionally distinct cluster upon oncogenic KRAS expression, making them a strong candidate for the cell of origin for LUAD (Dost *et al.*, 2020).

1.2.6 Adenocarcinoma progression on molecular level

In a multitude of studies both in mice and humans, researchers have tried to define the molecular drivers of LUAD tumorigenesis. Two important notions are now well established: First, TFs that play important roles in developmental processes in the lung are also drivers of tumor progression and metastasis. Second, a subset of LUAD cancer cells adapts transcriptional programs of other cell types and organs, referred to as lineage infidelity.

About 75-85% of pulmonary LUAD stain positive for the lung lineage TF NKX2-1 (Stenhouse *et al.*, 2004; Kunii *et al.*, 2011). Tumors that are negative for NKX2-1 often display areas with poorly-differentiated cells and patients have a worse prognosis (Berghmans *et al.*, 2006; Barletta *et al.*, 2009). Studies in the KP mouse model revealed that loss of NKX2-1 causes an upregulation of HMGA2, a TF important during lung development but mostly absent from healthy adult lung epithelial tissue (Winslow *et al.*, 2011). Cells that stain negative for NKX2-1 and positive for HMGA2 are mainly present in metastases and in late-stage metastatic primary tumors. Interestingly, another study conducted in the K mouse model, which produces less aggressive tumors, did not find an increase in HMGA2 expression after the deletion of *Nkx2-1* (Snyder *et al.*, 2013). Instead, the oncogenic KRAS expressing cells lose their AT2 identity and adopt a gastric-like phenotype that resembles mucinous ADC, a rare subtype of LUAD. Furthermore, there is no evidence for metastasis in the K model after *Nkx2-1* deletion, indicating that additional factors such as HMGA2 are necessary to adopt a metastatic phenotype.

The two SOX proteins determining proximal-distal lineage identities during lung development have also been shown to play a role in LUAD progression. High expression of SOX9 is correlated with higher stage, metastasis, and poor prognosis in LUAD patients (Zhou *et al.*, 2012; Huang *et al.*, 2019). A recent scRNA-Seq study of human LUAD tumors found that SOX2 and SOX9 are highly expressed in local and distal metastases (Laughney *et al.*, 2020). Although absent from primary tumors on protein level, the authors found subsets of cells in primary tumors that are transcriptionally similar to the metastatic cells. Furthermore, the study reported that tumor cells co-express a variety of genes that are usually restricted to one cell type in the adult lung, including club, basal, AT1, and AT2 genes, indicating lineage infidelity of cancer cells (Laughney *et al.*, 2020). Another recent scRNA-Seq study examined different time points of tumor progression in the K and KP mouse models (Marjanovic *et al.*, 2020). In this study, the authors identified a high-plasticity cell state (HPCS) population that emerges about 12

weeks after tumor initiation. They showed that this population can give rise to other cell states, including cells that transcriptionally resembled gastric cells, embryonic liver cells, and cells that undergo epithelial mesenchymal transition (EMT). Furthermore, they identified a mixed AT1/AT2 cell state and hypothesize that this population emerged early during tumor progression and gives rise to the HPCS cells (Marjanovic *et al.*, 2020). In a parallel study, single cell Assay for Transposase-Accessible Chromatin-sequencing (scATAC-Seq) was used to determine epigenetic cell state transitions in the KP LUAD model (LaFave *et al.*, 2020). They observed that some cancer cells share open chromatin regions with both AT1 and AT2 cells, providing further evidence for a bipotent cell state. Furthermore, they found that the loss of lung identity and the emergence of a metastatic phenotype is accompanied by increased access to the Runt Box 2 (RUNX2) TF site.

The majority of published studies focused on late-stage LUAD and the metastatic phenotype. Unfortunately, much less is known about the early steps of tumorigenesis. In a recent study, inspired by the work of this thesis, we looked at the early transcriptional changes in two KRAS-driven LUAD stage IA patients with single cell resolution (Dost *et al.*, 2020). Strikingly, the data indicated that, even in early-stage tumors, AT2 lineage gene expression is decreased, indicating that a loss of identity might occur earlier than previously thought. In the same study, we showed a similar phenotype in the K mouse model. Our knowledge of the molecular changes in early-stage lung cancer is still in its infancy; however, new technologies such as scRNA-Seq and organoid cultures will facilitate the study of this important field in future.

1.3 Organoids

Organoids have rightly received increasing attention in the past decade. Their astounding resemblance to *in vivo* processes has made them an important tool for a wide range of basic research and clinical applications. The term organoid was used in several different contexts throughout the last century. During the 1950s and ‘60s, it was often used to describe intracellular structures, also referred to as organelles, or in the context of tumors and abnormal cell growth of the skin, termed organoid nevus (Smith and Cochrane, 1946; Mehregan and Pinkus, 1965). However, starting in the ‘80s, the term increasingly described three-dimensional (3D) aggregates of normal or cancerous tissue *in vitro*. The growth of cells in three dimensions was facilitated by work on the extracellular matrix (ECM) in the ‘60s and ‘70s, when a gel-like substance rich in laminin, collagen, and other ECM proteins was isolated from chondrosarco-

mas (Swarm, 1963). Similar basement membrane matrices are now commercially available, such as a product called Matrigel (Kleinman and Martin, 2005). The meaning of the term organoid as we use it today was redefined in 2009, when Sato *et al.* showed that a single adult intestinal stem cell plated in Matrigel is able to self-organize into intestinal crypt–villus units that expand into organoids. Because of the resemblance of these organoids to the intestinal structure, they are commonly referred to as "mini guts" (Sato *et al.*, 2009). Although the field has not decided on one definition, it is mostly understood that organoids consist of several organ-specific cell types that derive from a stem or progenitor cell, that self-organizes through cell sorting and spatially restricted lineage commitment (Clevers, 2016; Eiraku and Sasai, 2012; Lancaster and Knoblich, 2014).

1.3.1 Advantages of organoid research

There is a variety of applications for organoids, ranging from studying very basic questions of stem cell and developmental biology, various diseases including cancer, or testing drug response as part of a personalized medicine approach. In many ways, organoids build a bridge between two dimensional (2D) *in vitro* studies and *in vivo* animal studies, as well as between animal studies and human studies (figure 6).

The first bridge can be explained with the example of cancer, a complex disease that has been studied for a long time. The availability of cancer cell lines and other *in vitro* cancer models has come a long way since the establishment of the first cancer cell line in 1951, the HeLa cell line (Beskow, 2016). Cancer cells grown in 2D cultures are easy to maintain and manipulate, and can therefore be used in high-throughput assays. Cell lines greatly advanced our understanding of cancer; however, their establishment is difficult and they often fail to recapitulate *in vivo* processes. Because there is no limit to their growth, cancer cell lines have often been cultured and passaged for years. Due to the unstable nature of cancer cell genomes, mutations occur, and fast-growing clones get selected for. Studies have shown that the same cell line cultured in different labs has vastly different properties and drug responses, questioning the reproducibility of studies conducted with cancer cell lines (Ben-David *et al.*, 2018; Liu *et al.*, 2019b). Furthermore, cell signaling, metabolism, cell-cell interactions, and cell-ECM interactions more closely resemble physiological processes when cells are grown in a 3D ECM scaffold (Lee *et al.*, 2007; Mazzocchi *et al.*, 2019; Mazzoleni *et al.*, 2009; Pickl and Ries, 2009). Before the advance of 3D organoid cultures, animal models were the only way to

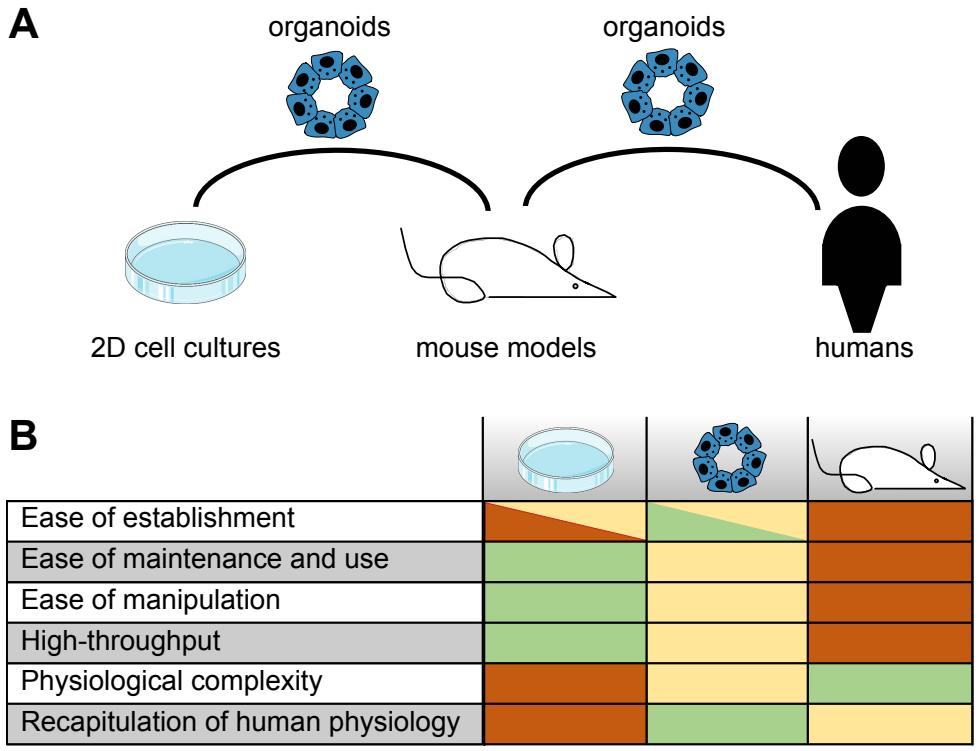

Figure 6: Comparison of 2D cell cultures, mouse models, and organoids
(A) Organoids represent bridge between 2D cell cultures and mouse models, and between mouse models and humans. **(B)** Tabular comparison of features of 2D cell cultures, mouse models, and organoids. Green = best, yellow = in between, red = worst.

study many of those relevant processes. However, mouse models are difficult to establish and costly, and are unsuitable for high-throughput screenings. Organoid cultures maintain some physiological complexity, are easier to establish, maintain, and manipulate than animal models, and are partially suitable for high- throughput applications. Therefore, organoids bridge 2D cell lines and *in vivo* animal models (figure 6).

The second bridge they represent is more obvious. Even though animal models such as GEMMs are invaluable tools for our understanding of physiological processes, they often lack in recapitulating human physiology. This is especially evident in the field of stem cell and developmental biology. For instance, the alveolarization process in the lungs takes place post-natally in mice, while it starts around week 36 of gestation in humans (Silva *et al.*, 2015). Furthermore, stem cell and progenitor populations differ between mice and humans, often questioning the human relevance of findings in mouse models (Chen and Fine, 2016; Nikolić *et al.*, 2017). Moreover, conditions for culturing human stem cells differ from mouse cultures, indicating that signals necessary to maintain stem cells can be species dependent (Kim *et al.*, 2020). Therefore,

being able to study biological processes with human cell-derived organoids, in a physiologically relevant context that can also be manipulated, bridges the gap between *in vivo* animal models and human studies (figure 6).

1.3.2 Organoids in stem cell and developmental biology

Since the groundbreaking work of Sato *et al.*, many groups have optimized culture conditions for organoids derived from adult stem cells of numerous tissues, including stomach, colon, pancreas, liver, prostate, lung, and others (Barker *et al.*, 2010; Sato *et al.*, 2011; Huch *et al.*, 2013a,b; Chua *et al.*, 2014; Karthaus *et al.*, 2014; Lee *et al.*, 2014; Rock *et al.*, 2009). Not only adult stem or progenitor cells can serve as cell of origin for organoids, but also tissue fragments, induced pluripotent stem cells, or embryonic pluripotent stem cells (Nikolić *et al.*, 2017; Kurmann *et al.*, 2015; Takebe *et al.*, 2013; Wong *et al.*, 2012; Eiraku *et al.*, 2011; McCracken *et al.*, 2014; Nakano *et al.*, 2012). In the stem cell field, one of the questions that researchers have pursued for a long time is how stem cells self-organize into complex structures through differentiation, proliferation, and signaling. This is relevant during embryogenesis and development, but also in adult organisms. The natural turn-over of cells in tissues, but also the repair of injuries, involve stem or progenitor cells. Being able to study these processes in a 3D *in vitro* system has greatly advanced our understanding of the underlying intrinsic and extrinsic signals involved. One impressive example of the power of organoids is the generation of cerebral organoids, or "mini brains". Lancaster *et al.* described in 2013 how human pluripotent stem cells self-organize into cerebral organoid structures that have an astounding resemblance to the developing brain and display neural activity (Lancaster *et al.*, 2013). A subsequent study compared the transcriptional landscape of cerebral organoids to fetal neocortex using scRNA-Seq and found that the genetic programs are very similar (Camp *et al.*, 2015). This is only one of many examples that highlights the utility of organoids to study developmental processes *in vitro*.

The lung field has used organoids to study stem cell behavior for over a decade. Basal cell-derived organoids, sometimes termed tracheospheres or bronchospheres, depending on the origin tissue, can be grown both from mice and humans (Rock *et al.*, 2009; Barkauskas *et al.*, 2013; Tata *et al.*, 2013; Danahay *et al.*, 2015). Club cells and BASCs can also be grown in organoid culture; however, they require a co-culture with supporting stromal cells to grow (McQualter *et al.*, 2010; Chen *et al.*, 2012; Lee *et al.*, 2014). Until very recently, this was also

true for AT2-derived organoids, often called alveolospheres or alveolar organoids. Different stromal cell types can be used for the organoid co-cultures, including fibroblasts, endothelial cells, and macrophages (Chen *et al.*, 2012; Barkauskas *et al.*, 2013; Lee *et al.*, 2014; Lechner *et al.*, 2017). However, in the past couple of years, stromal-free culture conditions have been described for murine and human AT2-derived organoids (Weiner *et al.*, 2019; Choi *et al.*, 2020; Katsura *et al.*, 2020; Salahudeen *et al.*, 2020). This was only possible through careful analysis of the signals and growth factors important for AT2 cell maintenance, discussed in detail in section 1.1.2.

1.3.3 Animal model-derived and *in vitro* engineered cancer organoids

There are two general approaches to generate murine cancer organoids. The first strategy is a holistic approach: Tumors are induced in GEMMs, and tumor cells are isolated from the *in vivo* tumors and cultured in 3D conditions (figure 7). This has been done for multiple cancer types, including intestinal cancer, pancreatic cancer, breast cancer, prostate cancer, and others (Sato *et al.*, 2011; Boj *et al.*, 2015; Duarte *et al.*, 2018; Wadosky *et al.*, 2019). While these studies deliver insights about tumor progression, niche requirements, signaling, or drug responses, tumor initiation and early steps of tumor progression take place *in vivo* and can therefore not be modeled *in vitro* using this strategy.

To overcome those limitations, some labs have started using a reductionist approach: Cancer organoids are engineered by isolating the cell of origin and introducing transforming mutations *in vitro* (figure 7). The requirements for this approach are that the cell of origin of the cancer is known and can be isolated, and that gene alterations can be introduced *in vitro* using inducible alleles or gene editing techniques. In one study, Li *et al.* made use of the already existing Cre inducible K, P, and KP mouse models. The researchers used the GEMMs to isolated the cells of origin for pancreatic and gastric cancer, and infected the cells with a Cre-expressing adenovirus *in vitro*, before plating them in 3D culture (Li *et al.*, 2014). In both cases, they found that the resulting organoids exhibit dysplasia and form ADCs when transplanted subcutaneously. Furthermore, to generate colorectal cancer (CRC) organoids, they isolated the cell of origin from $Apc^{fl/fl}$; *villin*-CreER mice and induced Cre expression by treating the cells with tamoxifen *in vitro*. Because CRC requires multiple hits in tumor suppressors and oncogenes, the authors then infected the *Apc* deleted cells with retroviruses encoding *Kras* G12D, and/or small hairpin ribonucleic acid (shRNAs) for *p53* and *Smad4* (Fearon and

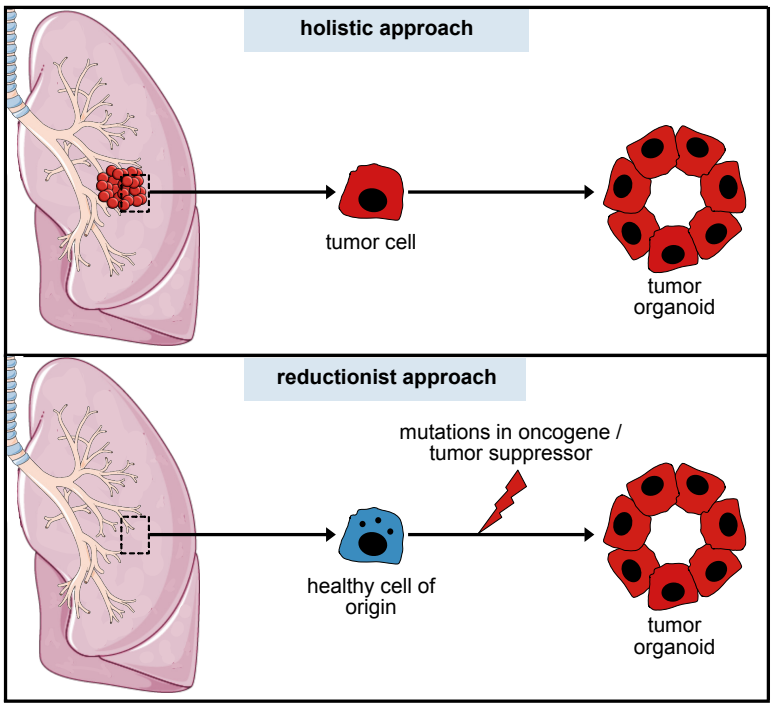

Figure 7: Holistic and reductionist approaches to obtain cancer organoids
In the holistic approach, tumor cells are isolated from an *in vivo* tumor and cultured as cancer organoids. In the reductionist approach, a healthy cell of origin is isolated and transforming mutations are introduced *in vitro*. The mutated cells are then cultured as cancer organoids. The schematic shows the lung as the origin for the cells, but this approach can be applied to any organ. Furthermore, cells can come from humans, mice, or other species.

Vogelstein, 1990). They found that different combinations of those gene alterations result in different degrees of transformation, with the 4-gene module resulting in severe colorectal ADC (Li *et al.*, 2014). In this study, the authors used a combination of Cre dependent GEMMs and shRNAs to engineer cancer organoids *in vitro*. Another widely used gene editing tool is the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system. In 2015, two groups used CRISPR/Cas9 to sequentially alter the same genes used in the previous study to generate CRC organoids (Drost *et al.*, 2015; Matano *et al.*, 2015). Because CRISPR/Cas9 is less efficient than Cre or shRNA mediated gene editing, the groups used different culturing conditions between the editing steps to select for cells that have successfully integrated the gene alterations. Both groups came to similar conclusions as Li *et al.*; however, their approach did not require a specific GEMM. In theory, CRISPR/Cas9 or other gene editing technologies can be used to introduce any gene alteration to engineer cancer organoids *in vitro*, as long as the cell of origin can be isolated. Similar approaches were used to model brain cancer in cerebral organoids, introducing as many as 18 single

gene alterations and 15 common combinations observed in brain tumors, bladder cancer in urothelial organoids, and others (Bian *et al.*, 2018; Mullenders *et al.*, 2019). In the lung field, it was demonstrated that multiple lung epithelial cells can survive adenovirus mediated *in vitro* induction of KRAS G12D in organoid cultures (Zhang *et al.*, 2017a). However, no further studies were done to characterize *in vitro* engineered lung cancer organoids.

1.3.4 Patient-derived cancer organoids and personalized medicine

Patient-derived organoids (PDOs) are 3D cultures of tumor cells or normal cells from patient biopsies or resected primary or metastatic tumors that grow in organoid culture indefinitely and can be established with high success rates. In the last years, protocols to establish PDOs have been published for a variety of cancer types, including colon cancer, CRC, prostate cancer, pancreatic cancer, glioblastoma, gastrointestinal cancer, lung cancer, and others (Sato *et al.*, 2011; Fujii *et al.*, 2016; Kondo *et al.*, 2011; Weeber *et al.*, 2015; Gao *et al.*, 2014; Boj *et al.*, 2015; Hubert *et al.*, 2016; Vlachogiannis *et al.*, 2018; Kim *et al.*, 2019; Sachs *et al.*, 2019). Almost all studies showed that PDOs derived from tumors maintain the heterogeneity, genetic profile, and histological features of the original tumor. Furthermore, there is accumulating evidence that PDOs can predict patient responses to drug therapy, an important step towards personalized medicine. For instance, one recent study tested chemotherapeutic and targeted agents on PDOs derived from metastatic gastrointestinal cancer and compared the results with the patient responses. They found a striking 88% positive predictive value (prediction that a drug will work) and a 100% negative predictive value (prediction that a drug will fail) (Vlachogiannis *et al.*, 2018). Many other groups have published promising data indicating that the predictive value of PDOs is indeed very high. In the case of lung cancer, two groups reported in 2019 that PDOs from primary and metastatic lung cancer responds to different targeted agents in line with their mutational backgrounds (Kim *et al.*, 2019; Sachs *et al.*, 2019).

Because of the overall promising results using PDOs, researchers have started to generate living organoid biobanks, large collections of PDOs that are characterized genetically and histologically. In 2015, Van De Wetering *et al.* established such a biobank from 20 CRC patients. The organoids closely resemble the tumor they were derived from and genetic changes were similar to previously published mutational analyses of this type of tumor (Van De Wetering *et al.*, 2015). Another example is a biobank of $>$100 breast cancer organoid lines derived from

primary and metastatic tumors, matching the histopathology and the receptor status of the original tumor (Sachs *et al.*, 2018). More organoid biobanks from a variety of cancer types are becoming available, including gastric cancer, bladder cancer, and others (Yan *et al.*, 2018; Lee *et al.*, 2018; Pauli *et al.*, 2017; Schutgens and Clevers, 2020).

1.3.5 Other applications of organoids

One of the first clinical applications for PDOs was in the context of the hereditary monogenetic disease cystic fibrosis (CF) (Riordan *et al.*, 1989). CF is caused by a defect in a chloride channel gene, which leads to thickening of mucus in the digestive system and the lungs (Guggino and Banks-Schlegel, 2004). In 2013, Dekkers *et al.* developed an assay to facilitate diagnosis and personalized medicine approaches in CF using intestinal organoids (Dekkers *et al.*, 2013). Indeed, this assay predicts patient drug responses with high accuracy and is now implemented in clinical diagnostics in the Netherlands (Berkers *et al.*, 2019). Other congenital conditions that were studied using PDOs include primary microcephaly and idiopathic autism spectrum disorder in cerebral organoids, α1-antitrypsin deficiency and Alagille syndrome in liver organoids, and more (Lancaster *et al.*, 2013; Mariani *et al.*, 2015; Huch *et al.*, 2015).

Another field of study that benefits from organoid technology is the field of infectious diseases. Many studies have shown that organoid responses to pathogens are similar to *in vivo* responses. Gastric organoids have been used to study infections of Helicobacter pylori (Bartfeld *et al.*, 2015; McCracken *et al.*, 2014); cerebral organoids for Zika virus infections with initial reports published merely three months after it was declared a global health emergency by the WHO (Cugola *et al.*, 2016; Garcez *et al.*, 2016; Qian *et al.*, 2016); intestinal and lung organoids to model SARS-CoV-2 infections with first reports published only four months after the WHO declared it a pandemic (Lamers *et al.*, 2020; Katsura *et al.*, 2020).

Highlighting the diversity of organoid applications, a study from 2020 described the first organoids derived from venom gland cells from nine different snake species, the first reptilian tissue derived organoids (Post *et al.*, 2020). These long-term organoid cultures allow for harvesting of snake venom components, facilitating the study of new therapies to treat snake bites.

1.3.6 Limitations and perspective of organoid research

Organoids have already proven to be a useful tool to study stem and developmental biology, cancer, hereditary diseases, infectious diseases, and more. PDOs are starting to be implemented in clinical practices, a hopeful step towards personalized medicine. Living biobanks containing PDOs derived from both cancerous tissue and matched healthy tissue can help select effective treatments that are not toxic to healthy cells, an important consideration when choosing a therapy. However, despite these promising advances the organoid field still stands at the very beginning and has limitations that need to be considered.

Most organoid conditions that are currently used allow for the growth of epithelial cells in an ECM scaffold without other supporting cells present. While these well-defined cultures allow for precise manipulation using growth factors or small molecules, important information gets lost because of the absence of other cell types that make up blood vessels, nerves, the immune system, or other stromal niches. Especially the lack of vascularization greatly limits the size that organoids can reach because nutrients can not enter the inside of large organoids efficiently and cells at the center become necrotic. In the brain organoid field, researchers partially overcame this issue by growing organoids in spinning bioreactors. Due to the forced movement of the media, nutrients reach the center of the organoids more easily and necrosis can be avoided (Lancaster *et al.*, 2013). However, this system is very costly and not an option for less well funded labs. To overcome this issue with other strategies, there are efforts to implement co-cultures of epithelial cells and supporting stromal cells. For instance, in 2017, two groups added neural crest cells to their intestinal organoid cultures to model a functional intestinal enteric nervous system (Schlieve *et al.*, 2017; Workman *et al.*, 2017). In the lung field, different types of lung organoids have been successfully cultured with endothelial cells, fibroblasts, and macrophages (Lee *et al.*, 2014; Barkauskas *et al.*, 2013; Lechner *et al.*, 2017).

Another limitation important for efficient drug testing in organoids is their scalability. 3D cultures are more expensive and time consuming than 2D cultures, and the use of animal-based ECM might lead to batch variations and concerns regarding reproducibility (Yin *et al.*, 2016). The optimization of organoid drug screening platforms in terms of robustness and sensitivity is therefore crucial in order to eventually implement these practices in clinics.

Despite these limitations, organoid technology is on the track of being a widely used tool both in basic research and clinical contexts. Complementary to already existing models, organoids

build a valuable bridge between 2D cultures and animal models, but also between animal models and human patients. With rapidly evolving culturing conditions, the emergence of co-cultures of different cell types, and the optimization of gene editing technologies, organoid technology holds great future promise for drug testing, disease modeling, and personalized medicine.

1.4 Aim and objectives of the work

KRAS-driven lung cancer is a common and deadly disease with no approved targeted therapies. The effect of oncogenic KRAS expression on primary epithelial cells is not well understood. Organoids are a promising new tool to study KRAS-driven tumorigenesis; however, current lung cancer organoid models are derived from already existing tumors, a strategy that does not allow for analysis of the early steps of cancer progression. With this work, I took a reductionist approach to establish a new LUAD organoid model to elucidate the transcriptional signals that lead to cancer progression *in vitro*. I used this new tool to study all steps of cancer progression, starting with the transformation of the cell of origin. This approach allowed for the study of basic biology questions, revealed transcriptional vulnerabilities of cancer cells, and provided a tool to test new drug targets.

Firstly, I generated and characterize cancer organoids and compared them to healthy counterparts. Organoids were derived from murine AT2 cells, the cell of origin of LUAD. I induced the transforming *Kras* G12D and *Tp53* mutations *in vitro*. Organoid cultures were analyzed for hallmarks of cancer, such as pleomorphic features and increased proliferation. To test if the *in vitro* transformed cells could initiate and propagate tumors *in vivo*, I transplanted them orthotopically. This establishment of a new cancer organoid tool allowed for further study of cancer cell biology (figure 8).

Secondly, I used the newly established cancer organoids to study the transcriptional landscape of KRAS G12D-driven cancer cells shortly after oncogene activation. This single oncogene-driven phenotype allowed for the study of genes that are important in early tumorigenesis, an understudied area of research with great promise. The analysis of transcriptional changes with single cell resolution allowed me to reveal heterogeneity in cell populations, an important consideration when studying cancer. With this approach, transcriptional vulnerabilities of the cancer cells were revealed and new potential targets were tested (figure 8).

Lastly, I expanded the single cell transcriptional investigation to a time course analysis of healthy and cancerous organoid cells, driven by KRAS G12D and loss of TP53. Understanding the progression of healthy alveolar organoids helped me point out differences to the progression of cancer organoids over time. With this time course approach, cell trajectories could be directly compared and the transcriptional drivers underlying organoid progression were revealed, allowing for the identification of new cancer cell vulnerabilities (figure 8).

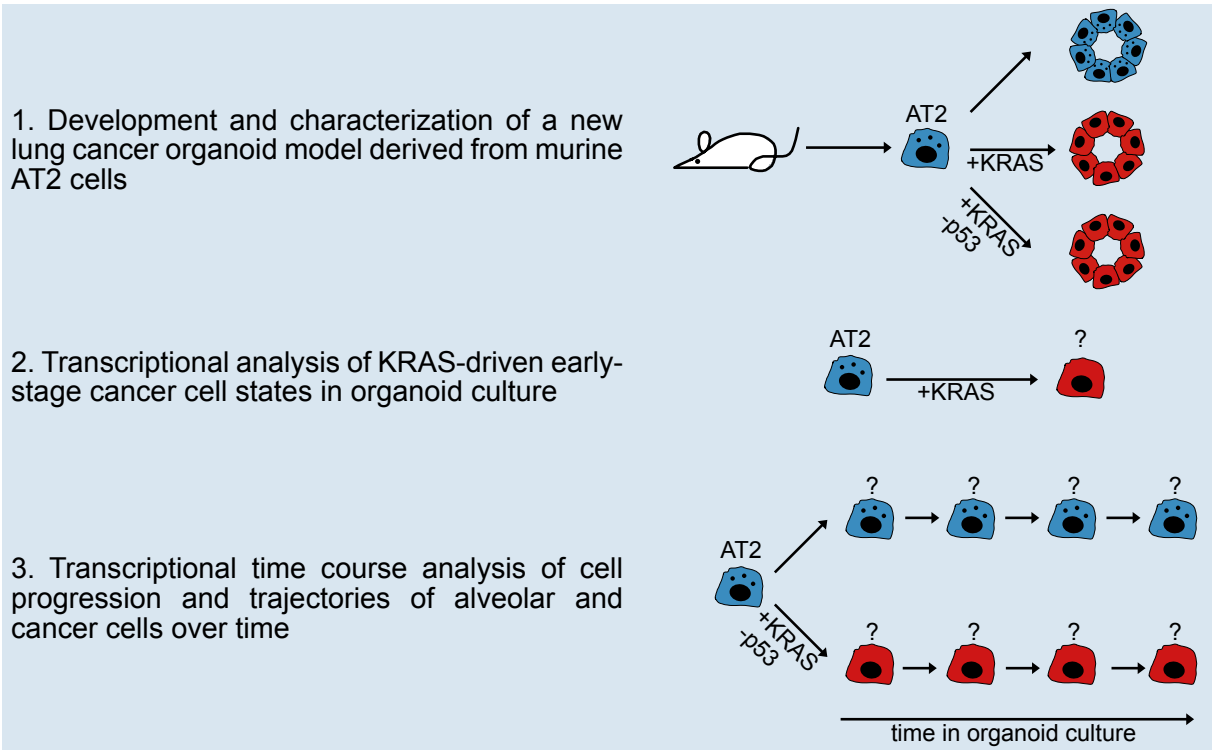

Figure 8: Schematic overview of objectives
Overview of the objectives of this thesis, divided into three parts. AT2=alveolar type 2 cell. Red cells indicate transformed cancer cells. Organoids made up of blue cells represent alveolar organoids, while red organoids represent cancer organoids. Questionmarks indicate unknown cell states that were explored in this thesis.

2 Materials

2.1 Mouse strains

Table 1

Mouse strain	Source
Gt(ROSA)26Sor$^{tm1(EYFP)Cos}$	The Jackson Laboratory; Cat#006148
Hsd:Athymic Nude-Foxn1^{nu}	ENVIGO; Cat#6903F
Kras$^{LSL-G12D/+}$	Jackson *et al.* (2001)
Kras$^{LSL-G12D/+}$; p53$^{fl/fl}$	Jackson *et al.* (2005)

2.2 Virus strains

Table 2

Virus strain	Company and Identifier
AdCMV-Cre	Viral Vector Core University of Iowa; Lot: Ad4117; Cat#VVC-U of Iowa-5
AdCMV-Empty	Viral Vector Core University of Iowa; Lot:Ad4154; Cat#VVC-U of Iowa-272

2.3 Antibodies and dyes

Table 3

Antibody or dye	Company and Identifier
4 ,6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich; Cat#D9542
Donkey anti-goat Alexa Fluor 488	Invitrogen; RRID:AB_2534102; Cat#A-11055
Donkey anti-goat Alexa Fluor 647	Invitrogen; RRID:AB_141844; Cat#A-21447
Donkey anti-mouse Alexa Fluor 647	Invitrogen; RRID:AB_162542; Cat#A-31571
Donkey anti-rabbit Alexa Fluor 488	Invitrogen; RRID:AB_141708; Cat#A-21206

Donkey anti-rabbit Alexa Fluor 594	Invitrogen; RRID:AB_141637; Cat#A-21207
Donkey anti-rat Alexa 594	Invitrogen; RRID:AB_2535795; Cat#A-21209
Goat polyclonal anti-GFP (YFP)	Abcam; RRID:AB_305643; Cat#ab6673
Mouse monoclonal anti-E-Cadherin	BD Biosciences; RRID:AB_397580; Cat#610181
Mouse monoclonal anti-Hmga2 [GT763]	GeneTex; Cat#GTX629478
Rabbit monoclonal anti-SP-C [EPR19839]	Abcam; Cat#ab211326
Rabbit monoclonal anti-TTF1 (Nkx2-1) [8G7G3/1]	Abcam; RRID:AB_1310784; Cat#ab76013
Rabbit monoclonal anti-Vimentin	Abcam; RRID:AB_10562134; Cat#ab92547
Rat monoclonal anti-CD31 APC [MEC 13.3, BD]	Thermo Fisher Scientific; Cat# BDB551262
Rat monoclonal anti-CD326 (EP-CAM) PE/Cy7 [G8.8]	BioLegend; RRID:AB_1236471; Cat#118216
Rat monoclonal anti-CD45 APC [30-F11, BD]	Thermo Fisher Scientific; Cat#BDB559864
Rat monoclonal anti-Ki67 [SolA15]	Thermo Fisher Scientific; RRID:AB_10854564; Cat#14-5698-82
Rat monoclonal anti-Ly-6A/E (Sca1) APC/Cy7 [D7]	Thermo Fisher Scientific; RRID:AB_1727552; Cat#560654

2.4 Chemicals, reagents, consumables, enzymes, media components

Table 4

Product	Company and Identifier
Basic fibroblast growth factor (bFGF)	R+D Systems; Cat#3139-FB/CF
Bleomycin Sulfate	Sigma-Aldrich; Cat#B2434
Bovine pituitary extract	Invitrogen; Cat#13028-014
Cell Strainer, 40 µm	Fisher Scientific; Cat#352340
Cell Strainer, 100 µm	Fisher Scientific; Cat#352360

Cholera Toxin	Sigma-Aldrich; Cat#C8052
Citrate-based antigen unmasking solution	Vector Laboratories; Cat#H-3300-250
Collagenase/Dispase	Roche; Cat#10269638001
Dispase	Corning; Cat#CB-40235
DMEM-Ham's F12	Thermo Fisher Scientific; Cat#11330032
DNAse	Sigma-Aldrich; Cat#D4527
EGF	Invitrogen; Cat#53003-018
Fetal Bovine Serum (FBS)	Corning; Cat#35-016-CV
Formalin, 10% neutralized buffer	Thermo Fisher Scientific; Cat#23245685
Growth factor reduced (GFR) Matrigel	Corning; Cat#356231
Hepes	Invitrogen; Cat#15630-080
HistoGel™	Thermo Fisher Scientific; Cat# HG-4000-012
Hoechst 33342	Thermo Fisher Scientific; Cat#62249
Insulin	Sigma Aldrich; Cat#I3146
Insulin/transferrin/selenium	VWR; Cat#45001-090
L-glutamine	Thermo Fisher Scientific; Cat#25030149
LysoTracker™-Red DND-99	Invitrogen; Cat#L7528
Penicillin-streptomycin	Invitrogen; Cat#15140122
Transwell, 24-well plate format	Corning; Cat#3470
Transwell, 96-well plate format	Corning; Cat#7369
Tribromoethanol (Avertin)	Sigma-Aldrich; Cat#T48402
Trypsin EDTA, 0.25%	Invitrogen; Cat#25200072

2.5 Small molecules

Table 5

Small molecule	Company and Identifier
ALW-II-41-27	MedChemExpress; Cat#HY-18007
GSTO1-IN-1	MedChemExpress; Cat#HY-111530
Talniflumate	Sigma-Aldrich; Cat#SML1710

2.6 Critical Commercial Assays

Table 6

Critical commercial assay	Company and Identifier
Apsolutely RNA Microprep kit	Agilent Technology; Cat#400805
Chromium™i7 Multiplex Kit, 96 rxns	10X Genomics; Cat#120262
Chromium™Single Cell 3' Library + Gel Bead Kit v2, 16 rxns	10X Genomics; Cat#120237
Chromium™Single Cell A Chip Kit, 48 rxns	10X Genomics; Cat#120236
KAPA Library Quantification Kit	Roche; Cat#07960140001
Qubit™dsDNA HS assay kit	Invitrogen; Cat#Q32851

2.7 Software and Algorithms

Table 7

Software or Algorithm	Source
CellRanger 3.0.0	10x Genomics; https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation
Enrichr in gseapy 0.9.13	Kuleshov *et al.* (2016); https://github.com/zq-fang/GSEApy/blob/master/docs/index.rst
FACSDiva™version v9.0	Becton and Dickinson
FlowJo version 10.5.3	Becton, Dickinson + Company; https://www.flowjo.com/
GraphPad Prism for MacOS version 8.2.1	GraphPad Software; https://www.graphpad.com/scientific-software/prism/
ImageJ	Schneider *et al.* (2012); https://imagej.nih.gov/ij/
Markov Affinity-based Graph Imputation of Cells (MAGIC) 1.5.5	van Dijk *et al.* (2018); https://github.com/KrishnaswamyLab/MAGIC
Matplotlib 3.0.2	Hunter (2007); https://matplotlib.org/index.html
1.4.4	Wolf *et al.* (2018); https://github.com/theislab/scanpy
scVelo 0.1.25	Theis lab; https://github.com/theislab/scvelo
Seaborn 0.9.0	https://seaborn.pydata.org/#

VDisseelocyto 0.17.16	Theis lab; https://github.com/theislab/scvelo
VIPER snakemake pipeline	Cornwell *et al.* (2018)

2.8 Equipment

Table 8

Equipment	**Company**
10X Chromium Controller	10X Genomics, Pleasanton, CA, USA
FACSAria II	Becton Dickinson, Franklin Lakes, NJ, USA
Evos™ FL Auto 2 imaging system	Thermo Fisher Scientific, Waltham, MA, USA
Illumina NextSeq500	Illumina, San Diego, CA, USA
LSR II Flow Cytometer	Becton Dickinson, Franklin Lakes, NJ, USA
Nikon eclipse 90*i* digital microscope	Nikon, Tokyo, Japan
Tapestation 2200	Agilent, Santa Clara, CA, USA
Qubit Fluorometric Quantification	Thermo Fisher Scientific, Waltham, MA, USA

3 Methods

3.1 Mouse work

3.1.1 Mouse husbandry

Kras$^{LSL-G12D/+}$ (K) and Kras$^{LSL-G12D/+}$; Trp53$^{fl/fl}$ (KP) mice were crossed to Rosa26 LSL-YFP mice to obtain Kras$^{LSL-G12D/+}$; Rosa26 LSL-YFP (KY) and Kras$^{LSL-G12D/+}$; Trp53$^{fl/fl}$; Rosa26 LSL-YFP (KPY) strains. Rosa26 LSL-YFP (Y) littermates of KY mice were used as a control. Mice were maintained in virus-free conditions. All mouse experiments were approved by the Boston Children's Hospital Animal Care and Use Committee, accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International, and were performed in accordance with relevant institutional and national guidelines and regulations.

3.1.2 Preparation of murine lung tissue

To dissect lung tissue, mice were anesthetized with tribromoethanol (Avertin), the rib cage was opened, and the lungs were perfused by administering 10 ml phosphate-buffered saline (PBS) into the right heart ventricle. Lungs were inflated by injecting 2 ml dispase (Corning) through the trachea. Then, inflated lungs were dissected out of the mouse. Working on ice, single lobes were separated from the trachea, bronchi, and remaining connective tissue. The lung lobes were minced using scissors, and incubated in 3 ml PBS per lung containing 25 U/ml DNAse (Sigma Aldrich) and 2 mg/ml collagenase/dispase (Roche) for 45 min at 37 °C while rotating. After the incubation, the tissue was filtered through 100 µm and 40 µm cell strainers (Fisher Scientific), and the flow-through was centrifuged for 5 min, at 1000 rpm and 4 °C. The supernatant was discarded and the pelleted cells were resuspended in red blood cell lysis buffer (0.15 M NH_4Cl, 10 mM $KHCO_3$, 0.1 mM EDTA) for 105 sec. The reaction was quenched with DMEM-Ham's F-12 (Gibco), cells were spun down 5 min at 1000 rpm and 4 °C, and finally resuspended in PBS containing 10% FBS (PF10) at 1 million cells per 100 µl.

3.1.3 Obtaining primary lung AT2 cells with fluorescence-activated cell sorting

After obtaining single cell suspensions from murine lungs as described in section 3.1.2, AT2 cells were isolated using fluorescence-activated cell sorting (FACS). For staining, cells were in-

cubated on ice at a concentration of 10 Mio cells per ml for a duration of 10 min with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich) as a viability dye and a combination of the following antibodies: anti-CD31-APC, anti-CD45-APC, anti-Ly6A/E-APC/Cy7 (stem cell antigen 1; SCA1) (all Thermo Fisher Scientific), and anti-CD326-PE/Cy7 (EPCAM; Biolegend) (1:100 dilutions for all antibodies). Single antibody staining controls and fluorophore minus one (FMO) controls were included for each experiment. FACS was performed on a FACSAria II with BD FACSDiva™ software, and subsequent analysis was done with FlowJo software.

3.1.4 Intratracheal transplantation of organoid-derived cells

For organoid transplantation studies, 8-10 weeks old athymic nude mice were used as recipients. To facilitate cell engraftment, the lungs of the mice were injected with 1.5 U/kg bleomycin sulfate (Sigma-Aldrich) intratracheally one day before transplantation. For transplantation assays, single cell suspensions were obtained from 14 - 21 days old organoid cultures as described in section 3.2.3. To ensure administration of equal cell numbers across samples, YFP+ cells were counted under a fluorescence microscope and 33,000 - 130,000 YFP+ cells in a volume of 45 µl PBS were injected into the lungs of the recipient mice through the trachea, as described elsewhere (DuPage *et al.*, 2009). For Y-CRE, KY-CRE, KPY-CRE derived cells the numbers of mice transplanted were 4, 6, and 4, respectively. After four weeks, mice were sacrificed, and lungs were fixed by administering 10% neutral-buffered formalin (Thermo Fisher Scientific) intratracheally. Fixed lungs were then embedded in paraffin and stained with hematoxylin and eosin (H+E). Histological evaluation was performed by the histopathologist Dr. Roderick Bronson at the Rodent Histopathology core at Harvard Medical School.

3.1.5 *In vivo* induction and preparation of tumors in the KY mouse model

The GEMM work presented in this thesis was done by Dr. Aaron Moye and detailed methods can be found in Dost *et al.* (2020). Briefly, to induce tumors, eight-week-old KY mice were infected with 2.5×10^7 particle forming units (PFU) AdCMV-Cre virus by intratracheal instillation as described elsewhere (DuPage *et al.*, 2009). A 1:1 ratio of male and female mice was used. After seven weeks, lungs were dissected as described in section 3.1.2 and tissue was prepared for FACS as described in section 3.1.3. DAPI-/CD31-/CD45-/EPCAM+/YFP-

control cells and DAPI-/CD31-/CD45-/EPCAM+/YFP+ KRAS G12D expressing cells were sorted and processed for scRNA-Seq.

3.2 Organoid work

3.2.1 *In vitro* induction of AT2 cells with subsequent organoid culture

FACS isolated murine lung CD31-/CD45-/EPCAM+/SCA1- cells (see sections 3.1.2 and 3.1.3) were, dependent on the experiment, split into equal aliquots, or not split, and spun down for 12 sec at 12,000g. The cell pellet was resuspended in 100 µl MTEC+ media (table 9) containing 6×10^7 PFU/ml of AdCMV-Cre or Ad5CMV-Empty adenovirus, or no virus in 100 µl per 100,000 cells. It followed an incubation step of 1 h at 37 °C and 5% CO_2 in tubes. After incubation, cells were pelleted for 12 sec at 12,000g and resuspended in PBS. This step was repeated twice more for a total of three washing steps. Live cells were counted using tryptan blue dye to stain dead cells, pelleted for 12 sec at 12,000g, and resuspended in 3D media (table 9) at a concentration of 5,000 live cells per 50 µl. As supporting cells, neonatal mesenchymal cells were isolated from mice as described elsewhere, and cultured in 2D conditions for 5-6 passages (Lee *et al.*, 2014). At the day of the organoid culture set-up, the supporting mesenchymal cells were trypsinized from the plate, spun down for 5 min at 1,000 rpm and resuspended in growth factor reduced (GFR) Matrigel at a concentration of 50,000 cells per 50 µl. The cells resuspended in 3D media and the supporting mesenchymal cells in GFR Matrigel were mixed in equal volumes and 100 µl of the cell mix was pipetted onto the membrane of a transwell (24-well-plate format, Corning). To let the Matrigel solidify, transwell plates were incubated for 20 min at 37 °C and 5% CO_2. Lastly, 500 µl of 3D media was added to the bottom of the transwell-containing well, and the media was changed every two to three days.

Table 9

MTEC+ media	3D media
DMEM-Ham's F-12	DMEM-Ham's F-12
10% FBS	10% FBS
4 mM L-glutamine	4 mM L-glutamine
1x penicillin/streptomycin	1x penicillin/streptomycin
5 µg/ml ITS	5 µg/mL ITS
10 µg/ml insulin	1 µM HEPES
0.1 µg/ml cholera toxin	
25 ng/ml EGF	
25 ng/ml bFGF	
30 µg/ml bovine pituitary extract	

3.2.2 Staining of tissue slides from organoid cultures

To fix organoid cultures, transwells were submerged in 10% neutral-buffered formalin overnight at room temperature. The next day, transwells were rinsed with PBS and the Matrigel plug was removed from the transwell and immobilized by embedding it first in HistoGel™ (Thermo Fisher Scientific), then in paraffin. Paraffin blocks were cut into sections (5 µm) and adhered to glass slides. To prepare the slides for staining, they were deparaffinized using xylene, and rehydrated using 100%, 95%, and 70% ethanol successively. Slides were then stained with H+E, or further processed for immunofluorescence (IF) staining. Antigen retrieval was achieved by incubating slides in citrate-based antigen unmasking solution (Vector Laboratories) for 20 min at 95 °C. It followed three washing and permeabilization steps with PBS containing 0.2% Triton-X (PBS-T). For blocking, slides were incubated for 1 h at room temperature using 10% normal donkey serum. Slides were then incubated in a humidified chamber at 4 °C overnight with a combination of antibodies for KI67 (EBioscience, 1:100), YFP (Abcam, 1:400), SPC (Abcam, 1:1,000), NKX2-1 (Abcam, 1:250), HMGA2 (GeneTex, 1:200), Vimentin (Abcam, 1:250), and E-Cadherin (BD Biosciences, 1:100). After three more washing steps with PBS-T, secondary antibodies were added and included donkey anti-rat Alexa 594, donkey anti-goat Alexa 488/647, donkey anti-rabbit Alexa 488/594, donkey anti-mouse Alexa 647 (all Invitrogen, 1:200). After three more washing steps with PBS-T, coverslips were mounted onto slides using Prolong Gold with DAPI (Invitrogen).

3.2.3 Preparation of single cell suspensions and FACS staining

To dissolve the Matrigel, 100 µl dispase (Fisher Scientific) was added to the transwells and incubated for 1 h at 37 °C and 5% CO_2. The organoids were washed out of the transwell with PBS and transferred into 15 ml conical tubes. Then, organoids were pelleted at 300 g for 5 min, and resuspended in 37 °C preheated trypsin EDTA (0.25%, Invitrogen). Organoids were incubated in trypsin for 7-10 min at room temperature with occasional pipetting up and down, to obtain a single cell suspension. Lastly, trypsin was quenched by adding PF10. To stain the cells for subsequent FACS, the suspensions were incubated with 1 µg/ml DAPI and EPCAM-PE/Cy7 (Biolegend) at a dilution of 1:100 for 10 min on ice. A single stain DAPI sample served as the FMO control for PE/Cy7. After the incubation period, the cells were spun down for 12 sec at 12,000 g, resuspended in PF10, and filtered into FACS tubes. FACS was performed using BD FACSDiva™ software on a FACSAria II. Subsequent analysis was done using FlowJo.

3.2.4 LysoTracker™ and Hoechst staining in organoid culture and FACS

LysoTracker™-Red and Hoechst 33342 were added to the media of 14 days old organoid cultures in 50 nM and 10 µM concentrations, respectively. After an incubation time of 60 min, wells were imaged using the Evos™ FL Auto 2 Imaging System. For LysoTracker™-Red staining in FACS samples, 14 days old organoids were made into single cell suspensions as described in 3.2.3. Cells were incubated with 50 nM of LysoTracker™-Red at a concentration of 250,000 cells per ml for 30 min at 37 °C. For the last 10 min of incubation time, 1 µg/ml DAPI was added. After, cells were spun down, resuspended in PF10, and analyzed on a LSR 2 Flow Cytometer. Cells were gated on single, DAPI-, YFP+ cells.

3.2.5 Treatment of organoid cultures with small molecules

Organoids were obtained as described in section 3.2.1. However, a 96-well-plate Transwell (Corning) format was used, changing the cell numbers and volumes as indicated (table 10). Right after plating of organoids, small molecules were added to the media at concentrations indicated in the respective figures. All small molecules were dissolved in dimethyl sulfoxide (DMSO). The final DMSO concentration for all conditions, including the DMSO control, was 0.1%. Media and small molecules were renewed every other day. Number of replicates is

indicated in figure legends. For biological replicates, experiments were done on different days, using different mice. For technical replicates, multiple wells were plated for each condition. Figures show representative images. Wells were imaged using the Evos™ FL Auto 2 Imaging System.

Table 10

Format	Transwell surface area	number of epithelial cells	number of MECCs	total volume in transwell	media volume in well
24-well plate	0.33 cm^2	5,000	50,000	100 µl	500 µl
96-well plate	0.14 cm^2	2,000	20,000	40 µl	200 µl

3.3 RNA Sequencing: Preparation and computational analysis

3.3.1 Organoid preparation for bulk RNA-Seq

To prepare samples for the RNA-Seq data, cells of three mice were pooled for each genotype. Cells were subsequently infected with virus and plated in organoid culture as described in section 3.2.1. For each of the 3-4 biological replicates, at least 10 wells of organoids were pooled and EPCAM+ cells were obtained from organoid cultures as described in section 3.2.3. RNA was extracted using the reagents and instructions of the Absolutely RNA Microprep Kit (Agilent). After RNA extraction, all downstream quality control steps, library preparation, sequencing, and differential gene expression analysis was performed by the Molecular Biology Core Facilities at Dana-Farber Cancer Institute (DFCI) in Boston, MA, USA. Briefly, from 2 ng of RNA starting material, complementary DNA (cDNA) was synthesized using Clontech SmartSeq v4 reagents. The full length cDNA was then fragmented to a mean size of 150 base pairs (bp) using a Covaris M220 ultrasonicator. From 2 ng sheared cDNA, Illumina libraries were prepared using Takara Thruplex DNAseq reagents according to the protocol of the manufacturer. To quantify the finished double strand DNA libraries, Qubit™ fluorometer, Agilent TapeStation 2200, and quantitative reverse transcription PCR (RT-qPCR) using the Kapa Biosystems library quantification kit were used. Uniquely indexed libraries were combined in equimolar ratios and finally sequenced at the DFCI Molecular Biology Core Facilities, using an Illumina NextSeq500 run with single-end 75 bp reads .

3.3.2 Computational analysis of bulk RNA-Seq

The computational analysis of the RNA-Seq data was performed by the Molecular Biology Core Facilities at DFCI in Boston, MA, USA. Briefly, sequenced reads were aligned to the UCSC hg19 reference genome assembly, and gene counts were quantified using STAR (v2.5.1b) (Dobin *et al.*, 2013). Differential expression (DE) analysis was performed with DESeq2 (v1.10.1) (Love *et al.*, 2014) and normalized read counts (Fragments Per Kilobase of transcript per Million mapped reads) were calculated using cufflinks (v2.2.1) (Trapnell *et al.*, 2010). Finally, downstream analysis was performed using the VIPER snakemake pipeline (Cornwell *et al.*, 2018).

3.3.3 Organoid preparation for scRNA-Seq

To prepare samples for the KY scRNA data, and the Y and KPY time course data, cells of three mice were pooled for each genotype. Cells were subsequently infected with virus and plated in organoid culture as described in section 3.2.1. At the indicated time points, 5 - 35 organoid wells were pooled for each genotype, and single cells were obtained as described in section 3.2.3. For the KY day 7 data, EPCAM+ cells were used. For the Y and KPY timecourse data, YFP+ and YFP-/EPCAM- cells were sorted. The cells were encapsulated using the 10X Genomics Chromium Controller Instrument and the Chromium Single Cell A Chip Kit. The following steps containing cell encapsulation, reverse transcription of RNA, amplification of cDNA, and library preparation were prformed using reagents from the Chromium Single Cell 3' Library and Gel Bead Kit v2, following the protocol of the manufacturer without modification (chromium single cell 3 reagent kits user guide v2 chemistry). Briefly, single cells were resuspended in PF10 at a concentration of 1,000 cells per µl for encapsulation. Total cDNA amount and quality following amplification and clean-up steps were determined using a Qubit™ dsDNA HS assay kit and the Agilent TapeStation High Sensitivity D5000 ScreenTape System. Prior to sequencing, the quality of the libraries was determined using Agilent TapeStation and qPCR, performed by the Biopolymers Facility at Harvard Medical School. Libraries were sequenced using an Illumina NextSeq500 using paired-end sequencing with single indexing (Read 1 = 26 cycles, Index (i7) = 8 cycles, and Read 2 = 98 cycles).

3.3.4 Computational analysis of scRNA-Seq

The computational analysis of the scRNA-Seq data for the GEMM and organoid experiments were performed by Dr. Aaron Moye (GEMM, KY, Y and KPY YFP+ epithelial data) and Dr. Preetida Bhetariya (Y and KPY YFP-/EPCAM- mesenchymal data). Sequencing reads were aligned to the mm10 reference genome and the count matrices were generated using CellRanger3.0.0 software (10X Genomics). Then, count matrices were read into the Python single cell analysis environment Scanpy (v 1.4.4) (Wolf *et al.*, 2018). For quality control purposes, cells with $> 10\%$ mitochondrial content were removed. High mitochondrial content generally correlated with low read counts. The data was normalized, logarithmized, and the significant number of principle components was determined using the in-built Scanpy functions. To de-noise the data, the data diffusion tool Markov Affinity-based Graph Imputation (v 1.5.5) was used with the following settings: Genes to return = all, k = 3 (KY, Y, KPY), k = 5 (EPCAM-/YFP-), t = 3 (KY), t = 5 (Y and KPY), t = auto (EPCAM-/YFP-) (van Dijk *et al.*, 2018). Gene Ontology enrichment analysis was performed using the Enrichr tool and the GSEAPY (v 0.9.13) python wrapper (Kuleshov *et al.*, 2016). In order to filter results by TFs, a list from the Animal Transcription Factor Database was used (Hu *et al.*, 2019). Gene signatures were taken from various sources, including the TRRUST database, the Panglao database, and publications (Han *et al.*, 2018; Franzén *et al.*, 2019; Barbie *et al.*, 2009; Bild *et al.*, 2006). Data was visualized using the in-built Scanpy plotting functions Seaborn (v0.9.0) and Matplotlib (v 3.0.2) (Hunter, 2007).

To perform trajectory analysis with RNA-Velocity, Velocyto (0.17.16) was run on the CellRanger output files, using the run10X shortcut and the mm10 genome annotation file provided with the CellRanger pipeline (La Manno *et al.*, 2018). The resulting Loom files were then concatenated into an anndata object. In order to visualize velocity on the original Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) embedding, the velocity and original anndata objects were merged using the utils.merge() function in scVelo (0.1.25). Using the merged anndata object, velocity was calculated using in-built velocity functions.

3.4 Obtainment and analysis of human patient data

The human patient data presented in this thesis was obtained and analyzed by the Dr. Jane Yanagawa and Dr. Steven Dubinett groups at University of California Los Angeles (UCLA); scRNA-Seq analysis was performed by Dr. Linh Tran. Detailed methods can be found in Dost *et al.* (2020). Samples of two patients with the diagnosis stage IA LUAD were obtained and analyzed. One patient was 74 years old, female, with a KRAS G12F mutation identified as driver mutation. The second patient was 77 years old, female, with a KRAS G12V mutation. The patients provided written informed consent and the studies were approved by the UCLA institutional review board. Briefly, resected tumor samples and matched healthy tissue was processed to single cell suspensions. For the scRNA-Seq preparation, the 10X Genomics platform was used. The bioinformatic analysis was performed using the Seurat pipeline (Stuart *et al.*, 2019).

3.5 Statistical analysis and definition of replicates

The methods section and figure legends contain the number of biological or technical replicates used for each experiment. For biological replicates of organoid experiments, different mice were used to obtain organoid cultures on different days. For technical replicates, multiple organoid wells were analyzed for each biological replicate. Statistical testing was performed using GraphPad Prism or Scipy 1.3.0 statistical functions (scipy.stats). The tests used to determine statistical significance are quoted in the appropriate figure legends. P values are indicated in the figures. P values < 0.05 were considered significant. Not significant (n.s.) p$\geq$0.05; *p < 0.05; **p < 0.005; ***p < 0.0005.

4 Results

4.1 Establishment and characterization of *in vitro* transformed lung cancer organoids

To facilitate the study of KRAS-driven lung cancer, I sought to develop an *in vitro* model to recapitulate LUAD progression in a controlled environment. In order to partially recreate *in vivo* conditions, I chose an organoid air liquid interphase (ALI) co-culturing system with supporting mesenchymal cells, previously developed for 3D lung stem cell cultures in our lab.

4.1.1 Lung cancer organoids can be generated by transforming AT2 cells *in vitro*

In order to generate organoids, I isolated AT2 cells from three different genetic backgrounds. For the generation of cancer organoids, I made use of the K; Rosa26 LSL-YFP (KY) and KP; Rosa26 LSL-YFP (KPY) GEMMs that contained a Cre-inducible yellow fluorescence protein (YFP) reporter (for more information on the GEMMs see section 1.2.3). For control alveolar organoids, I used Rosa26 LSL-YFP (Y) mice that were litter mates of the KY mice. Using adult mice older than eight weeks old, I dissected the lungs and isolated AT2 cells using an established FACS strategy. Briefly, lungs were digested and single cell suspensions were prepared. After a red blood cell lysis step, the remaining cells were stained and sorted using FACS. I excluded dead cells (DAPI), blood cells (CD45), and endothelial cells (CD31), then gated on epithelial cells using epithelial cell adhesion molecule (EPCAM), and enriched for AT2 cells which are mostly found in the stem cell antigen-1 (SCA1) negative population (henceforth, DAPI-/CD45-/CD31-/EPCAM+/SCA1- cells will be referred to as SCA1- or AT2 cells) (Kim *et al.*, 2005; Lee *et al.*, 2014). I then infected the SCA1- cells with AdCMV-Cre virus *in vitro* (henceforth, infected cells will be indicated by -CRE labeling). The expression of Cre-recombinase lead to expression of YFP in Y, KY and KPY samples, to the expression of KRAS G12D in KY and KPY samples, and to the homozygous loss of *Trp53* in the KPY sample. After the virus infection, I mixed the cells with supporting primary mesenchymal cultured cells (MECCs) in a 1:10 ratio in 50% GFR Matrigel, and plated the co-culture into a 3D ALI transwell system as described elsewhere (figure 9) (Lee *et al.*, 2014, 2017).

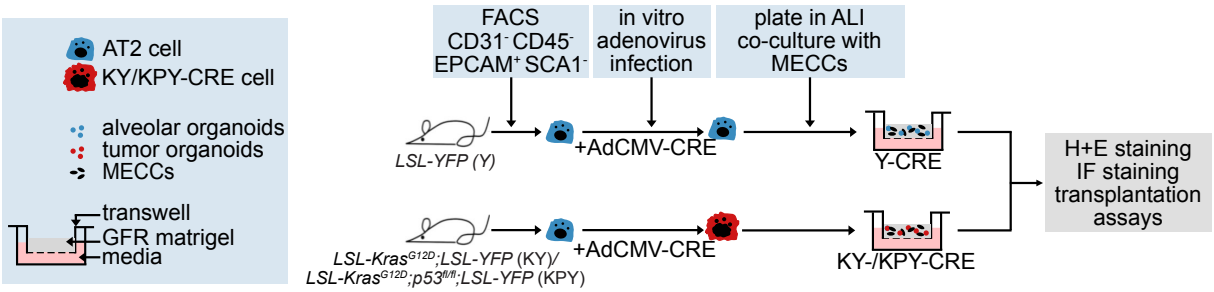

Figure 9: Workflow for the generation of AT2-derived organoids
Experimental strategy to isolate AT2 cells, enriched in the SCA1- population, from different mouse models using FACS. SCA1- cells were infected with AdCMV-Cre virus *in vitro*, and grown in air liquid interphase (ALI) organoid co-cultures with supporting mesenchymal cultured cells (MECCs) in growth factor reduced (GFR) Matrigel.

4.1.2 Oncogenic KRAS expressing organoid cells display hallmarks of cancer

In order to characterize alveolar and cancer organoids morphologically and histologically, I analyzed whole organoids and cells within organoids using microscopy. After 14 days of organoid culture, brightfield (BF) and fluorescence whole-well microscopy revealed that the majority of control and cancer organoids expressed the YFP color reporter, indicative of a high Cre-induction efficiency (figure 10 A). To determine organoid forming efficiencies (OFE), I quantified the number of YFP positive organoids and divided them by the epithelial cell number that was plated per well. Y-CRE and KY-CRE derived cells had comparable OFEs of 2.3% and 2.1%, respectively. In contrast, KPY-CRE derived cells had a higher OFE of 4.7% (figure 10 B). In order to assess if expression of KRAS G12D and loss of *Trp53* caused histological changes, I prepared sections of organoids on different time points and stained the slides with H+E (figure 10 C). Cells in the Y-CRE control organoids maintained normal nuclei in size and shape. In contrast, nuclei of KY-CRE and KPY-CRE organoids appeared enlarged and some nuclei were pleomorphic, a hallmark of cancer cells. Furthermore, giant, multinucleated cells were visible in the KPY-CRE organoids. The increased nuclear pleomorphism and the appearance of giant cells were reminiscent of the observed phenotypes in the K and KP mouse models, indicating that the organoids recapitulated *in vivo* tumor progression histologically (figure 5) (Jackson *et al.*, 2001, 2005).

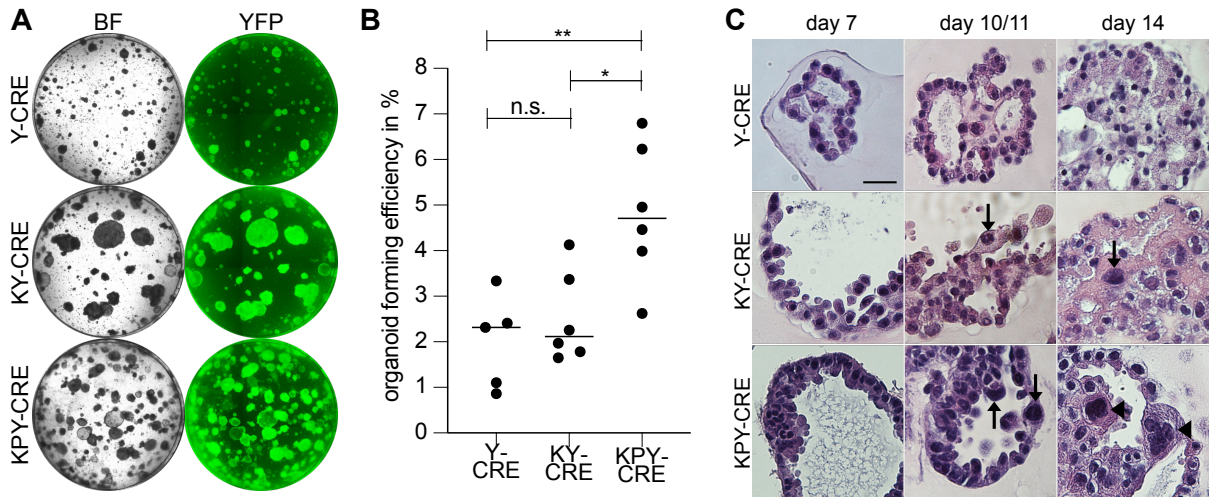

Figure 10: Morphological and histological analysis of control and cancer organoids
(A) Representative brightfield (BF) and YFP-channel images of whole wells on day 14 of organoid culture. **(B)** Organoid forming efficiency in %. One dot represents one biological replicate (Y-CRE: n=5, KY-CRE: n=6, KPY-CRE: n=6). The values of the biological replicates are the mean of 3-5 technical replicates. P values were calculated using an unpaired t-test. **(C)** Representative H+E stained organoid slides at indicated time points of organoid culture. Arrows: pleomorphic cells. Arrowheads: giant, multinucleated cells. Scale bar = 25 µm.

4.1.3 Oncogenic KRAS causes increased proliferation in cancer organoids

Because an increase in growth is usually observed in cancer cells, I next investigated if the KRAS G12D expressing cells were more proliferative. For this purpose, I stained day 7 and day 14 organoid sections for the nuclear proliferation marker KI67. Seven days after Cre-induction, there was no significant difference in the percentage of KI67+ cells per organoid between Y-CRE control (median 43%) and KY-CRE (median 35%) organoids. In contrast, KPY-CRE organoids contained a median of 56% KI67+ cells, a 1.3 fold and 1.6 fold increase compared to Y-CRE and KY-CRE, respectively (figure 11 A+B). 14 days after Cre-induction, Y-CRE organoids contained almost no KI67+ cells (median 0%), while both KY-CRE (median 14%) and KPY-CRE (median 27%) still contained proliferating cells (figure 11 C+D). Thus, organoids from all three genotypes had high numbers of proliferating KI67+ cells on day 7, but while control alveolar organoids had stagnated in growth by day 14, cells in organoids with KRAS G12D expression continued to proliferate.

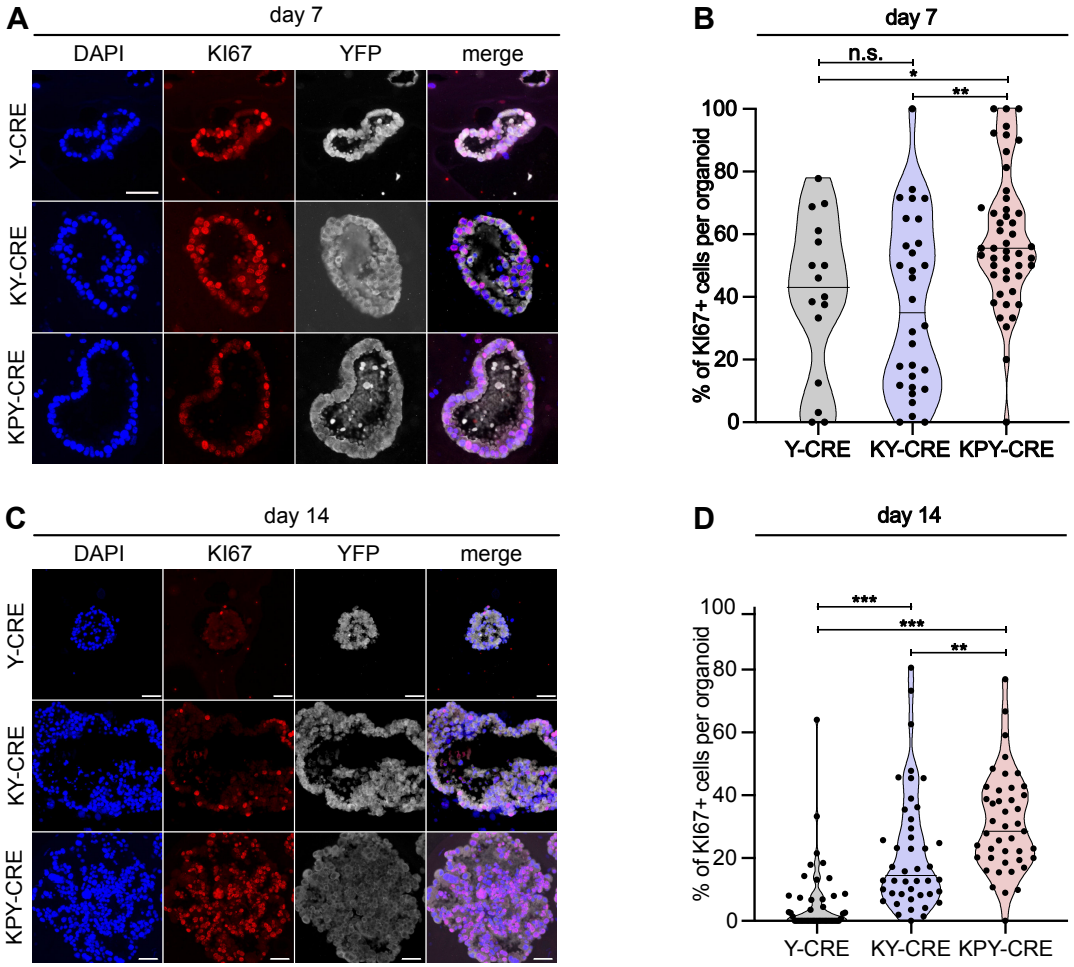

Figure 11: Proliferation marker staining and quantification in organoids
(A)+(C) Representative images of IF staining for DAPI (nuclei), KI67 (proliferating cells), YFP (Cre-induced cells), and merged channels on indicated days of organoid culture. Scale bar = 25 µm. **(B)+(D)** Quantifications of percentage of KI67+ cells per organoid field. One dot represents one organoid. Data is visualized using violin plots. P values were calculated using the Mann-Whitney rank test.

4.1.4 Cancer organoids initiate lung tumors in orthotopic transplantation assays

The KRAS G12D expressing organoid cells showed features of cancer cells, such are increased pleomorphisms and proliferation. To test if the cells were able to initiate tumors *in vivo*, I transplanted organoid-derived cells into lungs of recipient athymic nude mice. In order to enhance the engraftment of cells, I injected the chemical bleomycin intratracheally one day before transplantation, as often done in the lung field (Weiner *et al.*, 2019). The next day, I prepared single-cell suspensions from Y-CRE control, KY-CRE, and KPY-CRE organoids and administered the cells intratracheally into the lungs of the recipient mice. Four weeks after transplantation, I dissected the lungs and prepared tissue slides for histological evaluation by a histopathologist. In all three conditions, there were pockets of infiltrating lymphocytes typical for inflammation caused by bleomycin administration. However, there was no sign of tumor formation or aberrant epithelial cell growth in the lungs of the mice that received Y-CRE derived cells (figure 12 A). In contrast, mice that were transplanted with KY-CRE and KPY-CRE derived cells had tumors that were categorized as LUAD. In the KPY-CRE condition, the tumors contained pleomorphic cells and giant, multinucleated cancer cells comparable to the organoid cultures and the KPY GEMM *in vivo* phenotype (figure 12 A, 10 C, and 5). IF staining for YFP confirmed that the observed tumor lesions were comprised of the organoid-derived transplanted cells (figure 12 B). Overall, lungs with KY-CRE and KPY-CRE transplanted cells had tumor lesions that contained pleomorphic cells comparable to a grade 3 tumor, as defined for the KP model (Jackson *et al.*, 2005). Hence, I successfully transformed AT2 cells to cancer cells using *in vitro* induction and an organoid culture system. When transplanted orthotopically, cells derived from the cancer organoids formed tumors within four weeks, dramatically reducing the time required to model advanced stage lung cancer compared to traditional GEMMs.

4.2 Study of KRAS induced tumorigenesis with focus on early progression

The early steps of how KRAS transforms an epithelial cell into a cancer cell is poorly understood and difficult to study *in vivo*. Therefore, I decided to study the transcriptional landscape of early cancer progression *in vitro* using the new organoid model.

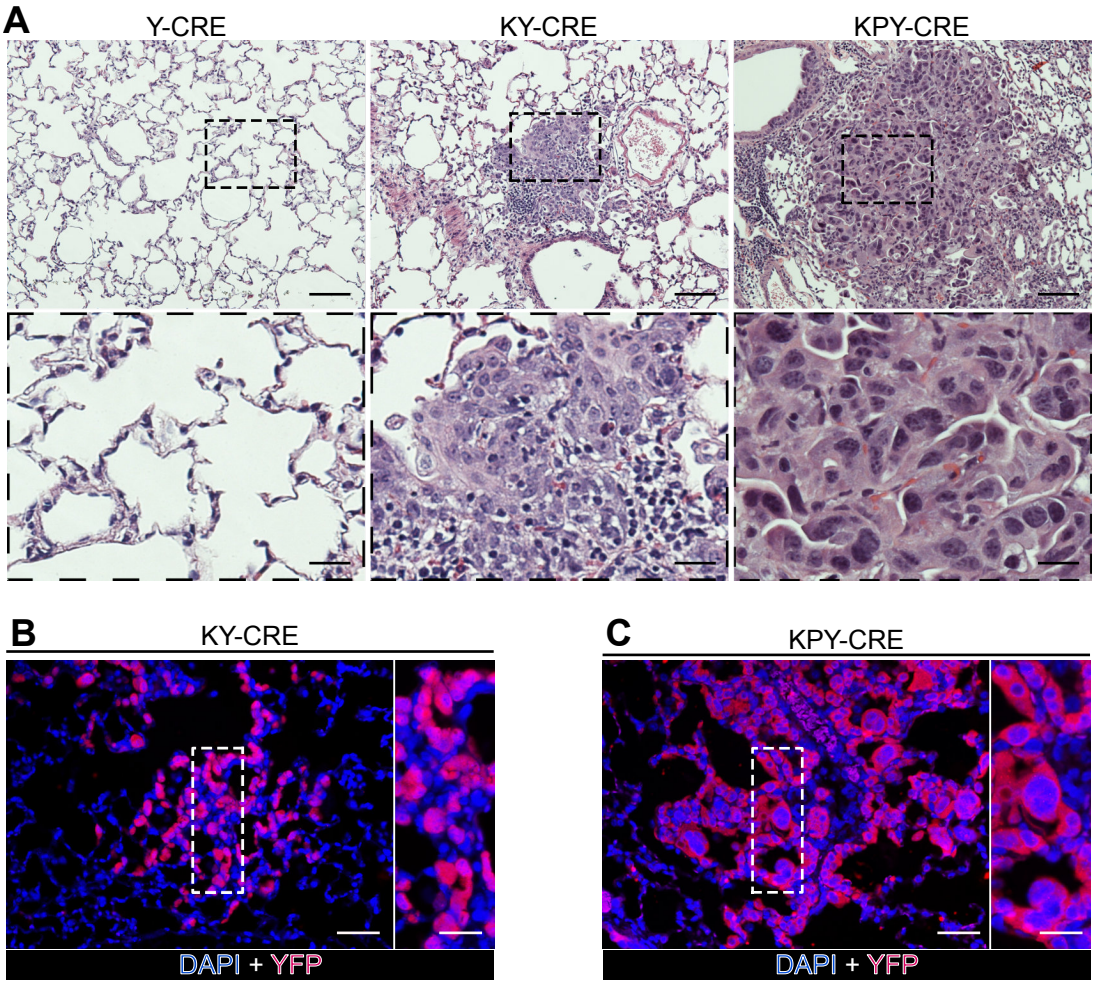

Figure 12: Orthotopic transplantation of organoid-derived cells
(A) Representative H+E images of recipient mouse lungs that were transplanted with organoid-derived cells. Number of transplanted mice: Y-CRE: n=4, KY-CRE: n=6, KPY-CRE: n=4. Scale bar low magnification = 100 µm; scale bar insets = 25 µm. **(B+C)** IF staining for DAPI (nulei) and YFP (transplanted cells) of lungs of recipient mice, confirming that the tumor lesions were comprised of organoid-derived cells. Scale bar low magnification = 50 µm; scale bar insets = 25 µm.

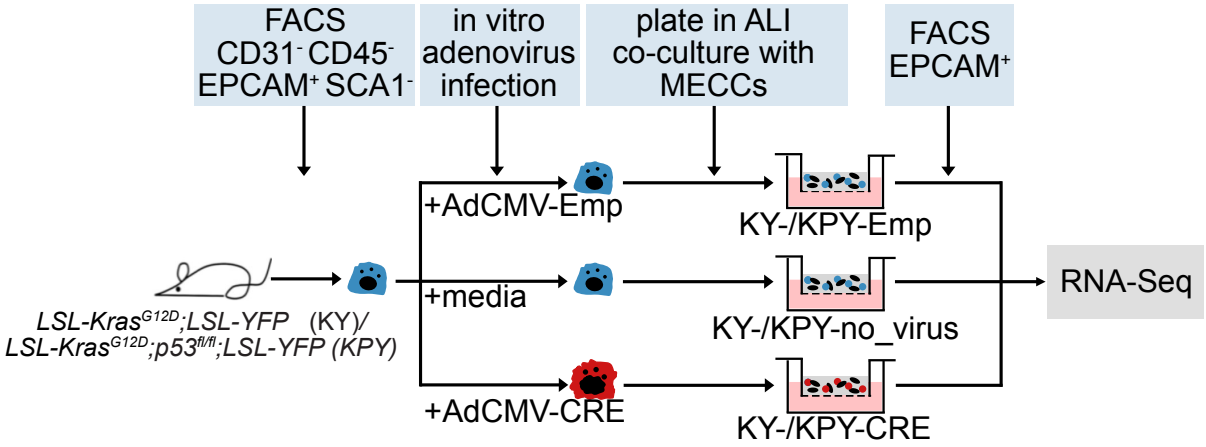

Figure 13: Workflow for the preparation of samples for RNA-Seq
Experimental strategy to prepare AdCMV-Emp, AdCMV-Cre, or no virus *in vitro* induced organoid cultures derived from SCA1- AT2 cells. After 7 days in culture, organoids were sorted for EPCAM+ cells and prepared for RNA-Seq. MECCs = mesenchymal cultured cells.

4.2.1 KRAS causes transcriptional downregulation of AT2 marker genes

In order to understand the early transcriptional changes that follow KRAS G12D expression, I performed bulk RNA-Seq on cells derived from the organoid cultures. There were two considerations in order to achieve comparable results between control and cancer organoids and to reduce background noise in the data. Firstly, I changed the controls to make them more comparable to the cancer samples. In detail, after FACS enrichment of KY- and KPY-derived AT2 cells, I split the cells into three arms that received either AdCMV-Cre (-CRE), AdCMV-Empty (-Emp, control), or no virus (no_virus, control). By using these controls instead of Y-CRE, the three arms were generated from the same pool of cells, reducing noise from differences in the genetic background. I then plated the infected cells in organoid culture as described before (figure 13). Secondly, I chose a time point at which proliferation was observed in all organoid conditions, 7 days after induction, mitigating the noise of transcriptional proliferation signatures (figure 11). At day 7 of culture, organoids were processed to single cell suspensions and enriched for epithelial cells by FACS sorting for EPCAM+ cells, removing supporting MECCs (figure 13). In the KY-CRE and KPY-CRE samples, 87% (+/- 7%) and 95% (+/- 2%) of the EPCAM+ cells, respectively, were YFP+, while the -Emp samples had 0% YFP+ cells, confirming high Cre induction efficiency (figure 14). After cell sorting, I extracted RNA from EPCAM+ cells and performed RNA-Seq on 3-4 biological replicates per condition.

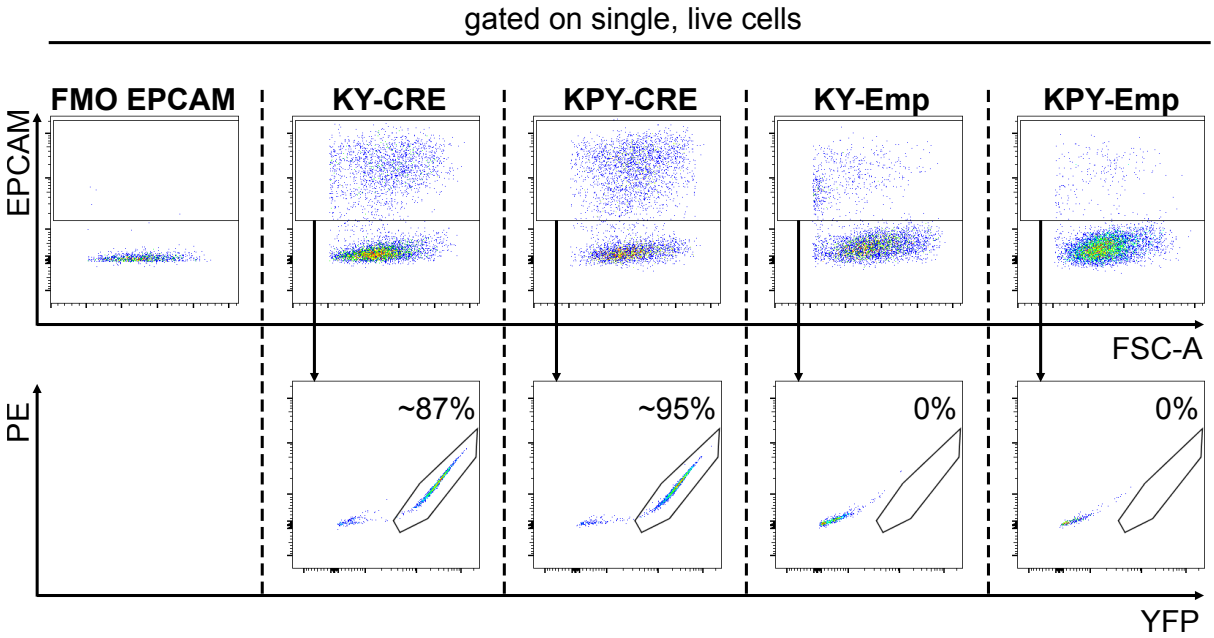

Figure 14: FACS strategy to enrich for epithelial cells for RNA-Seq
Representative FACS plots of organoid-derived cells. Cells were sorted for EPCAM+ cells before being processed for RNA-Seq. Percentage of YFP+ cells from EPCAM gate is shown to highlight the high Cre induction efficiency in the Cre samples. Fluorophore minus one (FMO) control is depicted to show specificity of the gate.

Sample-sample-correlation analysis revealed that all control samples clustered together, indicating high transcriptional similarity between the control samples regardless of origin (figure 15). In contrast, KY-CRE and KPY-CRE clustered closely together but were distinct from the control samples, indicating high transcriptional similarity between the -CRE samples but low correlation between the -CRE samples and the control samples (figure 15). Because there was very little transcriptional difference between the -Emp and no_virus control samples, downstream DE analysis was done comparing the -CRE samples to their respective -Emp control samples (KY-CRE compared to KY-Emp, henceforth: KY-Dif; KPY-CRE compared to KPY-Emp, henceforth: KPY-Dif). In order to find genes that were altered by the expression of oncogenic KRAS G12D, I looked for overlapping differentially expressed genes between KY-Dif and KPY-Dif and found that 1206 genes were shared upregulated and 1464 genes were shared downregulated (figure 16). To narrow down the list of genes to follow up on, I decided to focus on the overlap of the top 100 up- and downregulated genes for KY-Dif and KPY-Dif. Strikingly, among the shared downregulated genes were the well-known AT2 marker genes *Cd74* and *Lyz2*. Conversely, the lung developmental genes *Hmga2* and *Sox9* were amongst the shared upregulated genes, as was the lung progenitor marker *Ly6a* (SCA1) (figure 17 A+B;

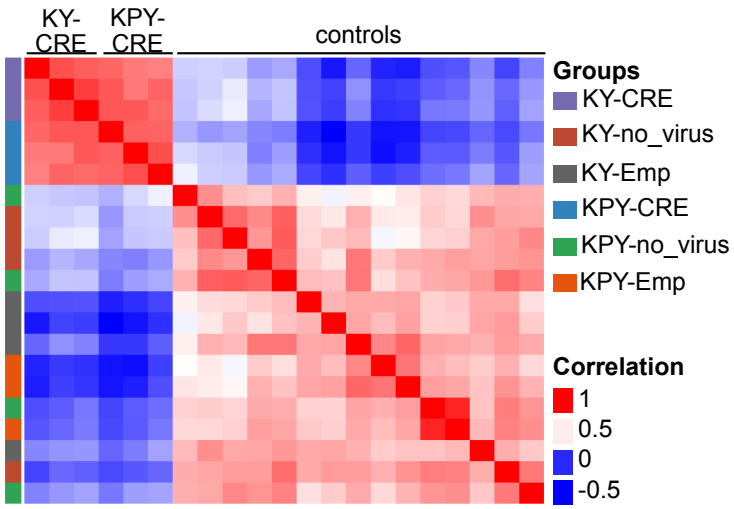

Figure 15: Correlation heatmap for RNA-Seq samples
Sample-sample-correlation heatmap of the organoid samples analyzed by bulk RNA-Seq. Each group had 3-4 biological replicates. All control samples clustered together and distinct from the CRE samples.

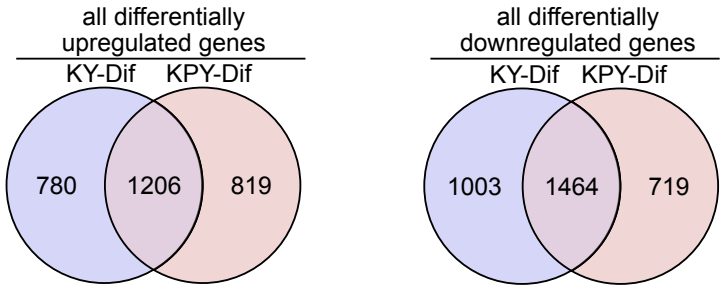

Figure 16: Overlap of differentially expressed genes in RNA-Seq analysis
Venn diagram depicting the number of genes uniquely and shared upregulated (left) and downregulated (right) in KY-CRE and KPY-CRE samples compared to their respective -Emp controls (KY-Dif and KPY-Dif).

table 14, appendix). Looking for the expression values of other well-known AT2 marker genes, I found that also *Sftpc* (SPC), and *Nkx2-1* were significantly downregulated both in KY-CRE and KPY-CRE compared to their respective controls (figure 17 B). These results indicated that upon KRAS G12D expression, AT2 cells lost the expression of genes associated with AT2 identity and acquired markers associated with the developing lung and progenitor cells.

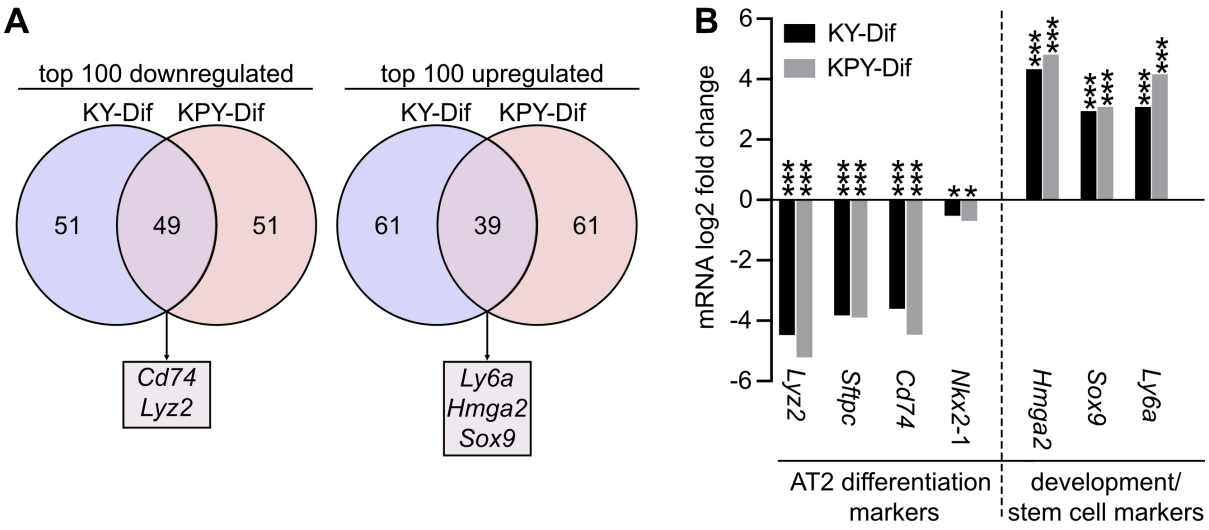

Figure 17: Differentially expressed genes in RNA-Seq data
(A) Number of top 100 genes uniquely and shared downregulated (left) and upregulated (right) in KY-CRE and KPY-CRE samples compared to their respective -Emp controls (KY-Dif and KPY-Dif), plotted as a Venn diagram. **(B)** Change in expression levels of selected genes as determined by bulk RNA-Seq analysis. P values were calculated as part of the RNA-Seq DESeq2 pipeline (Love *et al.*, 2014).

4.2.2 Expression of AT2 and developmental markers in cancer organoids is heterogeneous

Transcriptional analysis revealed that KRAS G12D expressing AT2 cells downregulated differentiation markers and upregulated development and progenitor cell markers. However, transcriptional changes do not always translate to changes on protein level. Therefore, I validated the findings using IF staining of organoid slides. On day 7 of organoid culture, the median percentage of SPC+ cells per organoid slide field was 100% in the Y-CRE control organoids, while the KY-CRE and KPY-CRE organoids had a 6.7-fold (median=15%) and 20-fold (median=5%) decrease, respectively (figure 18). On day 14, the same trend was observed, albeit more subtle, with a 1.1-fold decrease in KY-CRE and a 1.6-fold decrease in KPY-CRE compared to the Y-CRE control organoids (medians: Y-control=100%, KY-CRE=93%, KPY-CRE=61%) (figure 18). YFP staining was used to distinguish organoids from MECCs. SPC is the most commonly used marker for AT2 cells. However, SPC expression alone does not indicate if the AT2 cell is mature and functional. In order to test for mature AT2 cells, I used the live dye LysoTracker™ that stains acidic organelles such are lysosomes, which are frequent in AT2 cells (Van der Velden *et al.*, 2013). Cells in cancer organoids had lower intensity staining for LysoTracker™ in organoid culture (figure 19 A). In order to rule out that the

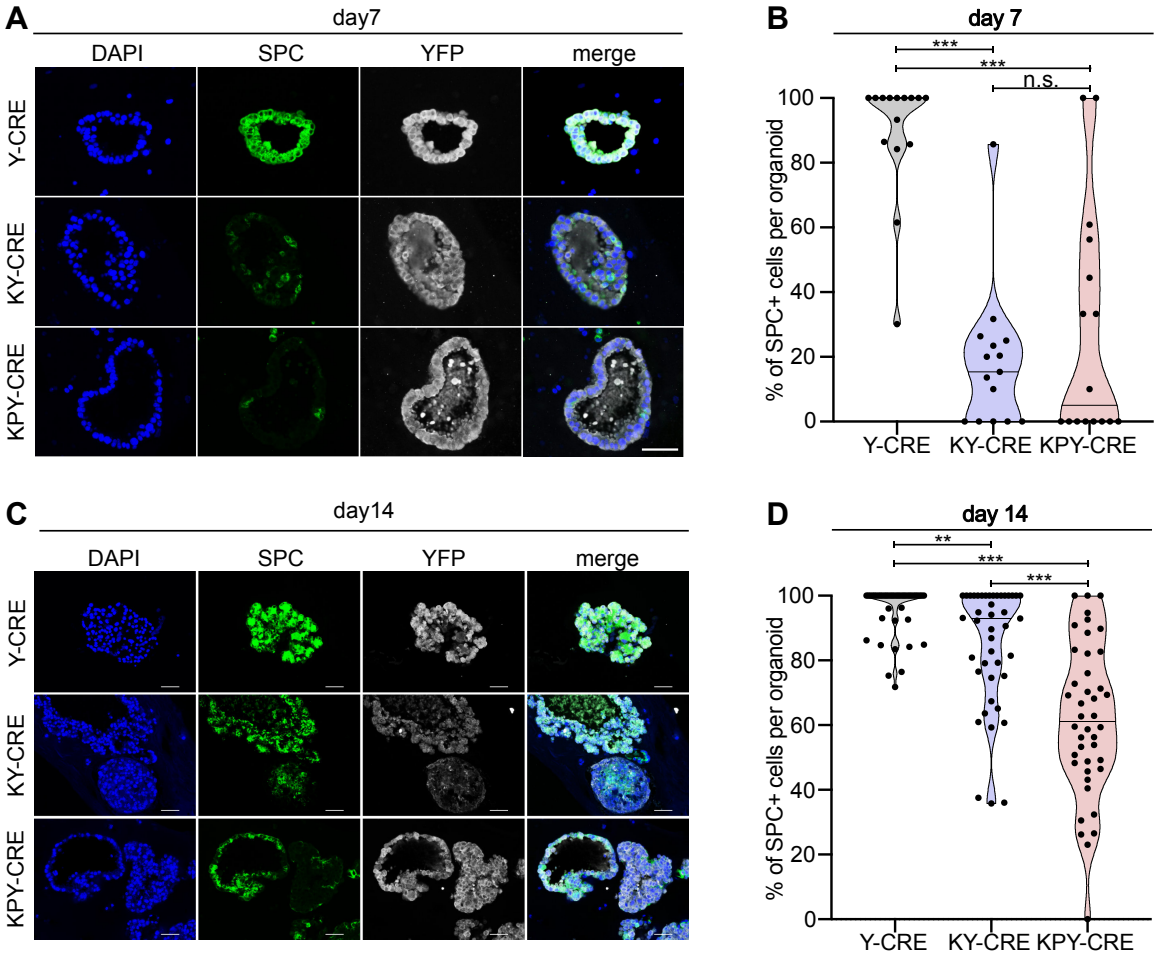

Figure 18: SPC stainings and quantifications
(A)+(C) Representative images of IF staining for DAPI (nuclei), SPC (AT2 marker), YFP (Cre-induced cells), and merged channels on indicated days of organoid culture. (A) Scale bar = 100 µm. (B) Scale bar = 50 µm. **(B)+(D)** Quantifications of percentage of SPC+ cells per organoid field. One dot represents one organoid. Data is visualized using violin plots. P values were calculated using the Mann-Whitney rank test.

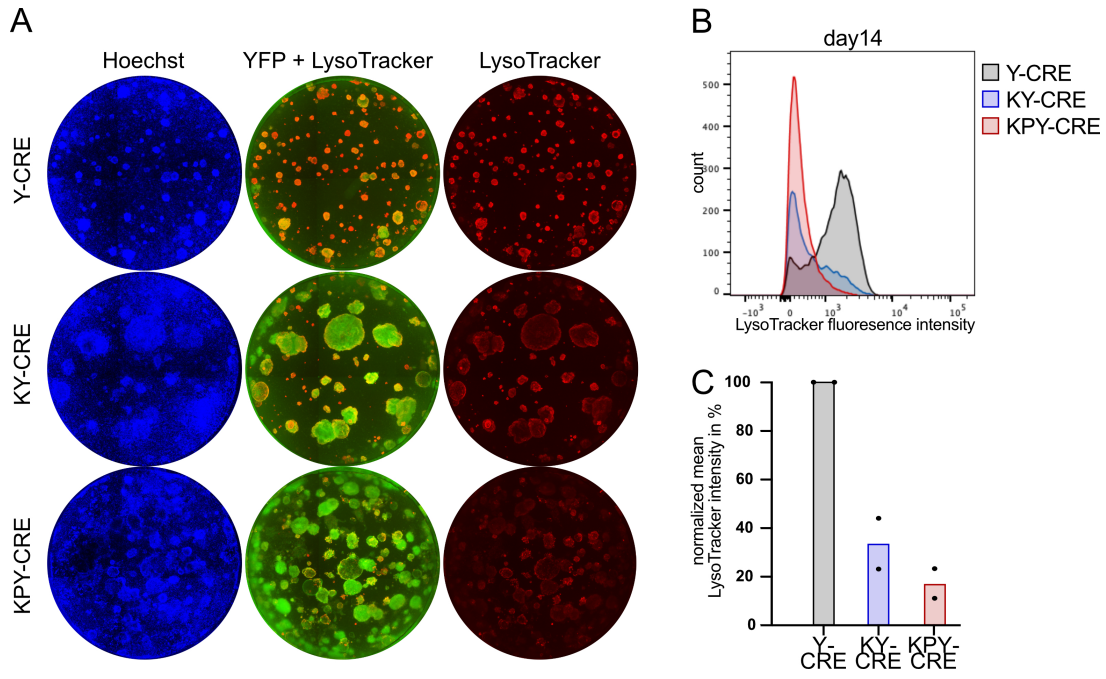

Figure 19: LysoTracker analysis of organoid cells
(A) Hoechst (nuclei), YFP (Cre reporter), and LysoTracke fluorescent images of whole, live organoid wells. **(B)** FACS histogram of LysoTracker fluorescence intensity. Cells were derived from organoids and processed to single cell suspensions before staining. Cells were gated on live, YFP+ cells. **(C)** Quantifications of LysoTracker mean fluorescence intensity as determined by FACS, gated on live, YFP+ cells. Biological replicates n=2.

lower staining intensity was due to differences in organoid size and number, I stained organoid-derived single cell suspensions and analyzed the fluorescence intensity using flow cytometry. There was a clear trend of reduced LysoTracker™ signal in the cancer organoid-derived cells compared to control cells, indicating that the cancer organoids contained fewer mature AT2 cells (figure 19 B+C).

Next, I stained for the AT2 and lung epithelial marker NKX2-1 and the development marker HMGA2, found to be transcriptionally downregulated and upregulated, respectively. I found that organoids stained heterogeneously for those markers. Some of the cancer organoids had high expression of NKX2-1 and low or no expression of HMGA2, while others had low NKX2-1 levels and high HMGA2 levels (figure 20). Staining of those proteins seemed to be negatively correlated within a cell, an observation that had been published previously (Winslow *et al.*, 2011). The stainings confirmed that AT2 cells lost differentiation markers and upregulated developmental markers as an early response to KRAS G12D activation; however, staining patterns were heterogeneous.

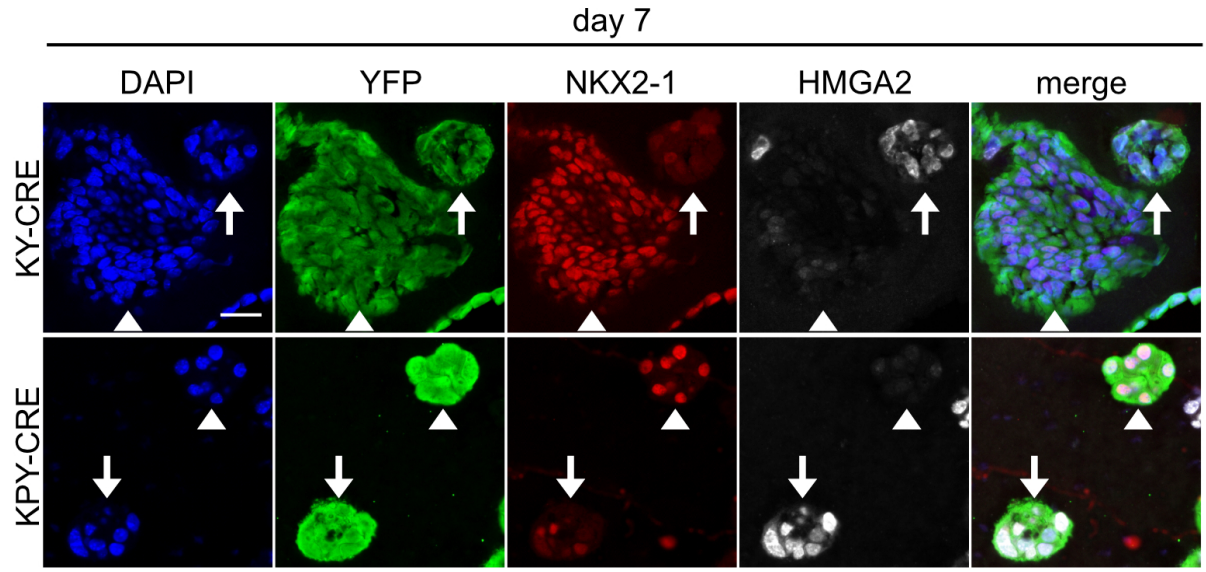

Figure 20: NKX2-1 and HMGA2 stainings in organoids
Representative images of immunofluorescence staining for DAPI (nuclei), YFP (Cre-induced cells), NKX2-1, HMGA2, and merged channels on indicated day of organoid culture. Scale bar = 25 µm.

4.2.3 Single cell transcriptional analysis of cancer organoids shows cell heterogeneity

While there was a general trend for dowregulation of AT2 differentiation genes and upregulation of developmental genes, staining revealed that the expression of those markers in cancer organoids was heterogeneous. To understand the transcriptional changes in early cancer progression better, I decided to characterize the transcriptional landscape of KY-CRE organoids with single cell resolution. As described before, I focused on day 7 EPCAM+ cells from KY-CRE and KY-Emp organoids (figure 21). ScRNA-Seq was performed using the 10xGenomics platform and the computational analysis was done using scanpy (Wolf *et al.*, 2018). After filtering of cells with low read count and high percentage of mitochondrial genes, the data was imputed using the AT2 markers *Sftpc* and *Lyz2* as an example for restored correlation (figure 22). The cells were then clustered using the UMAP algorithm. Clusters (C) from the KY dataset will be indicated by a C_{KY} nomenclature. With the chosen resolution, three clusters were identified, $C_{KY}0$, $C_{KY}1$, and $C_{KY}2$ (figure 23 A). $C_{KY}1$ mostly contained cells that were derived from the KY-Emp sample (this cluster will be referred to as control cluster), while $C_{KY}0$ and $C_{KY}2$ mostly contained KY-CRE cells (these clusters will be referred to as Cre-clusters) (figure 23 B+C). The small overlap between the control and the Cre clusters confirmed that KRAS G12D expressing AT2 cells had a distinct transcriptional profile from normal AT2s, an observation that was also made in the previous bulk RNA-Seq data (fig-

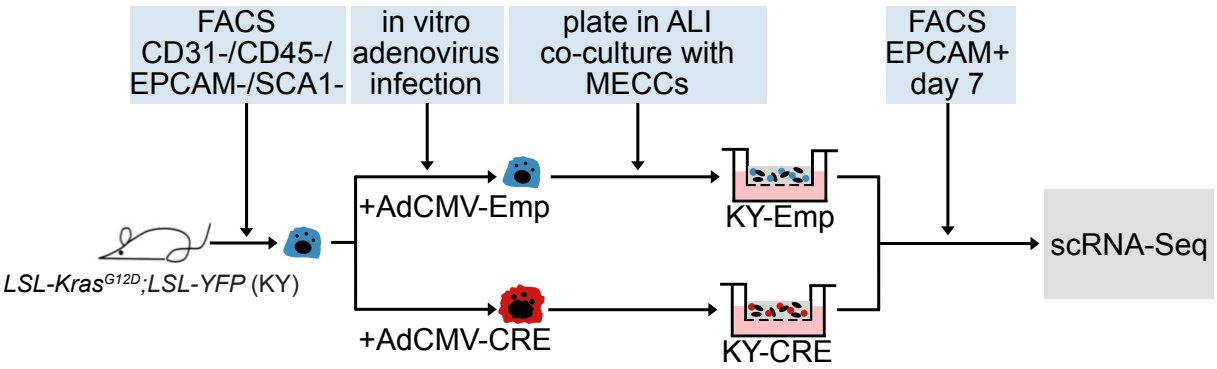

Figure 21: Workflow for preparation of samples for KY scRNA-Seq
Experimental strategy to prepare AdCMV-Emp and AdCMV-Cre *in vitro* induced organoid cultures derived from SCA1- AT2 cells from KY mice. After 7 days in culture, organoids were sorted for EPCMA+ cells and prepared for scRNA-Seq. MECCs = mesenchymal cultured cells.

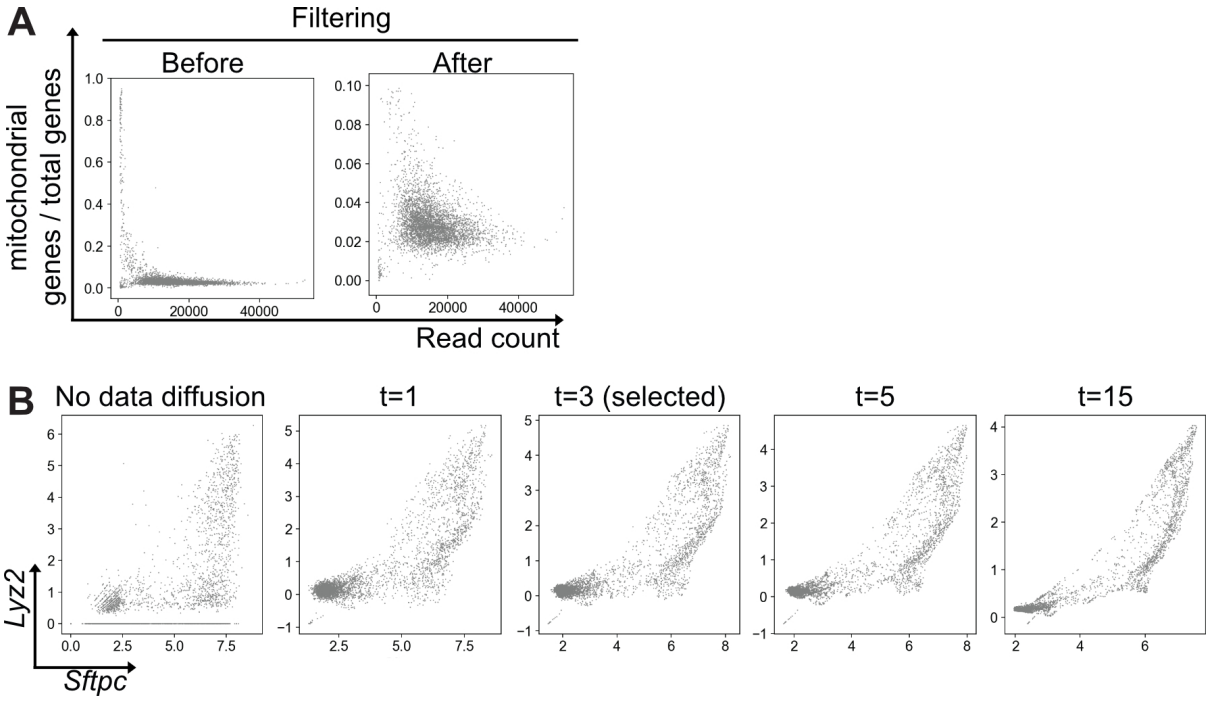

Figure 22: Filtering and imputation of KY scRNA-Seq data
(A) Correlation between read count and mitochondrial genes to total genes ratio depicted before and after filtering of cells with >10% mitochondrial genes. **(B)** Correlation between the two AT2 genes *Sftpc* and *Lyz2* after t rounds of data diffusion (van Dijk *et al.*, 2018). T=3 was selected because of the restored correlation between the two selected genes.

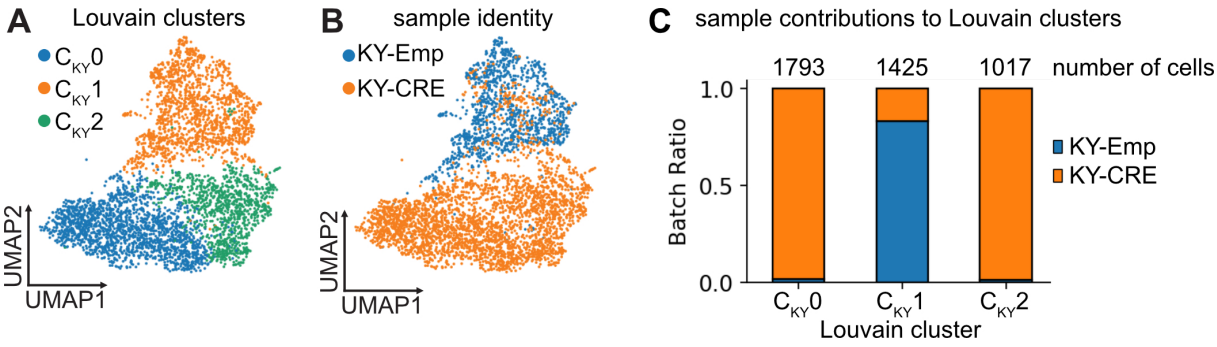

Figure 23: Clustering of KY organoids
(A) UMAP representation of scRNA-Seq data of KY organoids colored by Louvain clusters. Three clusters were identified. **(B)** Cells are colored by sample identity. KY-Emp represents control sample. **(C)** Bar graph representation of sample contributions to Louvain clusters and number of cells in each cluster.

ure 15). First, we checked if the cancer organoids had similar pathway activation to reported LUAD *in vivo* signaling. For this purpose, we used gene signatures from the literature known to be upregulated in LUAD (figure 24). Unsurprisingly, the two Cre clusters $C_{KY}0$ and $C_{KY}2$ had an increased Kras signature score, confirming high expression levels of Kras downstream targets (Bild *et al.*, 2006). A NF-κB signature was low in $C_{KY}2$ and high in $C_{KY}0$ compared to the control cluster $C_{KY}1$, indicating that only one of the Cre clusters had activated the NF-κB pathway (Han *et al.*, 2018). Strikingly, the proliferation signature showed the opposite trend and was only upregulated in $C_{KY}2$, despite a higher Kras activation score in $C_{KY}0$ (Travaglini *et al.*, 2020). This showed that activation of oncogenic KRAS had transcriptional effects on the cell state that went beyond activation of proliferation. Next, we checked the Hallmark_Wnt_beta_catenin_signaling signature, indicative of activated Wnt signaling, and found that $C_{KY}0$ had transcriptional activation of the Wnt pathway (Subramanian *et al.*, 2005; Liberzon *et al.*, 2015). Lastly, an AT2 signature was downregulated in both Cre clusters, confirming the bulk RNA-Seq results and the observations using IF staining (figure 24) (Franzén *et al.*, 2019).

In order to analyze the scRNA-Seq data unbiasedly, we performed DE analysis on the three clusters (table 15, appendix). To get a broader picture of the transcriptional changes within the clusters, differentially expressed genes were filtered for TFs and co-factors. Strikingly, control cluster $C_{KY}1$ had elevated expression of *Etv5*, a TF known to be important for maintenance of AT2 identity (figure 25) (Morrisey and Hogan, 2010; Zhang *et al.*, 2017b). This finding provided further evidence that while the control organoid cells maintained their AT2 identity, KRAS G12D expressing cells downregulated genes associated with AT2s. In contrast, both

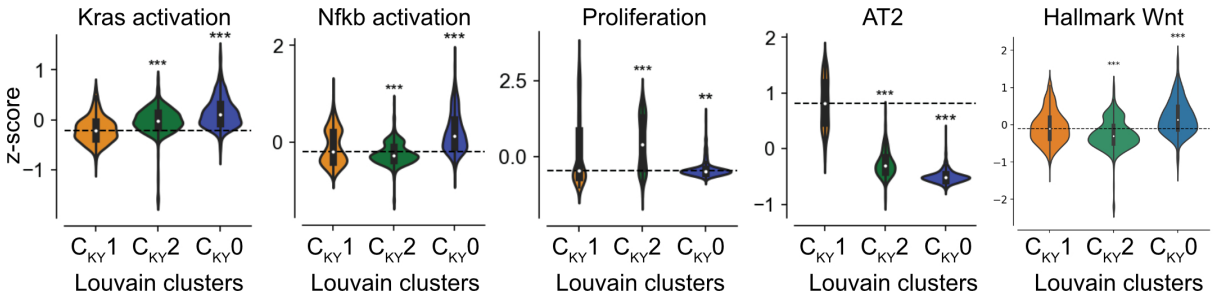

Figure 24: Expression of transcription signatures associated with LUAD
Z-scores of indicated previously published transcription signatures in each Louvain cluster. Dashed line marks median expression of control cluster $C_{KY}1$. Data is represented using violin plots. P values were calculated using the Mann-Whitney rank test.

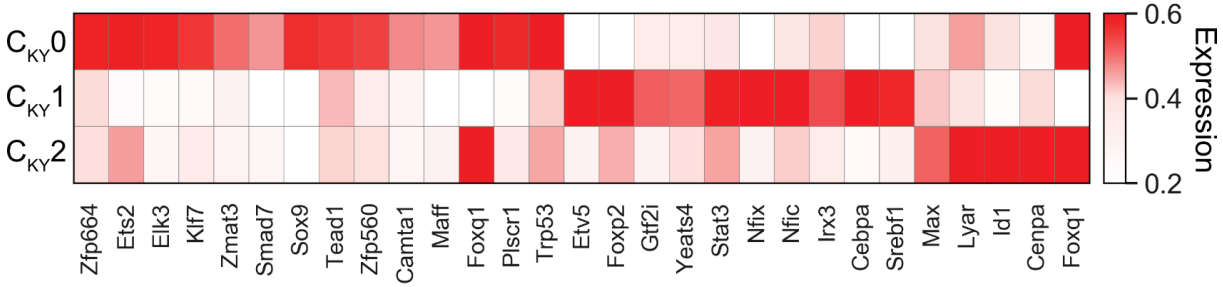

Figure 25: Differentially expressed transcription factors and co-factors
Heatmap depicting expression levels of differentially expressed TFs in the three Louvain clusters identified in the KY scRNA-Seq data. $C_{KY}1$ represents control cluster.

Cre clusters had high expression of *Foxq1*, a TF found to be increased in NSCLC tumor tissue compared to paired adjacent tissue (Li *et al.*, 2020). Cre cluster $C_{KY}2$ had increased expression of *Id1*, which was shown to promote NSCLC cell proliferation and metastasis, and correlated with worse outcome in patients (Antonângelo *et al.*, 2016; Cheng *et al.*, 2011; Pillai *et al.*, 2011). In contrast, $C_{KY}0$ had elevated levels of *Sox9*, an important lung development TF that was also identified in the bulk RNA-Seq analysis. Furthermore, *Trp53* and *Smad7* were upregulated in the same cluster, indicative of Tgfb and p53 signaling, respectively. In agreement with the decreased AT2 signature and the reduced expression of the AT2 maintenance TF *Etv5*, the AT2 markers *Lyz2*, *Sftpc*, and *Nkx2-1* were also downregulated in both Cre clusters (figure 26). In contrast, expression of the developmental gene *Hmga2* was elevated, confirming previous bulk RNA-Seq and staining results.

Gene Ontology (GO) classifies genes and annotates them to terms to facilitate the analysis of large gene lists. GO analysis of the organoid clusters revealed that terms enriched in the $C_{KY}1$ control cluster were connected to lipid, cholesterol, and alcohol metabolism pathways,

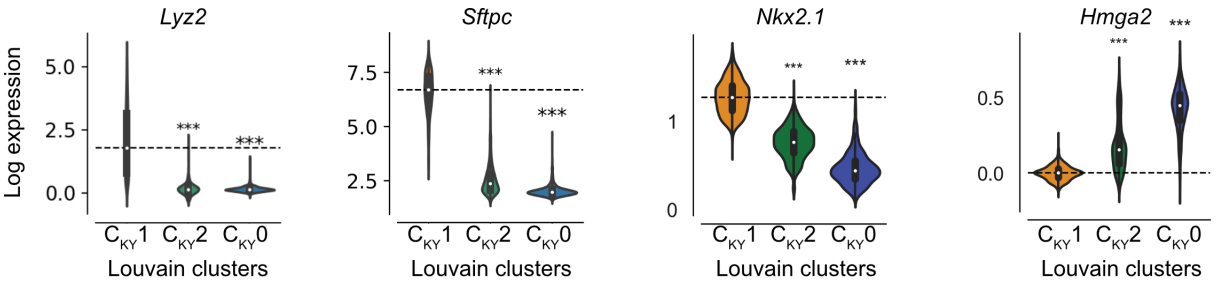

Figure 26: AT2 marker and lung development gene expression
Gene expression values of selected genes in the three Louvain cluster. Data is represented as violin plots. Dashed line marks median expression of control cluster $C_{KY}1$. P values were calculated using the Mann-Whitney rank test.

indicating that these pathways play a role in AT2 function (table 16, appendix). In contrast, $C_{KY}0$ was enriched for "Regulation of I-κkinase/NF-κB signaling", consistent with the elevated NF-κB signature in this cluster (figure 24), and "ERBB signaling", which has been shown to be involved in KRAS G12D lung cancer tumorigenesis elsewhere (Kruspig *et al.*, 2018). Enriched pathways in $C_{KY}2$ were related to mRNA processing, G1/S transition, and translation, likely connected to increased proliferation in this cluster as seen by the proliferation signature (figure 24)

4.2.4 Oncogenic KRAS causes cell transition to a less differentiated state

Transcriptional analysis of KY-CRE organoids on day 7 of organoid culture revealed that KRAS G12D expressing AT2 cells split into two trancriptionally distinct clusters. $C_{KY}2$ had a high proliferation score and expressed TFs connected to early-stage LUAD. In contrast, $C_{KY}0$ was marked by a low proliferation signature and by high expression levels of developmental genes, such as *Sox9* and *Hmga2*. Previous studies have shown that these developmental genes correlate with high stage, metastasis, and worse survival in LUAD patients and GEMMs (Jiang *et al.*, 2010; Zhou *et al.*, 2012; Winslow *et al.*, 2011). To test if the organoids model cancer progression from early- to late-stage transcriptional states, we analyzed the data using RNA-velocity, a computational pipeline that infers cell trajectories based on the ratio of unspliced to spliced mRNA transcripts (La Manno *et al.*, 2018). This pseudotime analysis indicated the direction of cell transitions by vectors on the UMAP plot. There was a clear transition from the proliferative $C_{KY}2$ to the developmental-like $C_{KY}0$ state in the Cre-clusters (figure 27 A). Visualization of *Sox9* expression levels revealed that the vectors followed a *Sox9* gradient, indicating that the TF might drive this transition (figure 27 B). To further evaluate if *Sox9*

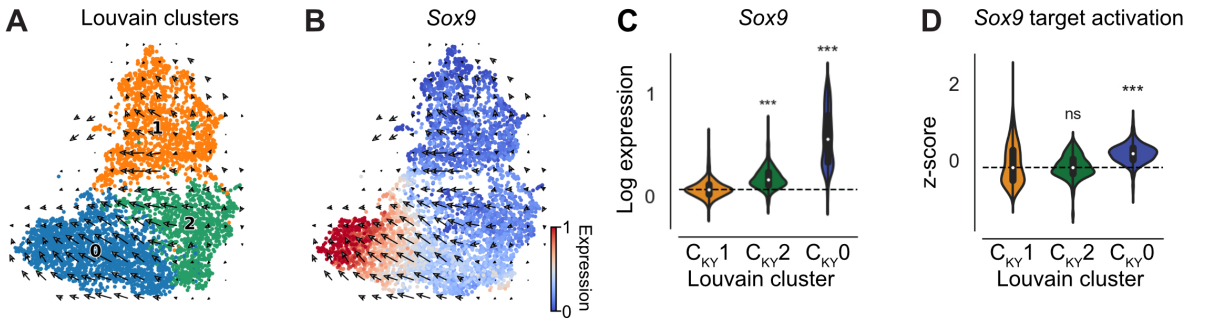

Figure 27: Cell trajectory analysis with RNA-velocity
(A) UMAP showing Louvain clusters with velocity indicated by vectors on the plot. **(B)** *Sox9* expression levels are highlighted on velocity plot. **(C)+(D)** Expression levels of **(C)** *Sox9* and **(D)** *Sox9* target activation signature in each Louvain cluster. Dashed line marks median of control cluster $C_{KY}1$. Data is represented using violin plots. P values were calculated using the Mann-Whitney rank test.

was active in $C_{KY}0$, we checked the expression of *Sox9* target genes using a published gene signature (Han *et al.*, 2018). Although *Sox9* itself was significantly upregulated in both Cre-clusters, only $C_{KY}0$ had an elevated target gene activation signature, indicating that *Sox9* downstream targets were only activated in $C_{KY}0$ (figure 27 C+D). These data suggested that AT2 cells transitioned from a more proliferative, differentiated state to a quiescent, less differentiated state upon oncogenic KRAS induction.

To test if lineage infidelity could be observed, we checked the expression of lineage genes for club cells (*Scgb1a1, Scgb3a2*), BASCs (*Ly6a*), ciliated cells (*Foxj1*), AT1 cells (*Pdpn, Aqp5*), and AT2 cells (*Etv5, Lyx2, Sftpc*) (figure 28). As expected, control cells ($C_{KY}1$) had high expression of AT2 markers, and some expression of AT1 markers in a subset of cells. In contrast, $C_{KY}0$ and $C_{KY}2$ had very low expression of those markers, confirming previous observations. Strikingly, the expression of *Ly6a* (SCA1), a BASC and cancer stem cell (CSC) marker, was upregulated, as were club and ciliated cell markers in a subset of cells (figure 28) (Kim *et al.*, 2005; Curtis *et al.*, 2010). This indicated that KRAS G12D expression in AT2 cells caused the expression of markers of other cell lineages, corroborating findings from human LUAD (Laughney *et al.*, 2020)

4.2.5 Loss of AT2 identity occurs *in vivo* and in patient samples

After only seven days of organoid culture, I found a significant downregulation of AT2 marker genes in the KRAS G12D expressing cancer organoid cells. To rule out that this observation

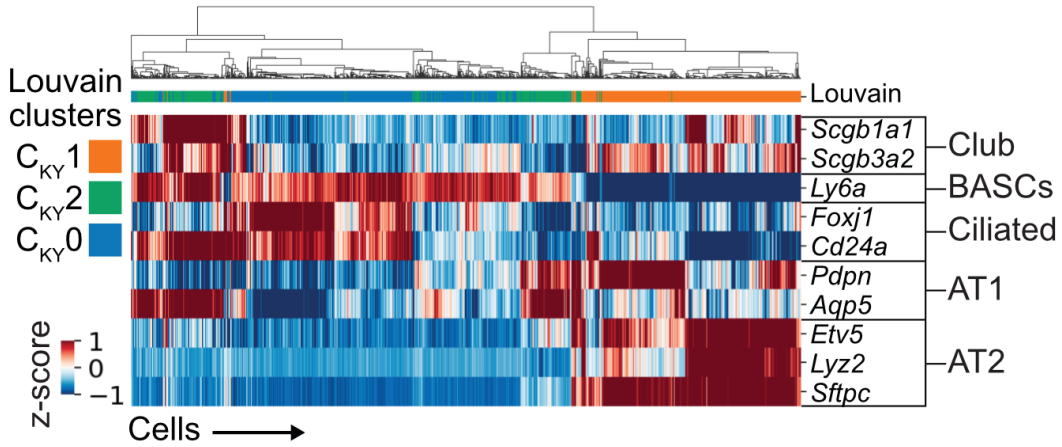

Figure 28: Analysis of lineage infidelity in organoids
ScRNA-Seq data presented as correlation heatmap of individual cells (x-axis) and gene expression values of indicated lineage genes (y-axis). Cells are ordered based on correlation distance calculation. Louvain clusters were color annotated for reference. $C_{KY}1$ represents the control cluster.

was an artifact of organoid culture conditions, I corroborated my findings with *in vivo* and patient data. In collaboration with Dr. Aaron Moye, we analyzed early-stage lesions in the KY GEMM. Briefly, mice were infected with an AdCMV-Cre virus intratracheally as described elsewhere (DuPage *et al.*, 2009). After seven weeks, small hyperplastic areas with clusters of YFP+ cells could be observed in the alveolar region (figure 29 A). At this time point, scRNA-Seq was performed on YFP- and YFP+ epithelial cells. Using the same AT2 signature as used previously, we found that control YFP- AT2 cells had higher expression of AT2 marker genes compared to oncogenic KRAS-expressing YFP+ AT2 cells (figure 29 B).

To test if loss of AT2 differentiation could also be observed in patients, we set up a collaboration with Drs. Jane Yanagawa and Steven Dubinett from University of California Los Angeles. Through this collaboration, we obtained scRNA-Seq data from two patients with stage IA LUAD lesions and matched normal tissue. Both patients harbored oncogenic KRAS mutations. Multiple epithelial and stromal cell types were identified in the dataset; however, only AT2 cells formed a transcriptionally distinct cluster depending on their sample identity (figure 30 A+B). When we checked for levels of the AT2 signature, we found that AT2 cells derived from stage IA lesions had a lower score compared to their normal AT2 counterparts (figure 30 C).

These results showed that the loss of AT2 differentiation genes early after oncogenic KRAS expression was not unique to organoid culture, but could also be observed in mice *in vivo*, and in stage IA LUAD patient samples.

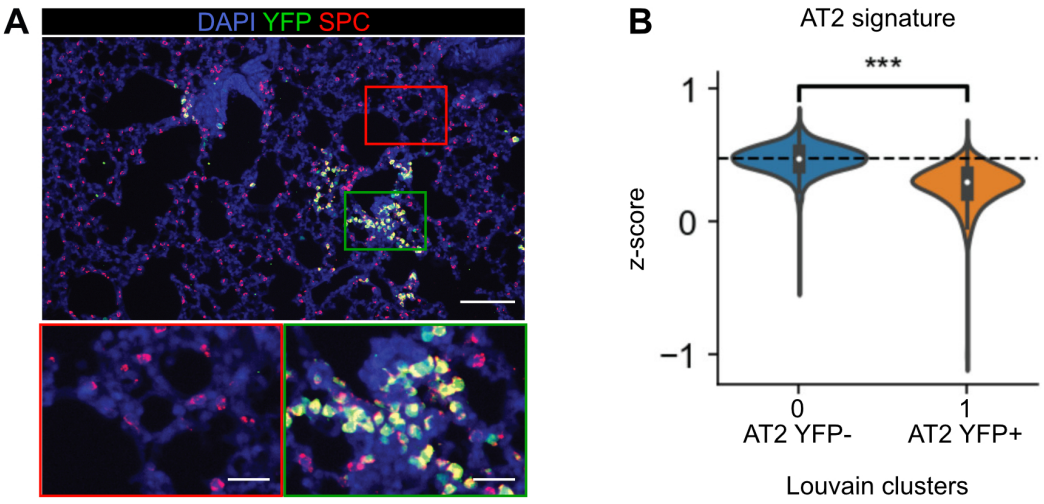

Figure 29: Analysis of AT2 lineage identity in early-stage lesions *in vivo*
Analysis of early-stage LUAD lesions in the KY mouse model seven weeks after AdCMV-Cre induction *in vivo*. Figure was adapted from Dost *et al.* (2020). **(A)** Representative IF image of a lung with an early-stage lesion stained with DAPI (nuclei), YFP (Cre expression), and SPC (AT2 cells). Depicted is a normal alveolar region (red box) and a hyperplastic region with clusters of YFP+ cells (green box). Scale bar low magnification = 100 µm. Scale bar insets = 25 µm. **(B)** AT2 signature score of control AT2 cells (YFP-) and KRAS G12D expressing AT2 cells (YFP+), taken from scRNA-Seq dataset.

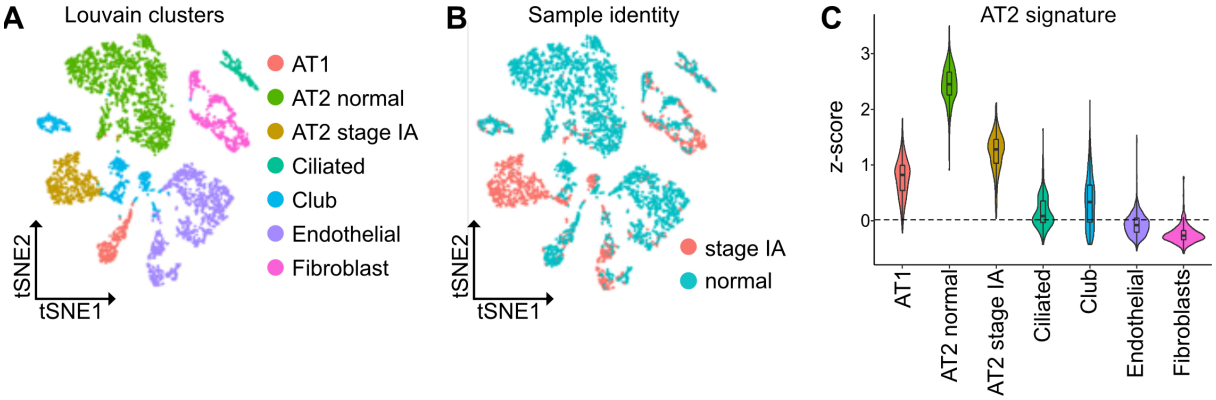

Figure 30: Analysis of AT2 lineage identity in early-stage lesions in patients
ScRNA-Seq analysis of two patients with stage IA KRAS-driven LUAD lesions and matched healthy tissue. Figure was adapted from Dost *et al.* (2020). **(A)** tSNE plot colored by Louvain clusters. Alveolar type 1 (AT1), alveolar type 2 (AT2), AT2 stage IA, ciliated cells, club cells, endothelial cells, and fibroblasts were annotated. **(B)** tSNE plot colored by sample identity. Cells were derived from stage IA LUAD lesions and matched normal tissue. **(C)** AT2 signature expression in annotated Louvain clusters. The AT2 stage IA cluster had a lower z-score than the AT2 normal cluster.

4.2.6 Cancer organoids can be used to identify and functionally validate targets

To test if the organoids can be used to find cancer targets, I made use of the transcriptional data to find commonly deregulated genes. I reasoned that genes upregulated in all organoid datasets were important for the first steps of cancer progression. I used the DE analysis for KY-CRE organoids and compared the bulk RNA-Seq and the two CRE Louvain clusters $C_{KY}0$ and $C_{KY}2$ of the scRNA-Seq analysis. I selected genes that were upregulated in all three datasets, excluding TFs, which are difficult to target with small molecules (figure 31). From the remaining 42 genes, I selected three targets with readily available small molecule inhibitors. Glucosaminyl N-Acetyl Transferase 3 (*Gcnt3*) is a core enzyme of mucin-synthesis that plays a role in KRAS driven pancreatic cancer; the enzyme can be inhibited selectively by the compound Talniflumate (Gupta *et al.*, 2020; Rao *et al.*, 2016). Glutathione S-Transferase Omega 1 (*Gsto1*) is an enzyme involved in the metabolism of xenobiotics and carcinogenes that is overexpressed in several cancers; it was shown to be inhibited by the small molecule GSTO1-IN-1 (Ramkumar *et al.*, 2016). Ephrin type-A receptor 2 (*Epha2*) encodes a member of the ephrin receptor family, known to be deregulated in a variety of cancers, including NSCLC (Ieguchi and Maru, 2019; Brannan *et al.*, 2009b). This family of receptor tyrosine kinases can be inhibited by the small molecule ALW-II-42-27 (Choi *et al.*, 2009; Song *et al.*, 2017).

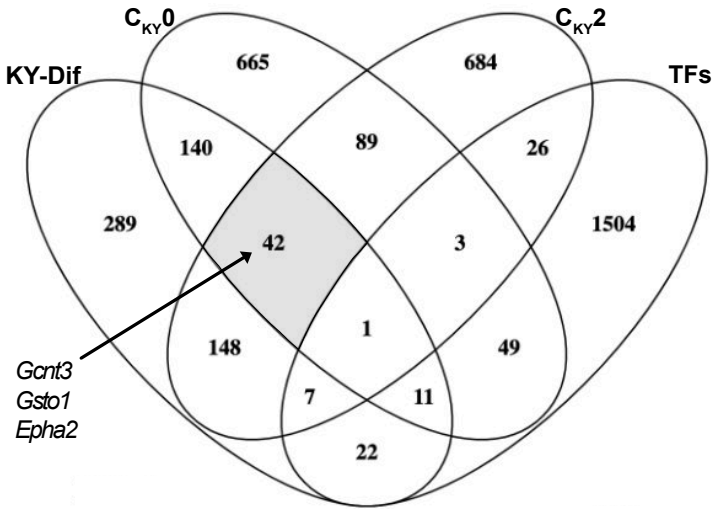

Figure 31: Venn diagram of upregulated genes in KY cancer organoids identified in transcriptional analysis
Overlap of genes upregulated in KY-CRE organoids compared to their respective controls, identified in bulk RNA-Seq (KY-Dif) and in the two Louvain clusters $C_{KY}0$ and $C_{KY}2$ of the scRNA-Seq analysis. Transcription factors (TFs) were also included in the Venn diagram. The shaded area highlights the overlap between the bulk RNA-Seq analysis, and both scRNA-Seq analysis clusters, excluding TFs. The three genes *Gcnt3*, *Gsto1*, and *Epha2* were picked for further analysis.

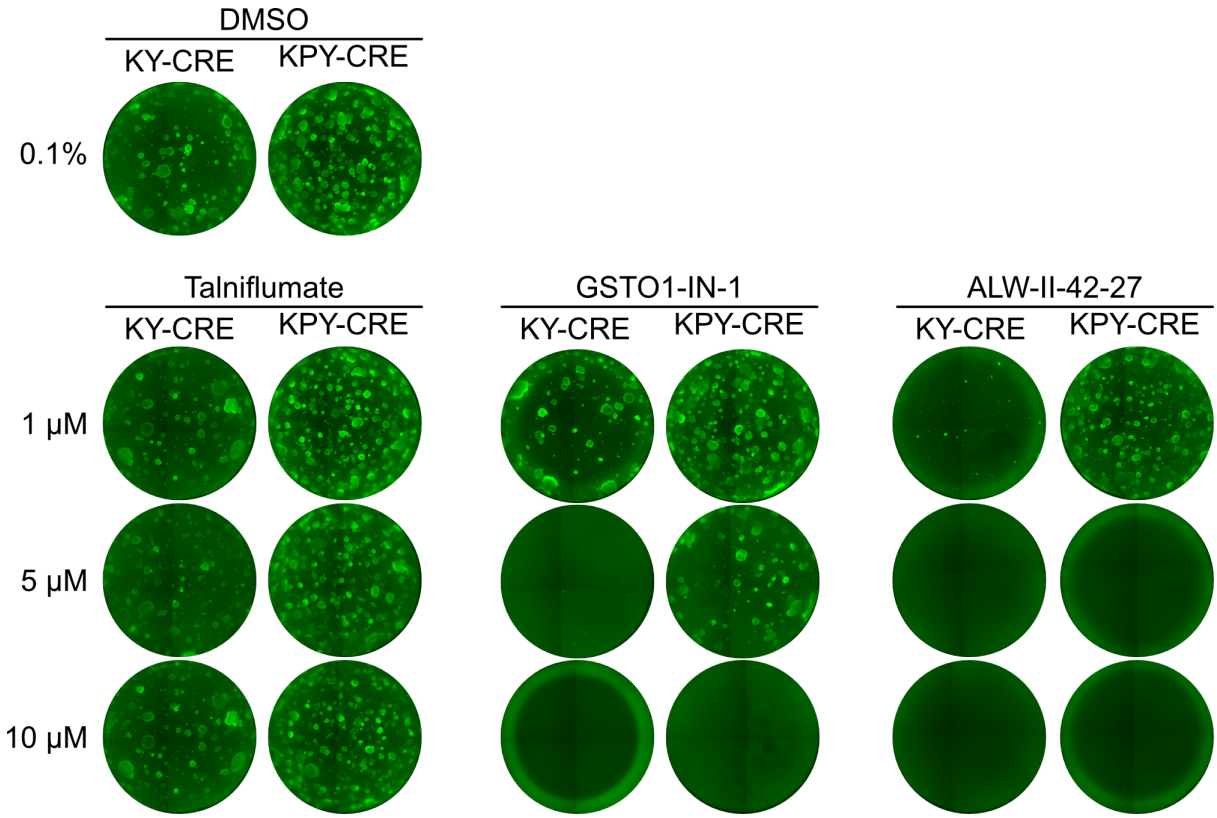

Figure 32: Treatment of cancer organoids with small molecules
Representative images of cancer organoids that were treated with DMSO (control) or three different concentrations of indicated small molecules from day 0 of organoid culture. Media and small molecules were replaced every other day for 14 days. Whole well images were taken in the YFP channel on day 14 of organoid culture. Biological replicates n=1. Technical replicates n=4.

To test if any of the small molecules could reduce or abolish cancer organoid formation, I plated KY-CRE and KPY-CRE organoids as described before and added the small molecules in different concentrations on day 0 of organoid culture, renewing them every other day for 14 days. Images of the wells on day 14 of organoid culture revealed that Talniflumate had no visible effect on organoid formation (figure 32). GSTO1-IN-1 abolished organoid formation at 5 µM in KY-CRE, and at 10 µM in KPY-CRE organoids. ALW-II-42-27 reduced organoid formation visibly at 1 µM and 5 µM in KY-CRE and KPY-CRE, respectively.

Because ALW-II-42-27 was the most promising candidate, having a visible effect on organoid growth at a concentration of 1 µM, I repeated the experiment including nanomolar concentrations of the small molecule. Moreover, I added Y-CRE control organoids to test if the effect on organoid growth was cancer cell specific, or due to general cell toxicity. Strikingly, at 100 nM both KY-CRE and KPY-CRE cancer organoids had a visible reduction in organoid size, while

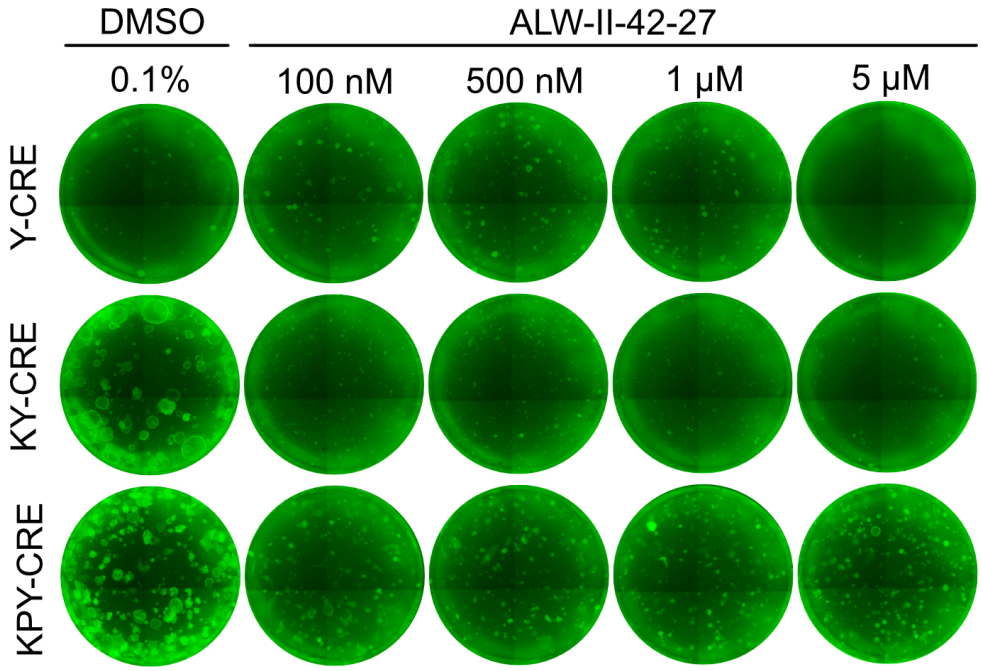

Figure 33: Treatment of organoids with the small molecule ALW-II-42-27
Representative images of organoids that were treated with DMSO (control) or four different concentrations of ALW-II-42-27 from day 0 of organoid culture. Media and small molecules were replaced every other day for 14 days. Whole well images were taken in the YFP channel on day 14 of organoid culture. Biological replicates n=2. Technical replicates n=4.

no effect was observed in the Y-CRE control cells up to 1 µM. This result indicated that the inhibition of the ephrin receptor family selectively targeted KRAS-driven lung cancer cells, but not their healthy epithelial cell counter-parts, at nanomolar concentrations.

4.3 Time course analysis of organoid progression with single cell resolution

The cancer organoids displayed histological and transcriptional hallmarks of cancer cell progression and were used to identify a target essential for cancer organoid growth. While the previous studies focused on the establishment of the system and the transcriptional changes of early steps of oncogenic KRAS driven cell transformation, I next sought to understand the whole transcriptional landscape of control and cancer organoid progression, including the role of co-cultured MECCs.

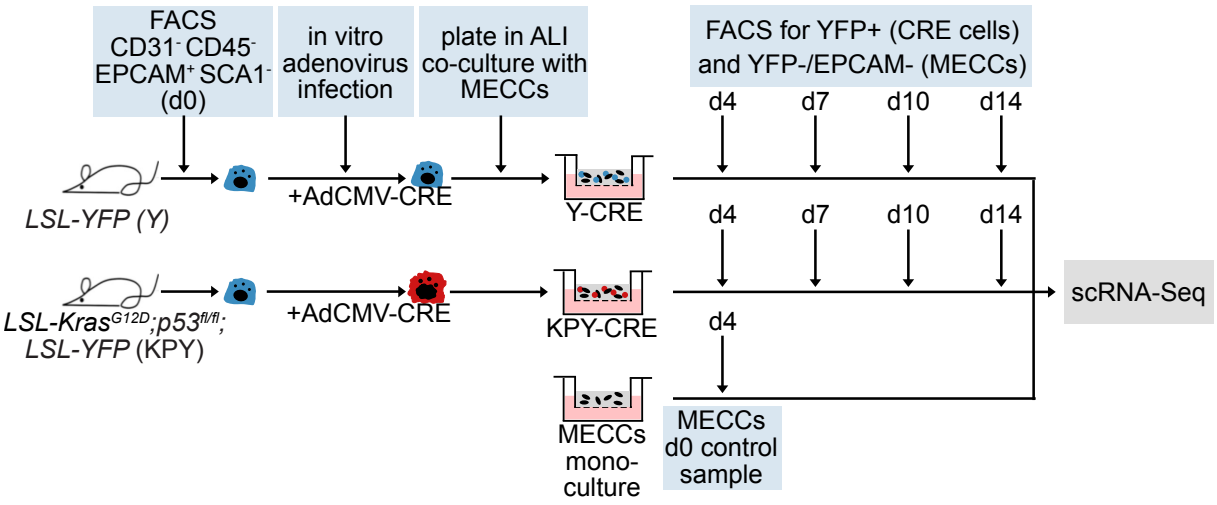

Figure 34: Workflow of sample preparation for scRNA-Seq time course
Experimental strategy to prepare organoid samples for Y and KPY scRNA-Seq time course. Day (d)0 epithelial cell samples represent SCA1- AT2 cells before virus infection. D0 mesenchymal cultured cells (MECCs) control samples represent MECCs from 3D monocultures, isolated at d4. D4, d7, d10, and d14 samples were taken from organoid cultures and either sorted for YFP+ cells (infected AT2 cells), or YFP-/EPCAM- cells (MECCs) from co-cultures. ScRNA-Seq for all samples was performed on the same day to avoid technical variation.

4.3.1 Transcriptional time course reveals differences in progression of alveolar and cancer organoids

Cancer progression is a complex process that involves increased proliferation, cell state changes, and the accumulation of mutations. While AT2-derived cells with an activating KRAS mutation were used to model the early steps of cancer progression, KRAS alone does not cause advanced tumors in GEMMs, but additional mutations such as the deletion of *Trp53* are required. Therefore, in order to study the whole range of cancer progression, I performed a scRNA-Seq time course using Y-CRE and KPY-CRE organoids. As described previously, I generated AT2 derived organoids from Y or KPY mice *in vitro*. On day 4, 7, 10, and 14 of organoid culture, I sorted YFP+ cells using FACS. In addition to the YFP+ epithelial cells, I also collected YFP-/EPCAM- MECCs that had been co-cultured as supporting cells (figure 34). To obtain a origin cell reference point, I additionally prepared samples of sorted AT2 cells before virus infection (day 0 control), and of MECCs that had been kept in 3D mono-culture for four days (day 0 MECCs control). As before, scRNA-Seq was performed using the 10xGenomics platform and analyzed using the scanpy environment (Wolf *et al.*, 2018). After filtering of low quality and contaminating cells, the data was imputated and 29,933 epithelial cells and 41,319 MECCs remained for analysis, with all time points represented (figure 35, 36 A+B).

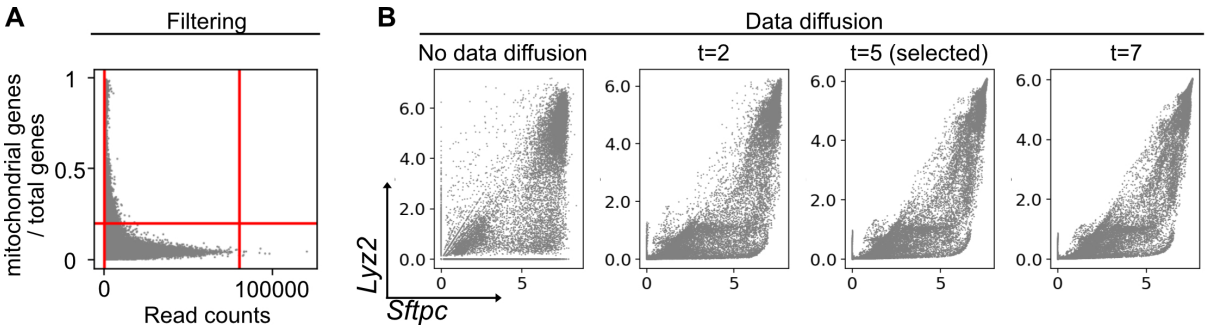

Figure 35: Filtering and imputation of scRNA-Seq time course data
(A) Correlation between read count and mitochondrial genes to total genes ratio. Red lines mark cut-offs: Cells with read counts < 200 and > 8000, and mitochondrial genes to total gene ratio < 0.2 were removed. **(B)** Correlation between the two AT2 genes *Sftpc* and *Lyz2* after t rounds of data diffusion (van Dijk *et al.*, 2018). T=5 was selected because of the restored correlation between the two selected genes.

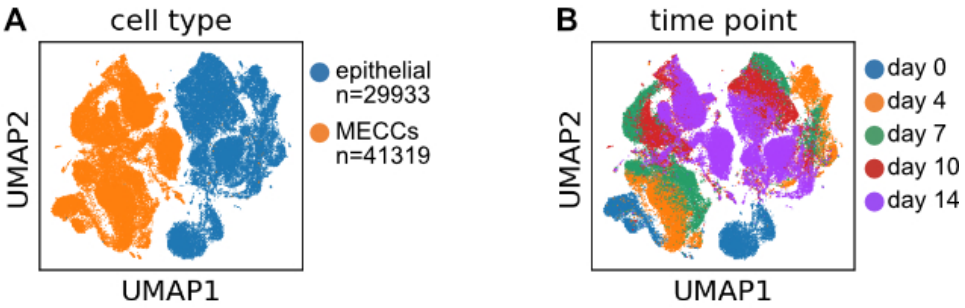

Figure 36: Clustering of all cells in scRNA-Seq time course data
UMAP representation of primary epithelial and organoid derived epithelial YFP+ cells, and mesenchymal cultured cells (MECCs) colored by **(A)** cell type with cell numbers indicated, or **(B)** time point.

To facilitate the analysis, we first re-clustered the epithelial cells without MECCs. As expected, day 0 cells from Y-CRE and KPY-CRE samples overlapped in the UMAP plot (figure 37 A). Day 0 cells represented FACS sorted AT2 cells before Cre induction and were therefore not expected to be distinct between the two genotypes. Despite the overlapping day 0 cells of the two samples, all other organoid time points split by sample origin with very little overlap (figure 37 B). Most strikingly, day 4 cells of Y-CRE and KPY-CRE samples clustered distinct from each other, indicating that oncogenic KRAS expression and loss of *Trp53* caused a significant transcriptional change within the first 4 days after induction (figure 37 A-C). Both organoid types underwent a transition over the 14 day time period, indicated by the mostly distinct transcriptional states at day 14 and the little overlap with the other time points. Because the previous KY-CRE organoid data revealed downregulation of AT2 differentiation genes (*Lyz2, Sftpc, Nkx2-1, Etv5*) and upregulation of lung developmental genes (*Sox9, Hmga2*) upon oncogenic KRAS expression, I next checked if those genes were also differentially expressed in the time course data-set. Strikingly, AT2 differentiation genes were expressed more highly in day 0 and in Y-CRE derived cells, while developmental genes were exclusively present in KPY-CRE organoid derived cells, confirming previous observations (figure 37 D).

4.3.2 Alveolar organoid cells follow AT1 and AT2 differentiation trajectories

The epithelial cell data pointed out transcriptional differences in KPY-CRE cells compared to Y-CRE cells in all analyzed time points. In order to understand the different progressions of the two organoid types, I first focused on understanding the cell dynamics in Y-CRE control organoids. For this purpose, we subset and re-clustered the control samples using ForceAtlas2 (FA2), to better visualize cell lineages (figure 38 A) (Jacomy *et al.*, 2014). As seen before, day 0 cells clustered distinct from the organoid-derived cells, indicating that the difference between *in vivo* and *in vitro* environments caused a transcriptional shift in the cells. Interestingly, on the FA2 projection it became clear that the organoid-derived cells split into two different trajectories, represented by the two day 14 tips of the plot. To understand the dynamics of cell transition better, we checked expression levels of the AT2 signature used before (Franzén *et al.*, 2019). As expected, day 0 cells had strong expression of AT2 genes, confirming the enrichment of AT2 cells in the SCA1- fraction (figure 38 B). Interestingly, day 4 organoid cells had a reduction in AT2 signature gene expression compared to day 0 cells, but the cells re-acquired the AT2 signature along one of the transition tips. Because AT2 cells are known to differentiate into AT1 cells *in vivo* and in organoid culture, we checked an AT1 signature

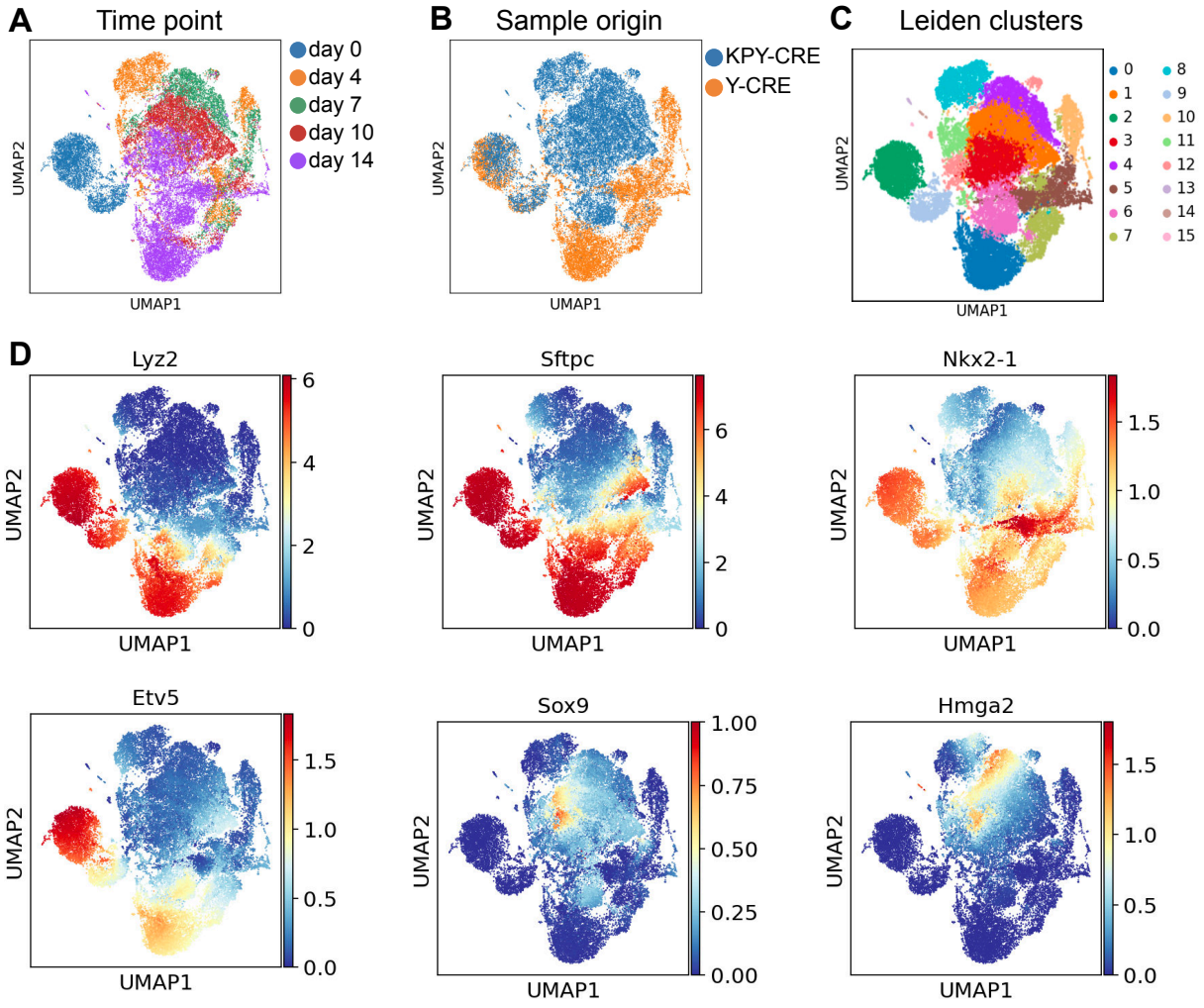

Figure 37: Clustering and gene expression of epithelial cells
UMAP representation of scRNA-Seq time course data of FACS sorted Y and KPY SCA1- cells (day 0), and Y-CRE and KPY-CRE organoid cells (day 4 - day 14) colored by **(A)** time point, **(B)** sample origin, and **(C)** Leiden clusters. (**D**) Gene expression levels of selected AT2 differentiation and lung developmental genes visualized on UMAP plots.

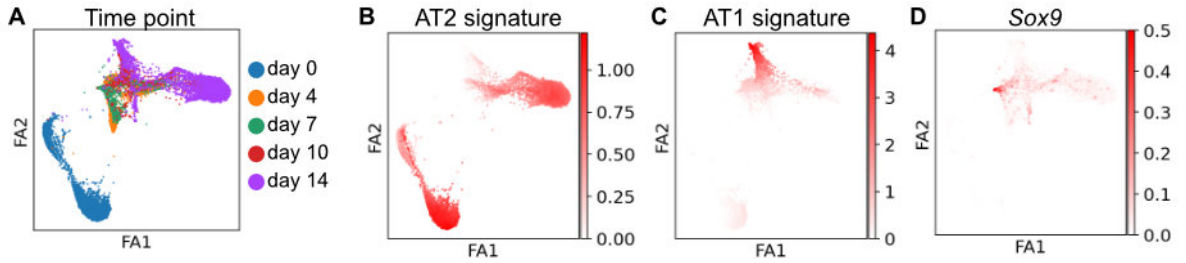

Figure 38: Clustering of time points and gene signature expression of Y control epithelial cells
FA2 representation of scRNA-Seq time course data of FACS sorted Y SCA1- cells (day 0), and Y-CRE organoid cells (day 4 - day 14) colored by **(A)** time point, **(B)** AT2 signature, **(C)** AT1 signature, and **(D)** *Sox9* expression.

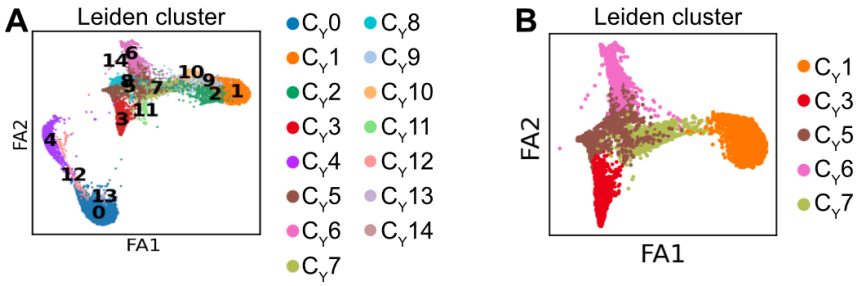

Figure 39: Clustering and selected clusters of Y control epithelial cells
FA2 plots of Y control sample. **(A)** 15 Leiden clusters were identified. **(B)** Five clusters, the root C_Y3, the tip clusters C_Y1 and C_Y6, and the branch clusters C_Y5 and C_Y7 were selected for further downstream analysis.

composed of genes associated with AT1 cell identity (Franzén *et al.*, 2019). Unsurprisingly, cells that formed the non-AT2 tip of the plot had high expression of the AT1 signature (figure 38 C). This indicated that AT2 cells initially downregulated AT2 signature genes in organoid culture, then bifurcated and differentiated into AT2 and AT1 lineages. Because *Sox9* was a gene highly expressed in the KY-CRE organoid data and is known to play a crucial role in lung development, we checked *Sox9* expression in the Y-CRE control samples (figure 38 D). As expected, *Sox9* expression was very low, indicating that the control cells did not dedifferentiate to a more developmental-like state, as was previously observed in the KY-CRE cancer organoids.

To further delineate the factors that drove the two observed lineage trajectories, we used the Leiden algorithm to divide the cells into transcriptionally similar clusters; 15 Leiden clusters were identified (figure 39 A). In order to understand the trajectories better, we selected specific clusters to focus on. Clusters from the Y dataset will be indicated by a C_Y nomenclature. C_Y3 represented the cluster furthest from the day 14 AT1 and AT2 tips and was mainly comprised

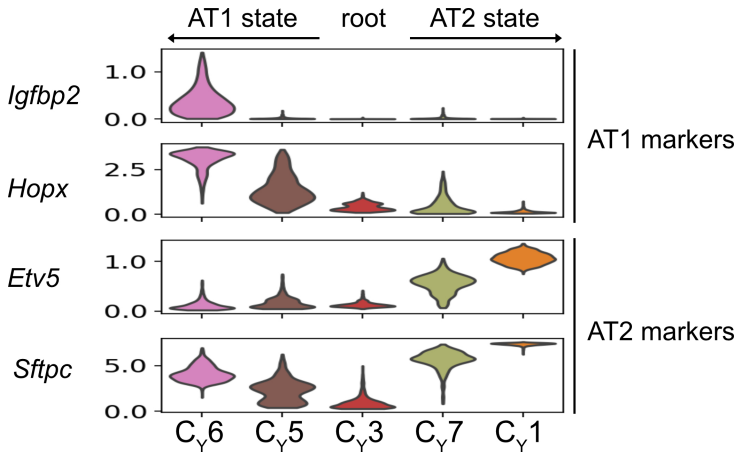

Figure 40: Gene expression in root, tip, and branch point clusters in alveolar organoids
Violin plot representation of expression levels of selected genes and transcription factors in root (C_Y3), AT1 branch point (C_Y5), AT1 tip (C_Y6), AT2 branch point (C_Y7), and AT2 tip (C_Y1) clusters of Y-CRE organoids. Plots are ordered by AT1 and AT2 states, bifurcating from the center root cluster.

of day 4 and day 7 organoid cells; I therefore termed it "root" cluster. C_Y6 and C_Y1 built the two day 14 trajectory "tips" and had high expression of AT1 and AT2 signatures, respectively. They appeared to be the most differentiated clusters in the control organoid data-set. C_Y5 and C_Y7 were situated at the trajectory "branch" point of the AT1 and AT2 tips, respectively, most likely comprising transitioning cells (figure 38 B+C and 39 B). We reasoned that analyzing the root, tip, and branch clusters might give insights into genes that drive the fate decisions of the respective trajectories. Therefore, we performed DE analysis of genes and TFs on the five selected clusters (table 17, appendix; table 11). In the AT1 trajectory clusters C_Y5 and C_Y6, we identified *Hopx*, a TF known to drive AT1 fate trajectory and highly expressed in AT1 cells (table 11; figure 40). The top differentially expressed gene in the AT1 tip cluster C_Y6 was *Ager*, a known marker for AT1 identity (Chung and Hogan, 2018). Interestingly, the branch C_Y5 had elevated expression of the Activator protein-1 (AP-1) TF components *Jun*, *Junb*, and *Jund*, indicating that they might play a role in AT1 differentiation. The AT2 tip and branch clusters C_Y1 and C_Y7 had high expression of *Etv5*, known to be important for AT2 identity maintenance (table 11; figure 40). C_Y1 had high expression of the AT2 genes *Sftpc* and *Lyz2*, but also expressed *Id2*, an important gene during lung development (Rawlins *et al.*, 2009a). Interestingly, the AT2 branch C_Y7 had elevated expression of Hypoxia-inducible factor 1-alpha (*Hif1a*), a gene that has recently been shown to be important for AT1 differentiation (Choi *et al.*, 2020). The root C_Y3 had very low expression of *Hopx* and *Igfbp2*, a terminally differentiated AT1 marker, and *Etv5* and *Sftpc* (figure 40). This cluster might therefore represent an uncommitted progenitor cell state. TFs highly expressed in the root

cluster included *Ybx1, Ybx3, Foxq1*, and *Id1*, indicating that those factors might play a role in maintaining a bipotent progenitor cell state (table 11; figure 41).

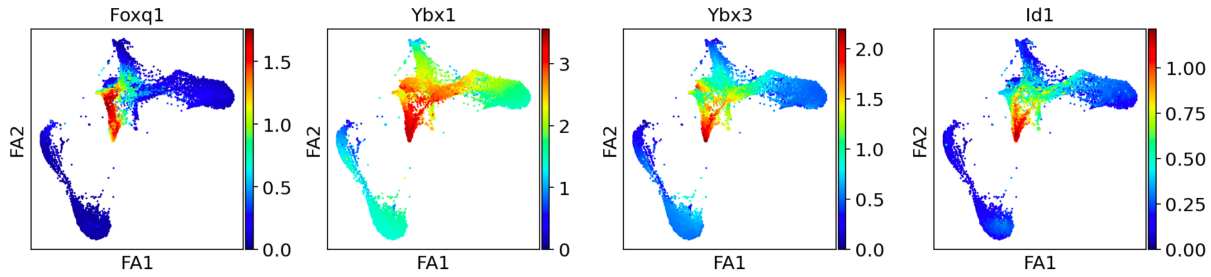

Figure 41: Transcription factors expressed in Y root cluster
FA2 representation of scRNA-Seq time course data of Y control cells. Cells are highlighted based on gene expression levels of indicated genes. All genes shown represent transcription factors highly expressed in the root cluster C_Y3.

Table 11: Top 10 upregulated transcription factors in Y control scRNA-Seq time course clusters. C_Y3 represents the root cluster, C_Y5 the AT1 transitioning cluster, C_Y6 the AT1-like cluster, C_Y7 the AT2 transitioning cluster, and C_Y1 the AT2-like cluster.

rank	C_Y1	C_Y3	C_Y5	C_Y6	C_Y7
1	Etv5	Ybx1	Zfp36l1	Hopx	Ybx1
2	Epas1	Ybx3	Jund	Tead1	Ssrp1
3	Id2	Foxq1	Sox4	Nkx2-1	Ybx3
4	Cebpa	Id1	Junb	Cux1	Smarcc1
5	Rbpjl	Dnajc2	Runx1	Tfdp2	Lyar
6	Srebf1	Lyar	Foxq1	Mxd4	Purb
7	Elf5	Ahr	Ahr	Foxn3	E2f1
8	Tfcp2l1	Sox2	Jun	Scx	Hif1a
9	Foxp1	Ssrp1	Hopx	Creb3l2	Nfic
10	Tsc22d3	Hmga1	Arid5b	Tef	Dnajc2

4.3.3 Cancer organoids transition to dedifferentiated state

In order to understand how the lineage trajectory of the control cells compared to cancer cells, we analyzed the KPY-CRE sample derived cells using FA2 projection. As seen previously, day 0 cells clustered distinct from all organoid-derived cells (day 4 to day 14) (figure 42 A). In contrast to the control cells, day 4 cells clustered separate from the other time points, suggesting that

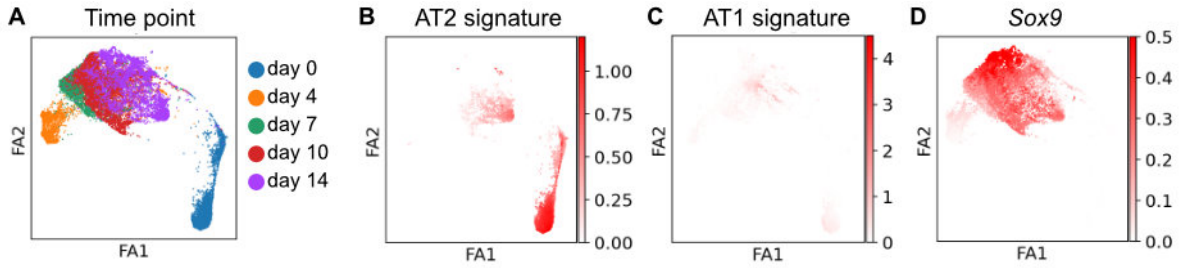

Figure 42: Clustering of time points and gene signature expression of KPY epithelial cells
FA2 representation of scRNA-Seq time course data of FACS sorted KPY SCA1- cells (day 0), and KPY-CRE organoid cells (day 4 - day 14) colored by **(A)** time point, **(B)** AT2 signature, **(C)** AT1 signature, and **(D)** *Sox9* expression.

there were significant transcriptional changes between day 4 and day 7 of cancer organoid culture. Furthermore, contrary to the control organoid trajectory, the transition of the cancer cells appeared to be unidirectional. Expression levels of the AT2 signature showed that day 0 cells had high expression of AT2 markers, while day 4, day 7, and day 10 cells had no or very low AT2 signature expression (figure 42 B). On day 14, the AT2 signature was partially reacquired by some cells at the tip of the plot. In stark contrast to the control cells, the cancer cells did not have an AT1 trajectory, indicated by the absence of the AT1 signature (figure 42 C). In accord with this observation and our previous observations in the KY organoids, the cancer cells had high expression levels of the developmental gene *Sox9*, especially at time points day 10 and day 14 (figure 42 D).

To focus our analysis, we selected five clusters, similar to our previous analysis of the control cells (figure 43 A+B). Clusters from the KPY dataset will be indicated by a C_{KPY} nomenclature. $C_{KPY}4$ was made up almost exclusively of day 4 cells and therefore represented the "root" cluster. $C_{KPY}2$ and $C_{KPY}5$ contained mostly day 14 cells and were therefore termed "late-stage" clusters, while $C_{KPY}0$ and $C_{KPY}8$ contained day 10 cells and represented "transitioning" clusters. Similar to before, we performed unbiased DE analysis on the five clusters and focused on TFs (table 18, appendix; table 12). Surprisingly, the root cluster $C_{KPY}4$ had elevated expression of similar TFs as the root control cluster, namely *Ybx1*, *Ybx3*, and *Id1*, indicating that similar transcriptional programs were activated in both root clusters (figure 44). However, the cancer root cluster additionally had increased expression levels of the well-known TF Krüppel Like Factor 4 (*Klf4*), one of the four factors essential for inducing pluripotency (Takahashi and Yamanaka, 2006). Interestingly, $C_{KPY}8$ and $C_{KPY}5$ expressed *Sftpc*, *Etv5*, and *Id2*, similar to the control AT2 tip cluster (table 18, appendix; table 12). This is in line with the observed re-acquisition of an AT2 program in those clusters (figure 42 B). In contrast,

$C_{KPY}2$ expressed *Hopx* and AP-1 TF components such as *Junb* and *Fos*, similar to the AT1 transitioning cluster in the control data. However, in contrast to the control data, an AT1 program was not observed in any of the clusters (figure 42 C). $C_{KPY}0$ expressed some of the TFs observed in the root clusters, such as *Foxq1* and *Ybx1*. Moreover, the lung development factor *Hmga2* was upregulated, as was *Sox4*, known to play a role in cancer cell stemness, EMT, and metastasis (Lourenço and Coffer, 2017; Moreno, 2020). Although no clear branching points were visible in the KPY time course data, transcriptional analysis revealed that $C_{KPY}5$ and $C_{KPY}8$ partially reacquired an AT2 program, while $C_{KPY}0$ represented a cluster with increased expression of developmental and EMT genes that are associated with high-grade cancer. To test if EMT could be observed in KPY organoid cells, I stained day 14 organoid slides for the mesenchymal marker vimentin. Indeed, single YFP+ cells had a round appearance and looked detached from the other organoid cells, untypical for epithelial cells. Those cells expressed vimentin protein, indicating that they might be undergoing EMT (figure 45).

Table 12: Top 10 upregulated transcription factors in KPY scRNA-Seq time course clusters. $C_{KPY}4$ represents the root cluster, $C_{KPY}0$ and $C_{KPY}8$ transitioning clusters, and $C_{KPY}2$ and $C_{KPY}5$ late-stage clusters.

rank	$C_{KPY}0$	$C_{KPY}2$	$C_{KPY}4$	$C_{KPY}5$	$C_{KPY}8$
1	Sox4	Bhlhe40	Ybx1	Cebpa	Id2
2	Gtf3a	Maff	Ybx3	Epas1	Etv5
3	Carhsp1	Junb	Fosl1	Etv5	Cebpa
4	Hmga2	Mxi1	Id1	Nkx2-1	Rbpjl
5	Onecut3	Hopx	Lyar	Srebf1	Foxp2
6	Tead2	Tsc22d1	Zfp36l2	Rbpjl	Mbd3
7	Smarcc1	Fos	Ppard	Zfp467	Hhex
8	Foxq1	Elf3	Dnajc2	Id2	Sp5
9	Dnajc2	Aff4	Klf4	Elf5	Elf5
10	Ybx1	Ddit3	Ssrp1	Nfia	Yeats4

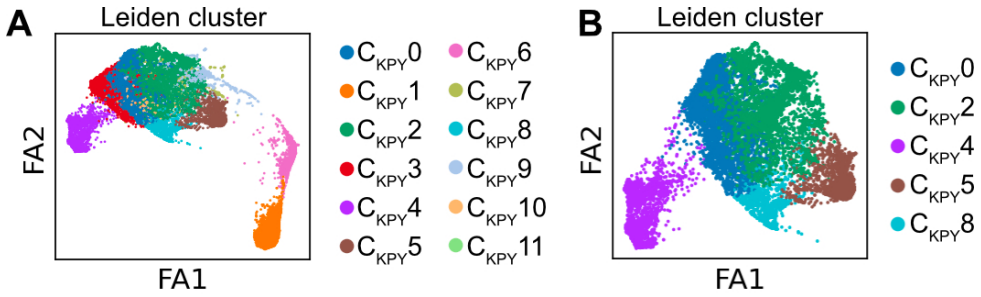

Figure 43: Clustering and selected clusters of KPY epithelial cells
FA2 plots of KPY samples. **(A)** 12 Leiden clusters were identified. **(B)** Five clusters, the root $C_{KPY}4$, the late-stage clusters $C_{KPY}2$ and $C_{KPY}5$, and the transitioning clusters $C_{KPY}0$ and $C_{KPY}8$ were selected for further downstream analysis.

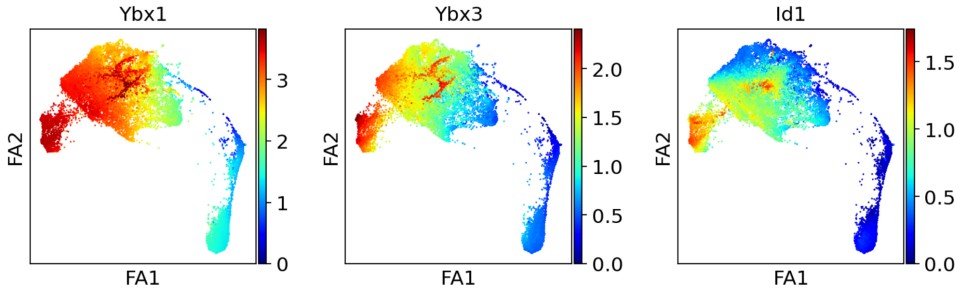

Figure 44: Transcription factors expressed in KPY root cluster
FA2 representation of scRNA-Seq time course data of KPY cells. Cells are highlighted based on gene expression levels of indicated genes. All genes shown represent transcription factors highly expressed in the root cluster $C_{KPY}4$.

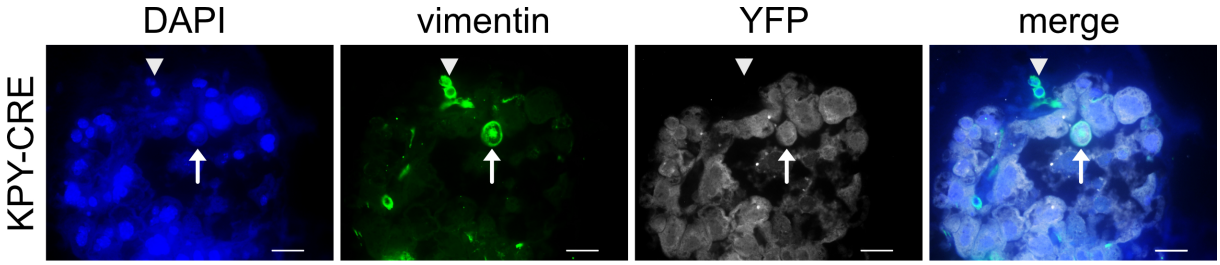

Figure 45: Vimentin stainings in KPY-CRE organoid culture
KPY-CRE day 14 organoids were stained for DAPI (nuclei), vimentin (mesenchymal marker) and YFP (Cre-expressing cells). Arrow points to a cell co-expressing YFP and vimentin. Arrowhead points to MECCs expressing vimentin but not YFP. Scale bar = 25 µm.

4.3.4 Direct comparison of root clusters reveals their differences

The root clusters of both Y-CRE and KPY-CRE organoids displayed transcriptional similarities. Firstly, both clusters had lower expression of AT2 signature genes compared to day 0 cells and compared to cells that later reacquired AT2 features. In order to test if the loss of AT2 signature was equal between the two root clusters, we compared them directly. We determined from the clustering shown in figure 37 A-C that C10 and C8 represented the control Y and the KPY root clusters, respectively. In a direct comparison, we found that C8 had a decreased AT2 signature compared to C10, indicating that the loss of AT2 identity was more severe in the cancer cell context (figure 46).

Secondly, similar TFs were elevated in the root clusters compared to their respective later time point clusters, including *Ybx1, Ybx3*, and *Id1*. Therefore, we decided to do a direct transcriptional comparison of C10 and C8 using DE analysis of TFs (table 13). We found that *Ybx1*, *Ybx3* and *Id1* were more highly expressed in the KPY root cluster compared to Y. Furthermore, *Klf6* and *Klf4* had elevated expression levels in KPY, the latter being important for pluripotency. In the control cells, *Foxq1* was the top differentially expressed TF, followed by *Trp53*, which was knocked out in the KPY sample. Moreover, the AT1 TF *Hopx* was upregulated, as was *Nkx2-1*, an important lung identity TF. Surprisingly, *Sox2*, a TF usually expressed in airway cells, also had elevated expression. Despite obvious similarities of the two root clusters, a direct comparison emphasized their transcriptional differences and highlighted factors that might drive tumorigenesis.

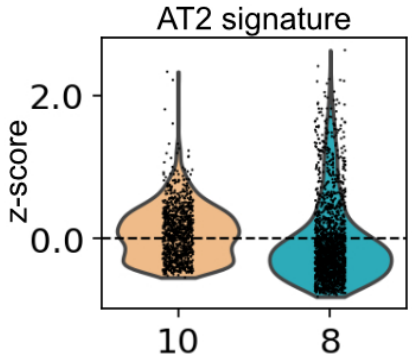

Figure 46: Comparison of AT2 signatures in Y and KPY root clusters
Violin plot representation of AT2 signature level in root clusters 10 (Y) and 8 (KPY). Refer to figure 37 for clustering. Dotted line marks z-score = 0.

Table 13: Top 10 upregulated transcription factors in Y and KPY root clusters. Cluster 10 represents the Y control root cluster in the scRNA-Seq time course data, while cluster 8 represents the KPY root cluster.

rank	Y root cluster 10	KPY root cluster 8
1	Foxq1	Ybx1
2	Trp53	Fosl1
3	Hopx	Ybx3
4	Sox2	Ppard
5	Ehf	Klf6
6	Nkx2-1	Tsc22d1
7	Zfp36l1	Foxp1
8	Hmga1	Klf4
9	Tcf4	Atf4
10	Pbx1	Id1

4.3.5 A proposed working model for alveolar and cancer organoid progression

Considering the transcriptional time course analysis of the Y-CRE control and KPY-CRE cancer organoids, I developed a working model for alveolar and cancer cell progression (figure 47). Because this model is solely based on the TF analysis, it requires future functional validation. The upper part of the schematic represents the progression of the control cells, while the lower part shows the cancer cell progression. In both cases, the cell of origin are AT2 cells. Taking stem cells out of their organ environment and putting them into organoid culture as single cells is similar to an organ injury. Therefore, it is not surprising that the first response of the cells was similar between the two genotypes; the TFs *Ybx1, Ybx3*, and *Id1* were upregulated. In the control cells, *Foxq1* also had elevated expression. In contrast, the cancer cells downregulated the lung identity TF *Nkx2-1* and upregulated the pluripotency factor *Klf4*, resulting in a cell state that was less differentiated than the control cell state. The control cells were in a state of a bipotent progenitor, and split into two distinct trajectories. On one trajectory, *Etv5* and *Id2* were upregulated and cells adapted an AT2-like transcriptional state. On the second trajectory, *Hopx* was increasingly expressed, resulting in an AT1-like state. In contrast, the cancer cells did not follow two clear trajectories. After the initial dedifferentiation, *Sox9* was expressed in almost all cells. Even though transcriptionally distinct cell states were less obvious compared to the control cells, the progression could be split into two paths. Some cells reacquired an

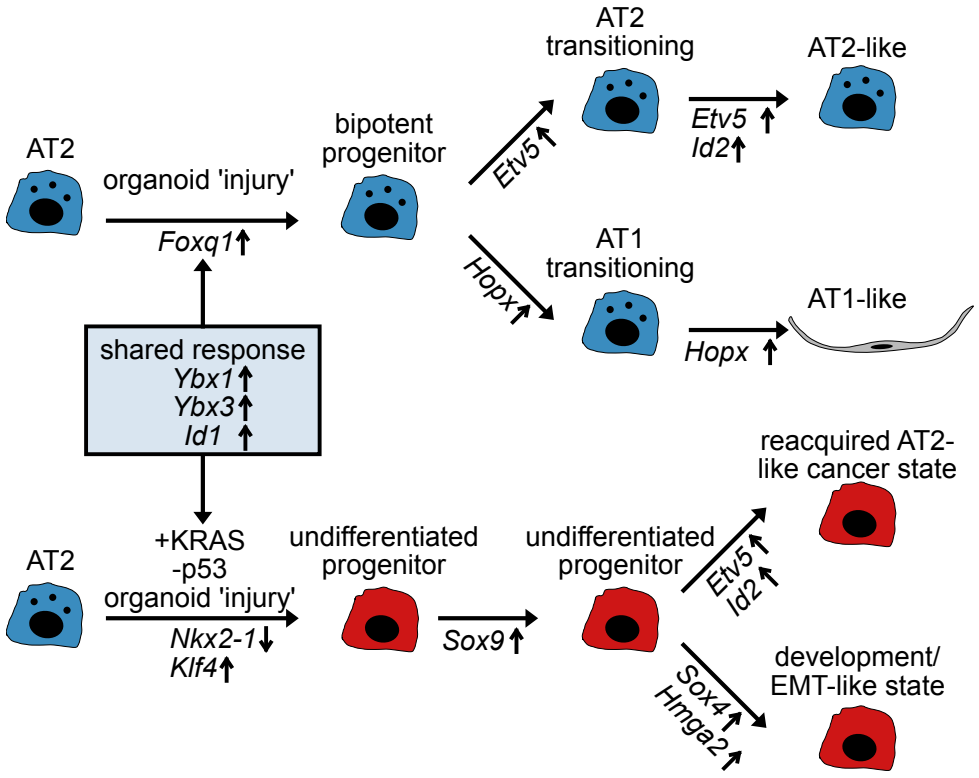

Figure 47: Working model for alveolar and cancer cell progression
This working model is based on the transcriptional time course analysis of Y-CRE and KPY-CRE organoids focused on the differential expression analysis of transcription factors. The arrows indicate if a gene is upregulated or downregulated. The upper part shows the cell transition in alveolar organoids; the lower part shows the transition in cancer organoids. In both cases, the initial cell state is a primary AT2 cell. After plating the AT2 cell in organoid culture, there is an initial shared injury response in both samples (box). However, already at this time point there are differences in the expression of TFs that lead to different transcriptional states. The control cells become bipotent progenitor cells that differentiate into AT2-like and AT1-like cells; in contrast, the cancer cells become dedifferentiated and transition to an AT2-like cancer state, and a development/ EMT-like state.

AT2-like program with increased *Etv5* and *Id2* levels. However, these cells did not express AT2 signature genes to the same extent as the control cells. Furthermore, they had high expression of *Sox9*, emphasizing their developmental-like cancer state. On the second path of the cancer cell progression, *Sox4* and *Hmga2* were upregulated, and the cells adapted a cell state distinct from any of the control states. They had high developmental gene expression, and expressed TFs associated with EMT and late-stage cancer. Overall, even though we observed parallels between the control and cancer cell progression, the transcriptional landscape and the cell trajectories were vastly different.

4.3.6 Supporting mesenchymal cells have distinct transcriptomes depending on the type of co-cultured epithelial cells

Because organs have complex environments, it is difficult to study the contributions of stromal cell populations on tumor progression *in vivo*. Organoids allow for the study of cell-cell-interactions in a simplified 3D environment. To test if the co-cultured MECCs interacted with cells in cancer organoid culture, I stained for the mesenchymal marker vimentin (Vim),

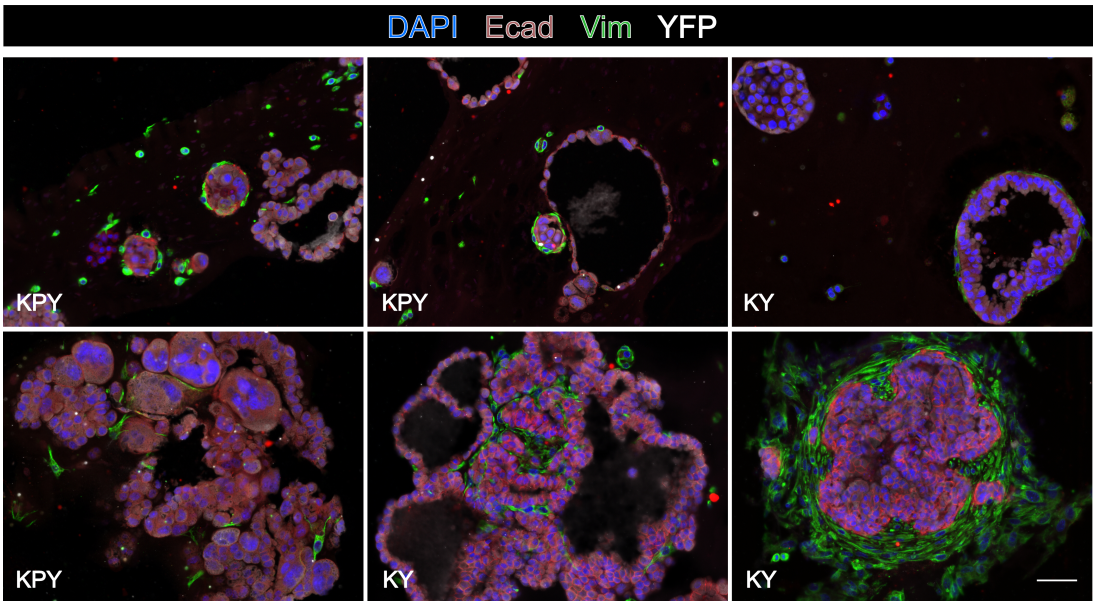

Figure 48: Staining for mesenchymal cells in cancer organoid cultures
Mesenchymal cells interact with organoids in various ways. Upper row: KY-CRE and KPY-CRE day 7 organoids. Lower row: KY-CRE and KPY-CRE day 14 organoids. Slides were stained for DAPI (nuclei), E-cadherin (Ecad, epithelial marker), vimentin (Vim, mesenchymal marker), YFP (Cre expression). Scale bar = 50 µm

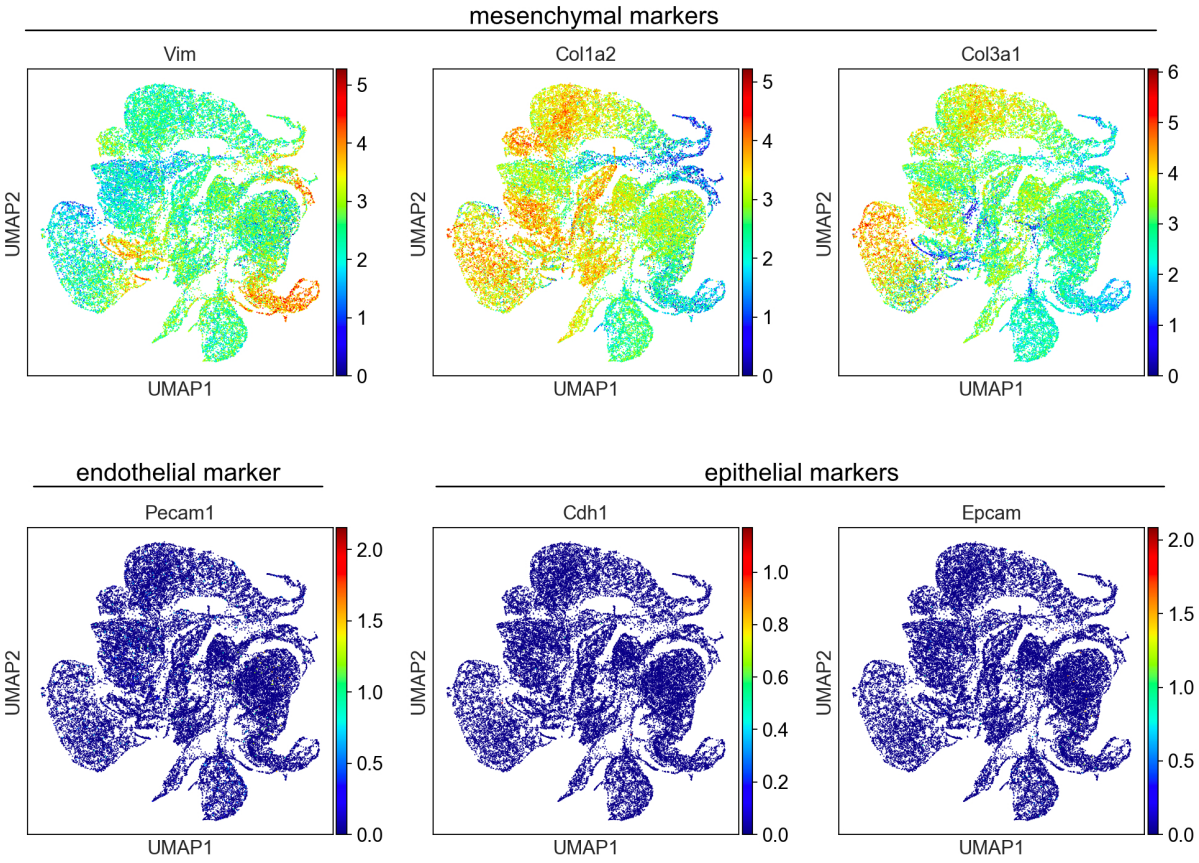

Figure 49: Cell type marker expression in mesenchymal co-cultured cells
UMAP representation of expression level of selected mesenchymal, endothelial, and epithelial markers in mesenchymal co-cutured cells.

the epithelial marker E-cadherin (Ecad), and YFP. Organoids displayed different levels of interactions with MECCs (figure 48). Some organoids were completely surrounded by MECCs, some had MECCs integrated into the organoid structure, while others did not seem to be closely associated. This observation emphasized the organoid heterogeneity when it comes to stromal cell interactions.

To test if MECCs changed their transcriptional landscape in organoid culture over time and if their transcriptomes were different when co-cultured with cancer cells, we analyzed the MECCs scRNA-Seq data separate from the epithelial data. First, I checked expression levels of commonly used mesenchymal markers (*Vim, Col1a2, Col3a1*), an endothelial marker (*Pecam1*), and epithelial markers (*Cdh1, Epcam*). Gene expression analysis confirmed that the co-cultured cells were indeed of mesenchymal nature (figure 49). Next, we used Louvain clustering to identify transcriptionally distinct clusters. Clusters from the MECCs dataset were indicated by a

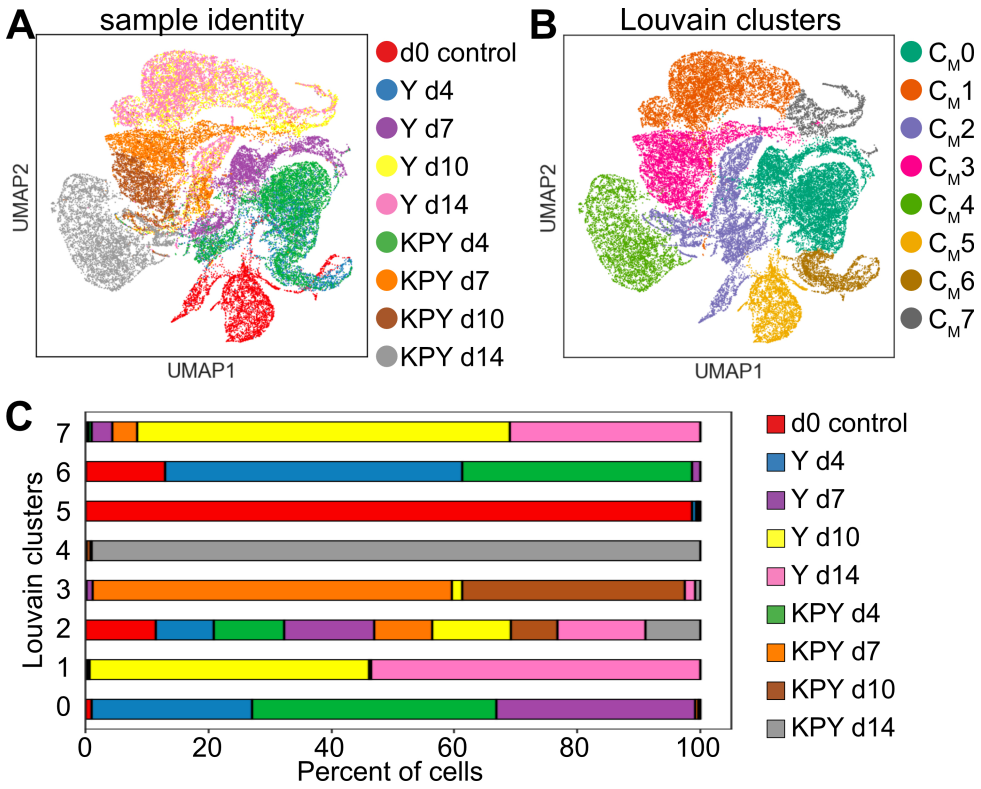

Figure 50: Clustering of mesenchymal cultured cells
Plots of mesenchymal cells co-cultured with Y-CRE or KPY-CRE epithelial cells. Samples were collected on days (d) 4-14 of organoid culture, or from 3D monoculture (d0 control). **(A)** UMAP plot colored by sample identity. **(B)** UMAP plot colored by Louvain clusters at a resolution of 0.5. 8 clusters were identified. **(C)** Percentage of cells per Louvain cluster colored by sample identity.

C_M nomenclature. Control day 0 MECCs represented cells that were kept in a 3D monoculture; they did not overlap with the other time points, indicating that the transcriptomes of MECCs cultured by themselves were distinct from MECCs that were in co-culture with epithelial cells (figure 50 A). Day 4 MECCs from Y and KPY samples mostly overlapped; the transcriptome of MECCs co-cultured with healthy epithelial cells was comparable to the ones co-cultured with cancer cells at this time point. Interestingly, at the later time points (day 7 to 14), the cells split into distinct clusters depending on their co-cultures (figure 50 A). This indicated that MECCs co-cultured with cancer cells had a different transcriptome compared to MECCs co-cultured with non-cancerous cells.

Using unsupervised clustering, we divided the data into eight Louvain clusters (figure 50 B). Interestingly, C_M2 comprised cells of every single sample, including the day 0 control time point, indicating that two different types of mesenchymal cells might be present in the MECCs (figure 50 C). In order to further delineate the different mesenchymal cell populations, we per-

formed DE analysis on all clusters (table 19, appendix). Genes associated with myofibroblasts such as Transgelin (*Tagln*), Smooth muscle actin alpha 2 (*Acta2*), and Myosin light chain 9 (*Myl9*) came up as highly differentially expressed in C_M2, identifying this cluster as a myofibroblast population (figure 51). To identify the other mesenchymal cell clusters present, I checked for genes that have been described in the literature as markers for different lung mesenchyme populations. The cells expressed only low levels of the pericyte markers Melanoma Cell Adhesion Molecule (*Mcam*), Chondroitin sulfate proteoglycan 4 (*Cspg4*), and Desmin (*Des*) (figure 51) (Xu *et al.*, 2017). Furthermore, no expression of Leucine-rich repeat-containing G-protein coupled receptor 5 and 6 (*Lgr5 and Lgr6*), and *Wnt2* was observed, markers of different mesenchyme subpopulations described in the literature (figure 51) (Zepp *et al.*, 2017; Lee *et al.*, 2017). Both platelet-derived growth factor receptor alpha and beta (Pdgrfa and Pdgrfb) were ubiquitously expressed (figure 51) (Zepp *et al.*, 2017). Interestingly, a small subset of cells that was part of C_M1, mainly comprised of day 10 and day 14 Y-CRE co-cultured cells, expressed *Axin2*, an important Wnt target gene that marks several distinct mesenchymal populations (figure 50 A and 51) (Zepp *et al.*, 2017). Even though some of the markers described in the literature were present in the MECCs dataset, it is difficult to make conclusions regarding the exact identity of the clusters. The MECCs used in this study were derived from neonatal mice and cultured *in vitro* for multiple passages, which makes a comparison to the adult mesenchymal subpopulations described in the lung *in vivo* challenging.

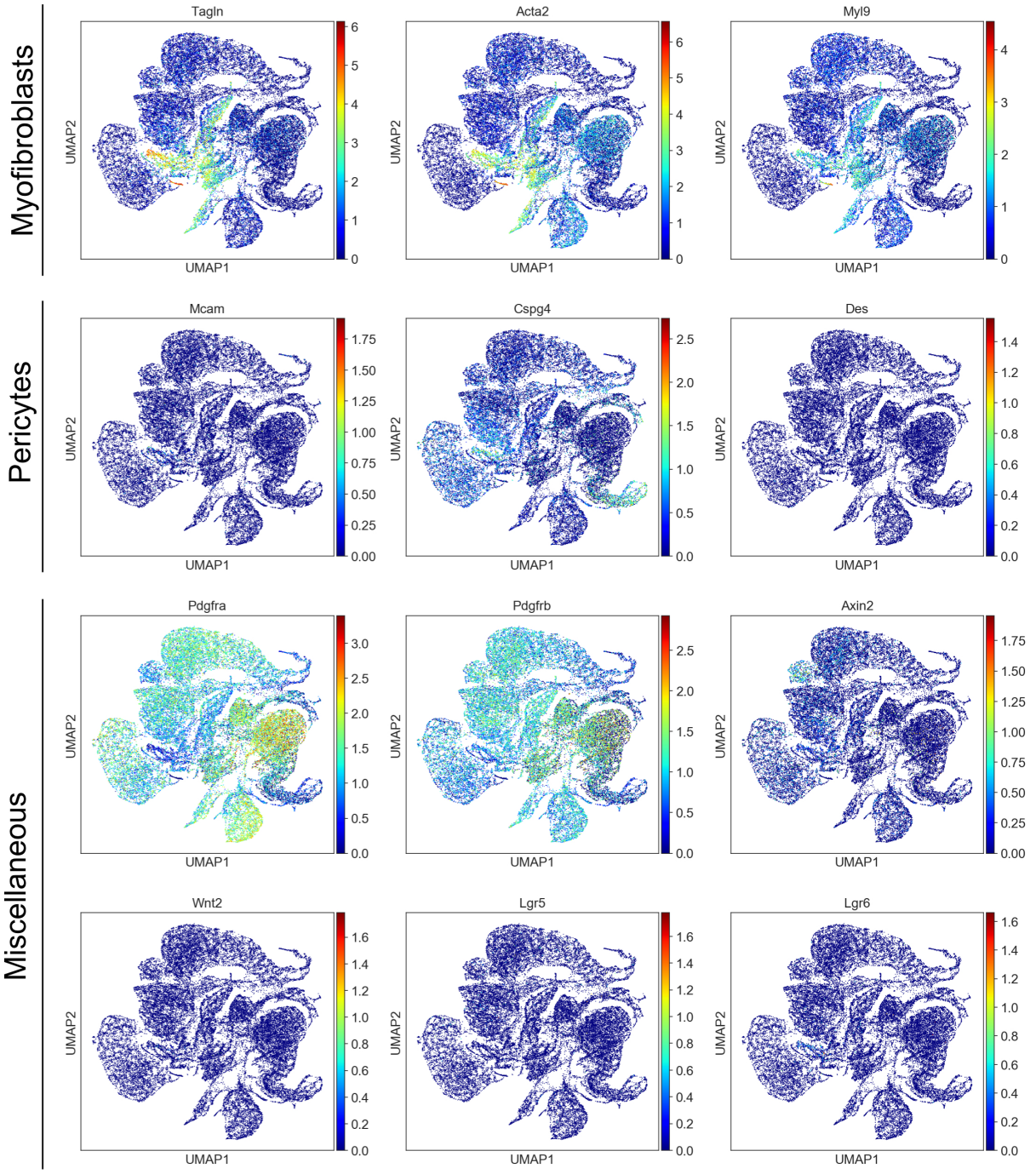

Figure 51: Marker expression of mesenchymal cell types
UMAP representation of expression level of selected mesenchymal markers, including markers for myofibroblasts, pericytes, and miscellaneous populations.

5 Discussion

KRAS is the most commonly mutated oncogene in epithelial cancers. Understanding how oncogenic KRAS alters cell states is crucial in order to develop new therapies. The LUAD organoid system introduced in this thesis presents a new tool to study KRAS-driven lung cancer *in vitro*. Instead of only causing increased proliferation, I showed that oncogenic KRAS changed the cellular state to a more dedifferentiated phenotype within a few days. Using transcriptional analysis with single cell resolution, I compared cancer cell progression to healthy counterparts, and provided TFs and genes that were upregulated during different stages of tumorigenesis.

5.1 Evaluation of the organoid system: research applications and limitations

The cancer organoid system presented in this thesis can be used for many different applications and to address a variety of questions. However, as with every tool there are limitations that need to be considered. In this section, I will discuss research areas that could be explored using this organoid system. I will point out the advantages, disadvantages, and the limitations of the system.

5.1.1 Organoid forming efficiency and the cell of origin for cancer

The organoids described in this thesis all came from the SCA1- population which is highly enriched in AT2 cells. A multitude of studies provided evidence that AT2 cells, the stem cells of the alveoli, are the cell of origin for LUAD (see section 1.2.5) (Mainardi *et al.*, 2014; Sutherland *et al.*, 2014; Xu *et al.*, 2012; Desai *et al.*, 2014; Sainz de Aja *et al.*, 2020). It should be noted that only a fraction of the isolated AT2 cells formed organoids both in the control and in the cancer setting. In previous work, it was shown that organoids are clonal, indicating that the number of organoids correlates with the number of cells with organoid forming potential present when plating (Lee *et al.*, 2014). A cell that can form an organoid has stem or progenitor cell properties, or is a cancer initiating cell or CSC (Wang *et al.*, 2019). In theory, a pure stem cell population would reach a OFE of close to 100%. Therefore, it can be concluded that indeed, not every AT2 cell has stem cell properties or the ability to initiate a tumor. There are multiple possible explanations for the observed low OFE in this organoid

system.

Firstly, it is possible that the tissue preparation, FACS, and virus infection caused significant cell death. Even though cell numbers plated referred to live cells after virus infection, apoptosis might have been induced after plating, so that only a fraction of cells survived and formed organoids. This could also explain why cells with *Tpr53* deletion, an important inducer of apoptosis, had a higher OFE. To rule out that a low OFE is due to cell death, a Rho-associated protein kinase (ROCK) inhibitor could be added to the culture in order to prevent apoptosis. Indeed, many AT2 cell culture protocols contain an anti-apoptotic compound (Weiner *et al.*, 2019; Katsura *et al.*, 2020).

A second possibility is that there are intrinsic differences within the AT2 population, and only a sub-population with increased stem or progenitor function is able to initiate organoid formation. Those differences could be subtle, or of epigenetic nature, which is not captured by scRNA-Seq. Indeed, in recent years subsets of AT2 cells with increased stem cell properties have been described, and it was speculated that those AT2 subpopulations might represent the cell of origin for LUAD (see section 1.1.2) (Nabhan *et al.*, 2018; Zacharias *et al.*, 2018). In order to test if enrichment for those subpopulations leads to a higher OFE in the control and cancer organoids, FACS sorting strategies developed by other labs could be tested. For instance, Zacharias *et al.* described a Wnt-responsive AT2 sub-population that can be isolated using the surface marker TM4SF1. It was found that this population had a higher OFE compared to the rest of the AT2 cells, indicating that it was enriched in AT2 stem cells (Zacharias *et al.*, 2018). To test if this sub-population also has a higher OFE compared to the bulk AT2 cells in a cancer setting, the same FACS strategy could be applied to KY and KPY mice and the cells could be transformed using the same *in vitro* infection strategy as presented here. This could provide evidence that progenitor-like subsets of AT2 cells are more prone to KRAS-driven transformation and give insight into the cell of origin of LUAD. Furthermore, it could be tested if the transformation of different AT2 subpopulations results in cancer cells with different properties, such as more invasive phenotypes, or cells more prone to proliferation or EMT

The third option is that stochastically, some AT2 cells are exposed to niches better suited to initiate an alveolar or cancer organoid. Niches in the organoid culture could be made up of the ECM scaffold, supporting MECCs, and other epithelial cells. Indeed, a revised CSC hypothesis states that each cell within a tumor can transition to a CSC in the presence of the right niche

signals (Batlle and Clevers, 2017). In order to assess this possibility, the niche within the organoid culture could be manipulated by using different ECM components, supporting cells, or by manipulating the supporting cells to produce signals shown to be relevant for AT2 and LUAD cancer cells, such as secreted Wnt pathway components (Nabhan *et al.*, 2018).

5.1.2 Modeling KRAS-driven early-stage lung cancer

Finding new treatments for early-stage cancer is of increasing importance due to ever improving detection methods that now allow for the identification of pre-cancerous and early-stage lesions (Heuvelmans *et al.*, 2017; De Koning *et al.*, 2020). Here, we used the cancer organoid system to study the transcriptional landscape of oncogenic KRAS-driven AT2 cells without any other mutational drivers, seven days after transformation. The organoids grew in a controlled and simplified microenvironment, lacking many external signals usually present *in vivo*. Therefore, we had the unique opportunity to study the intrinsic changes that occur in a primary epithelial cell as a result of oncogenic KRAS expression. We found that KRAS signaling did not simply increase proliferation early on, but that a more severe reprogramming of the cells occurred. We observed a striking loss of AT2 marker expression and an upregulation of developmental markers, indicating that KRAS signaling caused AT2 cells to transition to a more dedifferentiated state. The cells that appeared to be most dedifferentiated also displayed high Wnt signaling, indicating that the Wnt pathway might already play an important role early on in tumorigenesis. Furthermore, we found that genes from other epithelial cell lineages of the lung were upregulated. Both the loss of identity and lineage infidelity are processes known to occur at a later stage in lung cancer; however they have not been described for early-stage lesions thus far (Winslow *et al.*, 2011; Laughney *et al.*, 2020). To ensure that this observation is not an artifact of organoid culture, we confirmed our findings with *in vivo* data. Although a similar trend could be observed in the KY GEMM and in human stage IA patient samples, the phenotype was more pronounced in the cancer organoids (Dost *et al.*, 2020). There are multiple plausible explanations for this discrepancy.

Firstly, it is possible that the organoid setting allows for the enrichment of CSCs. Similar to stem cells, CSCs can self-renew and are therefore important for the long-term sustenance of the tumor. They are thought to be resistant to many cancer therapies and can cause relapse locally or at distant metastatic sites even years after treatment of the primary tumor. CSC are often rare subpopulations of cells, but can sometimes make up the bulk of the tumor (Clevers,

2011; Batlle and Clevers, 2017). Organoids are formed by stem or progenitor cells of normal tissue; therefore, it is conceivable that in a cancer setting, only CSCs are capable of forming and propagating cancer organoids. This might provide a unique opportunity to study the biology of CSCs and to find vulnerabilities of this difficult to treat cell population. The gold standard to test for the presence of CSC populations are serial transplantation assays *in vivo*. In future, it could be tested if transplanted cancer organoids can give rise to secondary and tertiary tumors. If this is the case, it would be evidence that CSCs are present in the organoids. Genes found to be upregulated in cancer organoids would therefore provide potential new therapy targets and could be examined experimentally using knock out studies, to test if tumor formation was ablated.

Secondly, it is possible that the timeline of cancer progression in organoid culture compared to *in vivo* is accelerated. Cancer cells in organoid culture have more space to expand faster than transformed cells inside of an epithelium. Because of the simplified microenvironment, they might lack inhibitory signals from other epithelial or stromal cell types. Most notably, there are no immune cells in organoid cultures that could recognize and eliminate transformed cells and keep their growth in check. Further studies will be necessary to precisely determine the alignment of organoid and *in vivo* early LUAD development timelines, and to what extent the cancer organoids recapitulate early-stage LUAD. One approach could be to compare the rate of cell division in the organoid culture to early-stage *in vivo* tumors.

5.1.3 The study of epithelial to mesenchymal transition and metastasis

Treatment of cancer that has metastasized is one of the most challenging parts of oncology. Therefore, the prevention and treatment of metastasis rightly receives a lot of attention. However, the study of metastasis *in vivo* is challenging and cell lines lack a physiologically relevant environment. Therefore, organoids might provide an *in vitro* tool to study metastasis-relevant processes, such as EMT (Mittal, 2018). When transplanted orthotopically, both KY-CRE and KPY-CRE derived cancer organoid cells formed tumors. However, we used a lung injury model and immunocompromised mice to facilitate the engraftments of the cells. In future, it should be tried to transplant the cancer organoids into the lungs of syngenic mice without injury. If tumor seeding is successful in the lung, it would provide evidence that the cancer organoids contain cells that are capable of spontaneous metastasis, describing the process of forming a tumor in the tissue of origin of the cancer cells (Price, 2014). In contrast,

models for experimental metastasis test the ability of tumor cells to extravasate, the process of leaving the circulatory system to enter a tissue, and to form tumors subsequently (Price, 2014). EMT is a cell feature often connected to metastases (Yang and Weinberg, 2008). We showed that important EMT-driving TFs such as *Sox4* were upregulated in the cancer organoids (Lourenço and Coffer, 2017). Furthermore, the mesenchymal marker vimentin was detected at the protein level in organoid cells, indicating that single cells were transitioning to a more mesenchymal state in culture. One experiment often used to test if cells have metastatic potential is the tail vein assay (Elkin and Vlodavsky, 2001). In order to do this assay, cancer organoid-derived single cell suspensions could be injected into the tail vein of an immunocompromised mouse. If the cells survive the blood stream and are capable of extravasion, they will form tumors at distant sites, often in the lungs.

If the proposed experiments show that the organoids indeed contain cells capable of metastasis, the organoid system would provide a convenient platform to further study this process. EMT was observed after only 14 days in KPY-CRE organoid culture, compared to 30 weeks in the KP mouse model, reducing the time of experiments to study EMT dramatically (Marjanovic *et al.*, 2020). Furthermore, manipulation of the organoids is easier than manipulation of GEMMs and could be followed by one of the previously mentioned metastasis assays to obtain *in vivo* validation.

5.1.4 Studying the mutational landscape of lung adenocarcinoma

Lung tumors have high numbers of somatic mutations (Wheeler and Wang, 2013). Defining the role of driver and passenger mutations on cell signaling and therapy success is one of the key challenges of lung cancer oncology. Therefore, it is crucial to understand the molecular changes that follow mutations in key oncogenes and tumor suppressors in a controlled manner, using GEMMs or *in vitro* systems. In the KP mouse model, protein-altering mutations are infrequent; however, copy number alterations are common (Chung *et al.*, 2017). This raises the question if mutations or other genetic alterations accumulate in the cancer organoid model, and how this process shapes the cancer cell heterogeneity. Even within a very short time frame of 14 days, cancerous AT2 cells underwent significant transcriptional and phenotypical changes. Especially in the KPY-CRE organoids, cells became pleomorphic and some cells appeared to be multinucleated. To test if genetic alterations occur in the organoid model, whole genome sequencing could be performed on organoid-derived cells. This study could give insights into

the process by which genetic alterations are acquired in KRAS-driven cancers, and improve our understanding about the connection between cancer drivers and cancer cell heterogeneity.

5.1.5 Improving the versatility of the organoid system

There are many applications and advantages of the organoid system presented here. However, there are ways to further increase the versatility of the tool. For instance, the focus of this work was on modeling and analyzing LUAD progression initiated by oncogenic KRAS G12D expression. However, even though G12D is the most common variant in never-smokers, G12C and G12V are more common in patients (Dogan *et al.*, 2012; Haigis, 2017). Because GEMMs of all three mentioned variants exists, the organoid system could be used to study the differences between the variants, and to directly compare therapy strategies.

Another way to improve the organoid system as a tool is to add temporal control over the genetic alterations, so that KRAS mutations act as the cancer initiating event, and loss of *Trp53* as a later event, as often observed in patients (Chen *et al.*, 2019; Nakanishi *et al.*, 2009). Temporal control in this system could be achieved by combining the Cre/loxP with another system such as Flp/FRT or Dre/rox, ideally with an inducible recombinase (Rodríguez *et al.*, 2000; Anastassiadis *et al.*, 2009). With such a system it would be possible to first grow alveolar organoids, then induce oncogenic KRAS only in a fraction of cells. This would more closely recapitulate *in vivo* tumor growth and would enable the study of cell-cell-interactions between healthy and transformed epithelial cells.

However, even with these improvements there are limitations that need to be considered. The cancer organoid system as presented in this thesis requires a Cre inducible GEMM such as the KY or KPY models used here. This has multiple advantages: Cre inducible GEMMs allow for the use of a Cre-carrying adenovirus to infect the cells. In contrast to a lentivirus, adenoviruses do not integrate into the cell genome, allowing for transient infection without causing off-target effects. Furthermore, infection efficiency with adenoviruses is usually high, rendering cell selection unnecessary. This makes it possible to study cell state alterations that occur early after the induction of an oncogene, such as KRAS. However, being dependent on already existing GEMMs limits the variety of mutations that can be studied, or would require lengthy mouse model generation to study different genotypes of interest. Using techniques such as CRISPR/Cas9 allows for the introduction or knock out of any mutation or gene and

therefore provides a more versatile tool. Of course, there are also drawbacks to this approach. Introducing a large protein such as Cas9 into cells can be challenging and the process is often inefficient, making it necessary to add selection steps in order to obtain the desired genotype (Drost *et al.*, 2015; Matano *et al.*, 2015). Hence, there is a trade-off between versatility and practicality, and the best system will depend on the research question studied.

5.2 Molecular mechanisms of alveolar and cancer organoid growth

In this thesis, I used the organoid model for extensive study of the transcriptional landscape of healthy alveolar organoids and their cancerous counterparts. Using pseudotime analysis and a time course, I described the molecular changes in alveolar regeneration and KRAS-driven tumorigenesis.

5.2.1 Organoids as a model for regeneration

The AT2-derived control organoids used in this study showed a clear bifurcating trajectory in the transcriptional time course data, representing AT2 and AT1 trajectories. Even though it was known that AT2 cells differentiate to AT1 cells in organoid culture, this is the first scRNA-Seq time course analysis of this process. A surprising finding was that at the earliest time point (day 4), AT2 cells had a low AT2 signature that was reacquired along the AT2 trajectory. The exact identity of the day 4 root cells is therefore not clear. In a recent paper, Choi *et al.* presented single cell transcriptional data of comparable organoid cultures from a late time point (day 21) and only detected two populations, AT2 cells and AT1 cells (Choi *et al.*, 2020). However, in the same paper they described different AT2 cell states that only occurred after lung injury *in vivo*. Those "primed AT2" cells and "damage-associated transient progenitors" (DATPs) had lower expression of AT2 marker genes and were shown to give rise to AT1 cells. This is reminiscent of the findings in this thesis. It is possible that the organoid culture is comparable to an injury model. Stem cells are taken from their native environments and put into an ECM scaffold with sparse supporting mesenchymal cells. Not sensing other epithelial cells in close proximity likely causes an injury response in the stem cells that is aimed at regenerating the epithelium (Nabhan *et al.*, 2018). Therefore, it is plausible that the root cells represent a regenerating AT2 cell state that is similar to the DATPs observed *in vivo* (Choi *et al.*, 2020). During regeneration, the cells then split into the AT1 trajectory, and into

an AT2 trajectory that is more representative of a homeostatic AT2 cell state.

The concept of organoids as a model for regeneration could also explain the shared response of control and cancer organoids early in organoid culture. As summarized in figure 47, many TFs changed expression levels in both day 4 samples, which was accompanied by a reduction in AT2 signature genes. One of those TFs was *Id1*. Id proteins are upregulated in multiple types of cancer and have been described as oncogenes and tumor markers because of their role in cellular processes connected to tumorigenesis (Hasskarl and Munger, 2002; Lyabin *et al.*, 2014). They inhibit differentiation and induce proliferation through various mechanisms. For instance, Id proteins have been shown to inhibit nucleoprotein complex formation by binding TFs that contain an ETS domain, such as ETV5 (Yates *et al.*, 1999). Since ETV5 is an important TF for AT2 identity, it can be speculated that the reduction in AT2 marker gene expression is a result of the interactions between ID1 and ETV5. Additional shared upregulated genes in day 4 organoid cultures were *Ybx1* (YB1) and *Ybx3*. YB proteins are involved in many cellular processes, including proliferation (Lyabin *et al.*, 2014). In lung cancer, YB1 was found to be overexpressed in human NSCLC lesions and correlated with poor prognosis (Zhao *et al.*, 2016; Zheng *et al.*, 2019). Our data suggest that it might also play a role in AT2 regeneration. Delineating the differences between normal regeneration responses and early tumorigenesis will be important to understand unique features of early-stage cancer cells better in future.

5.2.2 Early loss of AT2 identity as an advantage for cancer cells

Even though there was a striking overlap of genes upregulated in early organoid formation, the control and cancer cells did not cluster together, indicating that there were substantial transcriptional differences. For one, *Nkx2-1* was only downregulated in the cancer organoids, indicating that there was a more severe loss of AT2 identity in the cancer setting compared to the control. This might explain why the cancer cells appeared more dedifferentiated when directly compared to the control cells. However, if oncogenic KRAS caused downregulation of *Nkx2-1* directly or indirectly still needs to be determined. In contrast, a TF found upregulated in the cancer organoids was *Klf4*. *Klf4* became famous as one of the factors used to induce pluripotency in differentiated cells (Takahashi and Yamanaka, 2006). In cancer, *Klf4* has both been described as an oncogene and tumor suppressor, indicating its complex functions in tissue-specific contexts (Ghaleb and Yang, 2017). In lung cancer, *Klf4* was generally found to be decreased in tumors compared to normal tissue (Hu *et al.*, 2009; Fadous-Khalifé *et al.*,

2016). However, normal tissue analyzed in patients is usually homeostatic and contains all cell types present in the biopsy. A direct comparison between *Klf4* expression levels in normal and transformed AT2 cells has not been done. In a K; *Klf4*$^{fl/fl}$ mouse model, it was shown that the lung tumor burden was higher compared to the K model (Yu *et al.*, 2016). However, this experiment only determined the role of *Klf4* in tumor initiation in the K model and it did not assess if *Klf4* deletion had an effect on tumor stages. The role of *Klf4* in early tumor progression, especially in a *Trp53* deficient background, has not been assessed yet. Therefore, if *Klf4* plays a functional role in the observed dedifferentiation phenotype needs to be determined by further functional studies.

There are multiple explanations why a cancer cell would downregulate AT2 markers early after transformation. It could simply be that in order to proliferate at a higher rate, the cells cease to transcribe genes connected to a differentiated phenotype. Even though this could also explain why the initial response of the control cells showed the same trend, it does not explain the extent of the differences between control and cancer organoids. For instance, at day 7 of organoid culture close to 100% of the control cells stained positive for SPC, while less than 20% expressed the protein in the cancer organoids, despite comparable numbers of proliferating cells. Another explanation is that the dedifferentiated phenotype allows for more plasticity of the cancer cells. From a less differentiated state, cell subpopulations can turn on different transcriptional programs, resulting in lineage infidelity, as was observed in the cancer organoid system and also in human LUAD (Laughney *et al.*, 2020). A higher diversity of subclones increases the likelihood that the tumor contains cells with immune evasive properties, chemo-resistance, metastatic potential, or other CSC properties. The reason for the high incidence of KRAS mutations in cancer could therefore be its ability to reprogram the cell of origin quickly, creating enough cell diversity to evade immune-surveillance, while at the same time containing clones with high proliferative potential.

5.2.3 Comparison of alveolar regeneration and cancer progression

Some similarities between the control alveolar organoids and the cancer organoids go beyond the day 4 time point. After the initial downregulation of AT2 genes, an AT2 signature was reacquired by a subset of cells in both samples over time, but more strongly in the control organoids. The control cells likely followed a regeneration response, recovering the depleted AT1 and AT2 cells; in contrast, the cancer cells modeled tumor progression. Comparing

regeneration to tumorigenesis in this organoid model may have important advantages. It is a long-held notion that tumors are wounds that do not heal; indeed, cancer and injury response have striking parallels (Dvorak, 1986; Arwert *et al.*, 2012). In both cases, cells increase proliferation at the expense of differentiation. However, while the regeneration response is transient, cancers often hijack the same signals to expand in an unchecked manner, sometimes indefinitely (Ge *et al.*, 2017). In the lung, EGFR-KRAS signaling has been shown to be important during AT2 renewal; both genes are also the most common oncogenic drivers in LUAD (Desai *et al.*, 2014; Collisson *et al.*, 2014). The observation that the majority of KRAS G12D expressing AT2 cells reacquired AT2 markers transcriptionally, and SPC on protein level, could explain why human LUAD generally stains positive for AT2 markers (Jin *et al.*, 2018; Ye *et al.*, 2011; Zhan *et al.*, 2015). Because the very early steps of the transformation process usually occur before detection of the tumor, it is not well understood if the initial cellular response is a loss of those same markers. We were able to show only recently that in human LUAD, AT2 markers were indeed downregulated transcriptionally at an early stage (Dost *et al.*, 2020). One reason for the partial reacquisition of the AT2 program could be that the cells maintained their AT2 memory after the initial downregulation of the AT2 program, and entered a regeneration response to replenish AT2 cells. Another possibility is that they follow their developmental trajectories, as has been proposed for human LUAD progression (Laughney *et al.*, 2020). As outlined in figure 2, *Sox9* is an important driver of the distal fate trajectories during lung development and marks the bipotent progenitor cells that give rise to AT2 and AT1 cells. Instead of completing their trajectories, it is conceivable that the cells get stuck at this stage.

Apart from the re-acquisition of AT2 markers, the transcriptional trajectory of the cancer cells differed greatly from the control cells. In the cancer organoids, the lung developmental TF *Sox9* was expressed by the majority of cells after day 7, and the AT1 trajectory was completely lost, indicating a lasting reprogramming of the cells. Furthermore, the lung TF *Hmga2* and the EMT factor *Sox4* were gained in a subset of cells, indicating ongoing dedifferentiation to a more mesenchymal cell state, as was confirmed by stainings. The presence of AT2-like cancer cells and developmental-like or EMT-like cancer cells highlights the heterogeneity of tumors and is in line with published *in vivo* data (Winslow *et al.*, 2011; Marjanovic *et al.*, 2020; Laughney *et al.*, 2020; LaFave *et al.*, 2020). However, the emergence of such cell states, like the HPCS population described by Marjanovic *et al.* (2020) (see section 1.2.6) usually takes weeks to emerge *in vivo*, compared to only days in organoid culture. If a population with increased CSC features exists in the organoid cultures needs yet to be tested experimentally. The CSC

hypothesis describes a tumor cell hierarchy similar to stem cells in normal tissue (Clevers, 2011). CSC are often less proliferative than the bulk of the tumor, but can replenish the whole tumor after ablation of the bulk cells by chemotherapy or other treatments. Therefore, it is possible that a similar hierarchy exists in the cancer organoids. To test this hypothesis, the transcriptional data could be used to identify cell surface receptors that are specific for either the AT2-like or the developmental-like cancer cells. Those markers could then be used to isolate the respective cell clusters using FACS sorting. If a population is enriched in CSCs, it should have a higher OFE and give rise to all the subtypes of cells observed in the parent organoid generation.

5.2.4 The role of tumor stromal cells on organoid progression

The tumor microenvironment, or tumor stroma, is an important driving factor of tumorigenesis (Micke and Östman, 2005; Pietras and Östman, 2010). It contains non-neoplastic hematopoietic, endothelial, and mesenchymal cells that provide signals to the cancer cells or remodel the ECM. Especially immune cells play an important role in shaping cancer cell niches. For instance, the concept of cancer cell attraction to specific tissue sites is a long-held notion also referred to as inflammatory oncotaxis (DerHagopian *et al.*, 1978). In this non-random process, metastatic cells establish colonies in specific niches, which are often created by inflammation-causing immune cells (Mendoza and Khanna, 2009). Another abundant cell type of the tumor stroma are cancer-associated fibroblasts (CAFs). CAFs are distinct from their fibroblast counterparts by their perpetually activated state and their expression of myofibroblast markers (Sugimoto *et al.*, 2006; Micke and Östman, 2004). In lung cancer, multiple studies have looked at the prognostic value of CAFs. While studies that relied on a single marker for staining of tissue slides were often inconclusive, studies that used more comprehensive gene expression profiling showed that their CAF gene signatures had prognostic value (Paulsson and Micke, 2014). Furthermore, *in vitro* and *in vivo* studies have connected CAF activity to altered lung cancer cell metabolism, enhanced chemoresistance, and increased cancer cell plasticity (Cruz-Bermúdez *et al.*, 2019; Shintani *et al.*, 2016; Xiang *et al.*, 2020; Sung *et al.*, 2019; Alguacil-Núñez *et al.*, 2018). While targeting CAFs in cancer therapy sounds like a viable option in theory, it has proven to be a difficult task and clinical therapies are yet to be developed (Chen and Song, 2019).

The supporting stromal cells used in this thesis were of mesenchymal nature. We revealed that

they split into two populations when cultured by themselves, one of them being transcriptionally similar to myofibroblasts. Identifying the other population remains challenging. In an attempt to annotate the cells, I analyzed expression levels of marker genes published in the literature. MECCs did not express pericyte markers, or *Lgr5* and *Lgr6*, markers of distinct mesenchymal subpopulations in the murine lung (Xu *et al.*, 2017; Lee *et al.*, 2017). The majority of cells expressed *Pdgfra* and *Pdgfrb*, described to mark mesenchymal alveolar niche cells (MANCs) and *Axin2* myofibrogenic progenitor (AMP) cells, respectively (Zepp *et al.*, 2017). However, both of these cell populations have been reported to co-express *Axin2*, which was not expressed in the control MECCs cells in our dataset. Most of the markers described in the literature mark distinct mesenchyme populations in the adult murine lung. However, the mesenchymal cells used in the organoid co-cultures were isolated from neonatal mice, then cultured in 2D conditions for multiple passages. It is likely that the gene-expression profile of the cells does not match with the adult mesenchymal cells, making the annotation of the cells difficult. Nevertheless, we were able to show that the supporting cells co-cultured with cancer cells became transcriptionally distinct from the cells co-cultured with healthy AT2 cells over time. In future experiments, it needs to be determined if this observed change is due to more crowded wells in the cancer organoid setting or by direct alteration of the non-cancerous mesenchyme by secreted factors from cancerous AT2 cells. To achieve this, fewer cells should be seeded for the cancer organoids, so that the density of cells in the well is comparable to the control, also at later time points. If the transcriptional differences still hold up, an in-depth analysis of the potential interactions between the epithelial cells and stromal cells should be conducted. For instance, factors secreted by MECCs could be identified and blocked to test if they influence the way cancer organoids grow. Computational tools, such as CellPhoneDB, can be used to analyse cell-cell interactions mediated by receptor-ligand complexes (Efremova *et al.*, 2020). Briefly, the scRNA-Seq datasets of the control epithelial and mesenchymal cells, and the cancer epithelial and mesenchymal cells can be analyzed for the expression of receptor-ligand pairs. If for instance the cancer cells start expressing a receptor not present on the control cells, and the matching ligand is expressed by the mesenchymal cells, it could be evidence that new communication occurs between the cancer cells and their MECCs. The communication networks present in the cancer setting, but not in the control setting, could provide new therapeutic targets specific to the tumor stroma.

5.2.5 Ephrin receptors as targets for LUAD

One receptor identified as upregulated in the cancer organoid cells was the ephrin receptor *Epha2*. One of the many advantages of organoid systems compared to *in vivo* studies is that they provide a convenient platform to test small molecules. I found that inhibition of the ephrin receptor family ablated organoid formation at nanomolar concentrations of the inhibitor, while having no effect on the control cells. Ephrin receptors and their ligands are overexpressed or downregulated in a multitude of cancer types (Pasquale, 2010). Their signaling is complex and their role in cancer is puzzling and often contradictory. For instance, EPHA2 is frequently overexpressed and correlates with poor prognosis in NSCLC (Brannan *et al.*, 2009a,b). However, a study that was focused on stage I NSCLC found that high expression of EPHA2 and of its ligand ephrin-A1 predicted favorable outcome of cancer patients (Ishikawa *et al.*, 2012). Indeed, the interaction between ephrin receptors and their ligands is complex and can have both positive and negative effects on tumor progression (Ieguchi and Maru, 2019). While ephrin receptors do not pose a completely new target for LUAD, the identification of *Epha2* in the transcriptional data is evidence that the cancer organoids are valuable tools in finding relevant targets for cancer therapy. A commonly used small molecule inhibitor corroborated the relevance of ephrin receptors in cancer organoid progression, while having no effect on the control cells. In future assays, it should be confirmed that the observed phenotype is not due to off-target effects of the inhibitor. One way to show this is to use other compounds that target the receptor, such as the ephrin-A1 ligand or an EPHA2 agonist such as doxazosin mesylate (Petty *et al.*, 2012). A caveat of this approach is that it is possible that compounds added to the media will also have an effect on the MECCs population. One way to avoid this is to knock out *Epha2* only in the epithelial cells using shRNAs or CRISPR/Cas9 technology. Furthermore, it should be assessed which cell population produces the ligands that bind the receptor. If the same phenotype is observed when knocking out the potential ligands in the MECCs population, it would provide evidence that signaling from the stromal cells is crucial for cancer oganoid progression. The organoid system provides a tool to delineate the cell communication networks and to further study the complex roles that the ephrin receptors and their ligands play in tumorigenesis.

5.3 Outlook

New organoid culturing methods and improvements in single cell omics technology are advancing quickly and will facilitate the study of cancer biology in the years to come. The work presented in this thesis has the potential to be part of future discoveries that will ultimately lead to better cancer treatments. One approach to finding new therapies, using the herein described organoid system, is to set up a screen for new targets in the organoid culture. The transcriptional anaysis presented in this thesis provided a multitude of genes that were upregulated in cancer cells, many of which can be targeted with small molecules. Using the 96-well transwell setting, a small drug screen could be set up to test for new potential treatments for LUAD. In addition, a CRISPR/Cas9 screen could give further insight into vulnerabilities of the cancer cells.

The combination of organoid technology and scRNA-Seq provided insights into the transcriptional states of AT2 cells following KRAS G12D expression. Even though the cancer organoids were all generated from AT2 cells, heterogeneity was observed already after a few days in culture. One of the outstanding questions in cancer biology is how cancer cell heterogeneity arises, and the organoid system provides a tool to study this process in a controlled environment. In future, the transcriptional data could be complemented with other recently established single cell technologies. Multimodal single cell measurements allow for the profiling of more than one aspect of a cell state with single cell resolution, for instance mRNA and DNA sequences (Stuart and Satija, 2019). Integrating single cell DNA-Sequencing (scDNA-Seq) and scRNA-Seq could provide insights into how genetic alterations shape the transcriptional landscape of cells and cause cancer cell heterogeneity (Macaulay *et al.*, 2015). Similarly, single cell combinatorial indexing chromatin accessibility and mRNA (sci-CAR) combines epigenetic profiling with scRNA-Seq and could give insights into the role of epigenetics in cancer progression (Cao *et al.*, 2018). To get a better understanding of how different cancer cell lineages evolve, single cell lineage tracing technologies can be integrated with the transcriptional analysis (Raj *et al.*, 2018; Alemany *et al.*, 2018; Spanjaard *et al.*, 2018). In order to better understand the interactions between cancer cells and stromal cells, single cell spatial analysis combined with scRNA-Seq could provide insights into the spatial relationships between cells, including the mesenchymal cells co-cultured with the cancer organoids (Wang *et al.*, 2018a; Moffitt *et al.*, 2018).

In future, slowly increasing the complexity of the simplified organoid system by adding other

stromal cell types, in combination with the single cell analyses introduced above, will provide further valuable information about cell-cell interactions between cancer cells and other cell types. For instance, adding different types of immune cells will enable the study of immune evasion mechanisms in KRAS-driven cancers, and the role of inflammation in tumor progression (Gomes *et al.*, 2014; Muenst *et al.*, 2016). The presence of endothelial cells in the culture could provide insights into the angiogenesis process, or the role of endothelial cell signaling in EMT (Ramamoorthy *et al.*, 2019).

Eventually, a similar system as presented here could be established using human cells. However, human data is often difficult to interpret. There are many genomic variations in the human genome, such as single nucleotide polymorphisms (SNPs) and other genetic alterations. Therefore, it is best to start with a simpler model, such as an inbred mouse strain as used here, to then slowly build up the complexity of the system.

This is indeed an exciting time for cancer research. New tools such as organoid cultures in combination with ever improving single cell technologies now provide new platforms to tackle old questions regarding cancer biology. Eventually, our increasing understanding of the molecular changes in cancer cells will lead to new therapy discoveries, with the ultimate goal to not treat, but to cure cancer.

6 Summary

6.1 Summary (English)

Lung cancer is the leading cause of cancer-related death word-wide. To date, there is no approved targeted therapy for lung cancer driven by KRAS, the most frequently altered oncogene in epithelial cancers. Despite the high incidence of KRAS-driven lung adenocarcinoma (LUAD), the underlying mechanisms that cause tumor progression are not well understood. Especially the early steps of tumorigenesis have not been studied extensively, as the research focus has been on advanced stage cancer. With early detection methods improving, it will become more important to understand the molecular changes directly following oncogenic KRAS expression, to develop targeted therapies particularly for KRAS-driven early-stage LUAD. New technologies, such as organoid cultures and single cell RNA-Sequencing, now facilitate the modeling and study of tumorigenesis in primary cells.

For this reason, I developed an organoid system to faithfully model LUAD *in vitro*. I showed that transformed alveolar type 2 (AT2) cell-derived cancer organoids recapitulated LUAD progression histologically. When transplanted orthotopically, the cancer organoids gave rise to tumors *in vivo*. I characterized the transcriptional landscape of oncogenic KRAS expressing cells early after transformation. Most notably, I found that oncogenic KRAS alone was sufficient to reprogram the cells to a more dedifferentiated phenotype, with reduced AT2 identity and increased expression of development genes. Using the transcriptional data, I identified and confirmed the ephrin receptor *Epha2* as a potential target for early-stage LUAD.

To further characterize the molecular changes that occur during tumor progression, I performed a time course analysis of the transcriptional landscape of alveolar organoids and cancer organoids with single cell resolution. I found that the alveolar organoids followed a regeneration response; the cells transitioned to alveolar type 1 (AT1) and AT2 cells in culture, similar to their differentiation trajectories *in vivo*. In contrast, cancer organoid cells had lost the alveolar differentiation trajectories and instead expressed transcription factors connected to pluripotency and development, early after oncogene initiation. Even though the cancer cells partially recovered AT2 identity markers at a later time point, they remained in a dedifferentiated state.

Overall, I developed a new organoid tool that can be used for an array of applications. I characterized the system and explored the transcriptional changes that occured during cancer progression. I showed that oncogenic KRAS expression caused a lasting dedifferentiation response early after initiation. Delineating the molecular mechanisms that lead to this response will help understand KRAS biology, and provide new avenues for targeted treatments.

6.2 Zusammenfassung (Deutsch)

Lungenkrebs ist weltweit die häufigste krebsbedingte Todesursache. Bisher gibt es keine zugelassenen zielgerichteten Therapien für Lungenkrebs, der durch KRAS verursacht wird, dem am häufigsten veränderten Onkogen in Karzinomen. Trotz der hohen Inzidenz von KRAS-gesteuerten Lungenadenokarzinomen ist unser Verständnis der zugrunde liegenden Mechanismen, die Tumorprogression verursachen, bescheiden. Insbesondere die ersten Schritte der Tumorentstehung wurden nicht ausführlich untersucht, da der Forschungsschwerpunkt bislang auf den fortgeschrittenen Krebsstadien lag. Auf Grund der sich verbessernden Früherkennungsmethoden wird es immer wichtiger, auch die molekularen Veränderungen direkt nach der onkogenen KRAS-Expression zu verstehen und neue zielgerichtete Therapien vorallem für die frühen Krebsstadien zu entwickeln. Neue Technologien, wie Organoidkulturen und Einzelzell-RNA-Sequenzierung, können nun die Modellierung und Untersuchung des Fortschreitens von Krebs in Primärzellen erleichtern.

Aus diesem Grund habe ich ein Organoid-System entwickelt, das die Lungenadenokarzinom-Progression modelliert. Ich konnte zeigen, dass Krebsorganoide, die aus transformierten alveolären Typ 2 (AT2) Zellen hergestellt wurden, die Tumorprogression histologisch nachahmten. Bei orthotopischer Transplantation führten die Krebsorganoide zu Tumoren *in vivo*. Ich charakterisierte die Transkriptionslandschaft von Zellen, früh nach der Transformation durch onkogenes KRAS. Insbesondere fand ich heraus, dass onkogenes KRAS allein ausreichte, um die Zellen auf einen dedifferenzierten Phänotyp mit reduzierter AT2-Identität und erhöhter Expression von Entwicklungsgenen umzuprogrammieren. Unter Verwendung der Transkriptionsdaten identifizierte und bestätigte ich den Ephrinrezeptor *Epha2* als potenzielles Ziel für Lungenadenokarzinome im Frühstadium.

Um die molekularen Veränderungen, die während der Tumorprogression auftreten, weiter zu charakterisieren, führte ich eine Zeitverlaufsanalyse der Transkriptionslandschaft von Alveolar- und Krebsorganoiden mit Einzelzellauflösung durch. Ich fand heraus, dass die Alveolarorganoide einer Regenerationsreaktion folgten. Die Zellen differenzierten in Kultur zu alveolären Typ 1- (AT1) und AT2-Zellen, ähnlich ihrer Differenzierungsverläufe *in vivo*. Im Gegensatz dazu hatten Krebsorganoid-Zellen die alveolären Differenzierungswege verloren und exprimierten stattdessen Transkriptionsfaktoren, die mit Pluripotenz und Entwicklung zusammenhängen, früh nach der Initiierung des Onkogens. Obwohl die Krebszellen zu einem späteren Zeitpunkt teilweise AT2-Identitätsmarker wiedererlangten, blieben sie in einem

dedifferenzierten Zustand.

Zusammenfassend lässt sich festhalten, dass ich ein neues Organoid-System entwickelt habe, das für eine Reihe von Anwendungen verwendet werden kann. Ich charakterisierte das System und untersuchte die Transkriptionsänderungen, die auf die Krebsentstehung folgten. Ich zeigte, dass die onkogene KRAS-Expression früh nach der Initiierung eine anhaltende Dedifferenzierungsreaktion hervorrief. Die Beschreibung der molekularen Mechanismen, die zu dieser Reaktion führten, wird zum Verständnis der KRAS-Biologie beitragen und neue Wege für ziegerichtete Behandlungen eröffnen.

Bibliography

Alemany, A., Florescu, M., Baron, C. S., Peterson-Maduro, J., and van Oudenaarden, A. (2018). Whole-organism clone tracing using single-cell sequencing. *Nature*, 556(7699):108–112.

Alguacil-Núñez, C., Ferrer-Ortiz, I., García-Verdú, E., López-Pirez, P., Llorente-Cortijo, I. M., and Sainz, B. (2018). Current perspectives on the crosstalk between lung cancer stem cells and cancer-associated fibroblasts. *Critical Reviews in Oncology/Hematology*, 125:102–110.

Allemani, C., Weir, H. K., Carreira, H., Harewood, R., Spika, D., Wang, X. S., Bannon, F., Ahn, J. V., Johnson, C. J., Bonaventure, A., *et al.* (2015). Global surveillance of cancer survival 1995-2009: Analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *The Lancet*, 385(9972):977–1010.

Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., Buchholz, F., Edenhofer, F., and Stewart, A. F. (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. *DMM Disease Models and Mechanisms*, 2(9-10):508–515.

Antonângelo, L., Tuma, T., Fabro, A., Acencio, M., Terra, R., Parra, E., Vargas, F., Takagaki, T., and Capelozzi, V. (2016). Id-1, Id-2, and Id-3 co-expression correlates with prognosis in stage I and II lung adenocarcinoma patients treated with surgery and adjuvant chemotherapy. *Experimental Biology and Medicine*, 241(11):1159–1168.

Arwert, E. N., Hoste, E., and Watt, F. M. (2012). Epithelial stem cells, wound healing and cancer. *Nature Reviews Cancer*, 12(3):170–180.

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., Schinzel, A. C., Sandy, P., Meylan, E., Scholl, C., *et al.* (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*, 462(7269):108–112.

Barkauskas, C. E., Cronce, M. J., Rackley, C. R., Bowie, E. J., Keene, D. R., Stripp, B. R., Randell, S. H., Noble, P. W., and Hogan, B. L. (2013). Type 2 alveolar cells are stem cells in adult lung. *Journal of Clinical Investigation*, 123(7):3025–3036.

Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H. J., van Es, J. H., Sato, T., Stange, D. E., Begthel, H., van den Born, M., *et al.* (2010). Lgr5+ve Stem Cells Drive Self-Renewal in the Stomach and Build Long-Lived Gastric Units In Vitro. *Cell Stem Cell*, 6(1):25–36.

Barletta, J. A., Perner, S., Iafrate, A. J., Yeap, B. Y., Weir, B. A., Johnson, L. A., Johnson, B. E., Meyerson, M., Rubin, M. A., Travis, W. D., *et al.* (2009). Clinical significance of TTF-1 protein expression and TTF-1 gene amplification in lung adenocarcinoma. *Journal of Cellular and Molecular Medicine*, 13(8b):1977–1986.

Bartfeld, S., Bayram, T., Van De Wetering, M., Huch, M., Begthel, H., Kujala, P., Vries, R., Peters, P. J., and Clevers, H. (2015). In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology*, 148(1):126–136.e6.

Bassères, D. S., Ebbs, A., Levantini, E., and Baldwin, A. S. (2010). Requirement of the NF-κB subunit p65/RelA for K-Ras-induced lung tumorigenesis. *Cancer Research*, 70(9):3537–3546.

Batlle, E. and Clevers, H. (2017). Cancer stem cells revisited. *Nature Medicine*, 23(10):1124–1134.

Bellusci, S., Furuta, Y., Rush, M. G., Henderson, R., Winnier, G., and Hogan, B. L. (1997). Involvement of Sonic hedgehog (Shh) in mouse embryonic lung growth and morphogenesis. *Development (Cambridge, England)*, 124(1):53–63.

Ben-David, U., Siranosian, B., Ha, G., Tang, H., Oren, Y., Hinohara, K., Strathdee, C. A., Dempster, J., Lyons, N. J., Burns, R., *et al.* (2018). Genetic and transcriptional evolution alters cancer cell line drug response. *Nature*, 560(7718):325–330.

Berghmans, T., Paesmans, M., Mascaux, C., Martin, B., Meert, A.-P., Haller, A., Lafitte, J.-J., and Sculier, J.-P. (2006). Thyroid transcription factor 1—a new prognostic factor in lung cancer: a meta-analysis. *Annals of Oncology*, 17(11):1673–1676.

Berkers, G., van Mourik, P., Vonk, A. M., Kruisselbrink, E., Dekkers, J. F., de Winter-de Groot, K. M., Arets, H. G., Marck-van der Wilt, R. E., Dijkema, J. S., Vanderschuren,

M. M., *et al.* (2019). Rectal Organoids Enable Personalized Treatment of Cystic Fibrosis. *Cell Reports*, 26(7):1701–1708.e3.

Beskow, L. M. (2016). Lessons from HeLa Cells: The Ethics and Policy of Biospecimens. *Annual Review of Genomics and Human Genetics*, 17(1):395–417.

Bian, S., Repic, M., Guo, Z., Kavirayani, A., Burkard, T., Bagley, J. A., Krauditsch, C., and Knoblich, J. A. (2018). Genetically engineered cerebral organoids model brain tumor formation. *Nature Methods*, 15(8):631–639.

Bild, A. H., Yao, G., Chang, J. T., Wang, Q., Potti, A., Chasse, D., Joshi, M. B., Harpole, D., Lancaster, J. M., Berchuck, A., *et al.* (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature*, 439(7074):353–357.

Boers, J. E., Den Brok, J. L., Koudstaal, J., Arends, J. W., and Thunnissen, F. B. (1996). Number and proliferation of neuroendocrine cells in normal human airway epithelium. *American Journal of Respiratory and Critical Care Medicine*, 154(3 I):758–763.

Boj, S. F., Hwang, C. I., Baker, L. A., Chio, I. I. C., Engle, D. D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M. S., *et al.* (2015). Organoid models of human and mouse ductal pancreatic cancer. *Cell*, 160(1-2):324–338.

Braiman, A. and Priel, Z. (2008). Efficient mucociliary transport relies on efficient regulation of ciliary beating. *Respiratory Physiology and Neurobiology*, 163(1-3):202–207.

Branchfield, K., Nantie, L., Verheyden, J. M., Sui, P., Wienhold, M. D., and Sun, X. (2016). Pulmonary neuroendocrine cells function as airway sensors to control lung immune response. *Science*, 351(6274):707–710.

Brannan, J. M., Dong, W., Prudkin, L., Behrens, C., Lotan, R., Bekele, B. N., Wistuba, I., and Johnson, F. (2009a). Expression of the receptor tyrosine kinase EphA2 is increased in smokers and predicts poor survival in non -small cell lung cancer. *Clinical Cancer Research*, 15(13):4423–4430.

Brannan, J. M., Sen, B., Saigal, B., Prudkin, L., Behrens, C., Solis, L., Dong, W., Bekele,

B. N., Wistuba, I., and Johnson, F. M. (2009b). EphA2 in the early pathogenesis and progression of non-small cell lung cancer. *Cancer Prevention Research*, 2(12):1039–1049.

Camp, J. G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., *et al.* (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51):15672–15677.

Cao, J., Cusanovich, D. A., Ramani, V., Aghamirzaie, D., Pliner, H. A., Hill, A. J., Daza, R. M., McFaline-Figueroa, J. L., Packer, J. S., Christiansen, L., *et al.* (2018). Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science*, 361(6409):1380–1385.

Ceteci, F., Xu, J., Ceteci, S., Zanucco, E., Thakur, C., and Rapp, U. R. (2011). Conditional expression of oncogenic C-RAF in mouse pulmonary epithelial cells reveals differential tumorigenesis and induction of autophagy leading to tumor regression. *Neoplasia*, 13(11):1005–1018.

Chang, D. R., Alanis, D. M., Miller, R. K., Ji, H., Akiyama, H., McCrea, P. D., and Chen, J. (2013). Lung epithelial branching program antagonizes alveolar differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45):18042–18051.

Chapman, H. A., Li, X., Alexander, J. P., Brumwell, A., Lorizio, W., Tan, K., Sonnenberg, A., Wei, Y., and Vu, T. H. (2011). Integrin $\alpha 6\beta 4$ identifies an adult distal lung epithelial population with regenerative potential in mice. *Journal of Clinical Investigation*, 121(7):2855–2862.

Chen, F. and Fine, A. (2016). Stem Cells in Lung Injury and Repair. *American Journal of Pathology*, 186(10):2544–2550.

Chen, G., Korfhagen, T. R., Xu, Y., Kitzmiller, J., Wert, S. E., Maeda, Y., Gregorieff, A., Clevers, H., and Whitsett, J. A. (2009). SPDEF is required for mouse pulmonary goblet cell differentiation and regulates a network of genes associated with mucus production. *Journal of Clinical Investigation*, 119(10):2914–2924.

Chen, H., Carrot-Zhang, J., Zhao, Y., Hu, H., Freeman, S. S., Yu, S., Ha, G., Taylor, A. M., Berger, A. C., Westlake, L., *et al.* (2019). Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nature Communications*, 10(1).

Chen, H., Matsumoto, K., Brockway, B. L., Rackley, C. R., Liang, J., Lee, J.-H., Jiang, D., Noble, P. W., Randell, S. H., Kim, C. F., *et al.* (2012). Airway Epithelial Progenitors Are Region Specific and Show Differential Responses to Bleomycin-Induced Lung Injury. *STEM CELLS*, 30(9):1948–1960.

Chen, X. and Song, E. (2019). Turning foes to friends: targeting cancer-associated fibroblasts. *Nature Reviews Drug Discovery*, 18(2):99–115.

Cheng, Y. J., Tsai, J. W., Hsieh, K. C., Yang, Y. C., Chen, Y. J., Huang, M. S., and Yuan, S. S. (2011). Id1 promotes lung cancer cell proliferation and tumor growth through Akt-related pathway. *Cancer Letters*, 307(2):191–199.

Choi, J., Park, J.-E., Tsagkogeorga, G., Yanagita, M., Koo, B.-K., Han, N., and Lee, J.-H. (2020). Inflammatory Signals Induce AT2 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration. *Cell Stem Cell*, 27(3):366–382.e7.

Choi, Y., Syeda, F., Walker, J. R., Finerty, P. J., Cuerrier, D., Wojciechowski, A., Liu, Q., Dhe-Paganon, S., and Gray, N. S. (2009). Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorganic and Medicinal Chemistry Letters*, 19(15):4467–4470.

Chua, C. W., Shibata, M., Lei, M., Toivanen, R., Barlow, L. J., Bergren, S. K., Badani, K. K., McKiernan, J. M., Benson, M. C., Hibshoosh, H., *et al.* (2014). Single luminal epithelial progenitors can generate prostate organoids in culture. *Nature Cell Biology*, 16(10):951–961.

Chung, M. I., Bujnis, M., Barkauskas, C. E., Kobayashi, Y., and Hogan, B. L. (2018). Niche-mediated BMP/SMAD signaling regulates lung alveolar stem cell proliferation and differentiation. *Development (Cambridge)*, 145(9).

Chung, M.-I. and Hogan, B. L. M. (2018). Ager-CreER T2 : A New Genetic Tool for Studying Lung Alveolar Development, Homeostasis, and Repair. *American Journal of Respiratory Cell and Molecular Biology*, 59(6):706–712.

Chung, W.-J., Daemen, A., Cheng, J. H., Long, J. E., Cooper, J. E., Wang, B.-e., Tran, C., Singh, M., Gnad, F., Modrusan, Z., *et al.* (2017). Kras mutant genetically engineered mouse models of human cancers are genomically heterogeneous. *Proceedings of the National Academy of Sciences*, 114(51):E10947–E10955.

Ciriello, G., Miller, M. L., Aksoy, B. A., Senbabaoglu, Y., Schultz, N., and Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. *Nature Genetics*, 45(10):1127–1133.

Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nature Medicine*, 17(3):313–319.

Clevers, H. (2016). Modeling Development and Disease with Organoids. *Cell*, 165(7):1586–1597.

Collisson, E. A., Campbell, J. D., Brooks, A. N., Berger, A. H., Lee, W., Chmielecki, J., Beer, D. G., Cope, L., Creighton, C. J., Danilova, L., *et al.* (2014). Comprehensive molecular profiling of lung adenocarcinoma: The cancer genome atlas research network. *Nature*, 511(7511):543–550.

Cornwell, M. I., Vangala, M., Taing, L., Herbert, Z., Köster, J., Li, B., Sun, H., Li, T., Zhang, J., Qiu, X., *et al.* (2018). VIPER: Visualization Pipeline for RNA-seq, a Snakemake workflow for efficient and complete RNA-seq analysis. *BMC Bioinformatics*, 19(1).

Cruz-Bermúdez, A., Laza-Briviesca, R., Vicente-Blanco, R. J., García-Grande, A., Coronado, M. J., Laine-Menéndez, S., Alfaro, C., Sanchez, J. C., Franco, F., Calvo, V., *et al.* (2019). Cancer-associated fibroblasts modify lung cancer metabolism involving ROS and TGF-β signaling. *Free Radical Biology and Medicine*, 130:163–173.

Cugola, F. R., Fernandes, I. R., Russo, F. B., Freitas, B. C., Dias, J. L., Guimarães, K. P., Benazzato, C., Almeida, N., Pignatari, G. C., Romero, S., *et al.* (2016). The Brazilian Zika virus strain causes birth defects in experimental models. *Nature*, 534(7606):267–271.

Curtis, S. J., Sinkevicius, K. W., Li, D., Lau, A. N., Roach, R. R., Zamponi, R., Woolfenden, A. E., Kirsch, D. G., Wong, K. K., and Kim, C. F. (2010). Primary tumor genotype is an

important determinant in identification of lung cancer propagating cells. *Cell Stem Cell*, 7(1):127–133.

Cutz, E., Pan, J., Yeger, H., Domnik, N. J., and Fisher, J. T. (2013). Recent advances and contraversies on the role of pulmonary neuroepithelial bodies as airway sensors. *Seminars in Cell and Developmental Biology*, 24(1):40–50.

Danahay, H., Pessotti, A. D., Coote, J., Montgomery, B. E., Xia, D., Wilson, A., Yang, H., Wang, Z., Bevan, L., Thomas, C., *et al.* (2015). Notch2 Is Required for Inflammatory Cytokine-Driven Goblet Cell Metaplasia in the Lung. *Cell Reports*, 10(2):239–252.

Dankort, D., Filenova, E., Collado, M., Serrano, M., Jones, K., and McMahon, M. (2007). A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. *Genes and Development*, 21(4):379–384.

De Koning, H. J., Van Der Aalst, C. M., De Jong, P. A., Scholten, E. T., Nackaerts, K., Heuvelmans, M. A., Lammers, J. W., Weenink, C., Yousaf-Khan, U., Horeweg, N., *et al.* (2020). Reduced lung-cancer mortality with volume CT screening in a randomized trial. *New England Journal of Medicine*, 382(6):503–513.

Dekkers, J. F., Wiegerinck, C. L., De Jonge, H. R., Bronsveld, I., Janssens, H. M., De Winter-De Groot, K. M., Brandsma, A. M., De Jong, N. W., Bijvelds, M. J., Scholte, B. J., *et al.* (2013). A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nature Medicine*, 19(7):939–945.

DerHagopian, R. P., Sugarbaker, E. V., and Ketcham, A. (1978). Inflammatory Oncotaxis. *JAMA: The Journal of the American Medical Association*, 240(4):374.

Desai, T. J., Brownfield, D. G., and Krasnow, M. A. (2014). Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature*, 507(7491):190–194.

Ding, L., Getz, G., Wheeler, D. A., Mardis, E. R., McLellan, M. D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D. M., Morgan, M. B., *et al.* (2008). Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*, 455(7216):1069–1075.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21.

Dogan, S., Shen, R., Ang, D. C., Johnson, M. L., D'Angelo, S. P., Paik, P. K., Brzostowski, E. B., Riely, G. J., Kris, M. G., Zakowski, M. F., *et al.* (2012). Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clinical Cancer Research*, 18(22):6169–6177.

Dost, A. F., Moye, A. L., Vedaie, M., Tran, L. M., Fung, E., Heinze, D., Villacorta-Martin, C., Huang, J., Hekman, R., Kwan, J. H., *et al.* (2020). Organoids Model Transcriptional Hallmarks of Oncogenic KRAS Activation in Lung Epithelial Progenitor Cells. *Cell Stem Cell*, 27(4):663–678.e8.

Drost, J., Van Jaarsveld, R. H., Ponsioen, B., Zimberlin, C., Van Boxtel, R., Buijs, A., Sachs, N., Overmeer, R. M., Offerhaus, G. J., Begthel, H., *et al.* (2015). Sequential cancer mutations in cultured human intestinal stem cells. *Nature*, 521(7550):43–47.

Duarte, A. A., Gogola, E., Sachs, N., Barazas, M., Annunziato, S., R De Ruiter, J., Velds, A., Blatter, S., Houthuijzen, J. M., Van De Ven, M., *et al.* (2018). BRCA-deficient mouse mammary tumor organoids to study cancer-drug resistance. *Nature Methods*, 15(2):134–140.

Duchartre, Y., Kim, Y. M., and Kahn, M. (2016). The Wnt signaling pathway in cancer. *Critical Reviews in Oncology/Hematology*, 99:141–149.

DuPage, M., Dooley, A. L., and Jacks, T. (2009). Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols*, 4(7):1064–1072.

Dvorak, H. F. (1986). Tumors: Wounds That Do Not Heal. *New England Journal of Medicine*, 315(26):1650–1659.

Efremova, M., Vento-Tormo, M., Teichmann, S. A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature Protocols*, 15(4):1484–1506.

Eiraku, M. and Sasai, Y. (2012). Self-formation of layered neural structures in three-dimensional culture of ES cells. *Current Opinion in Neurobiology*, 22(5):768–777.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature*, 472(7341):51–58.

Elkin, M. and Vlodavsky, I. (2001). Tail Vein Assay of Cancer Metastasis. *Current Protocols in Cell Biology*, 12(1).

Engelman, J. A., Chen, L., Tan, X., Crosby, K., Guimaraes, A. R., Upadhyay, R., Maira, M., McNamara, K., Perera, S. A., Song, Y., *et al.* (2008). Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. *Nature Medicine*, 14(12):1351–1356.

Fadous-Khalifé, M. C., Aloulou, N., Jalbout, M., Hadchity, J., Aftimos, G., Paris, F., and Hadchity, E. (2016). Krüppel-like factor 4: A new potential biomarker of lung cancer. *Molecular and Clinical Oncology*, 5(1):35–40.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, 61(5):759–767.

Fisher, G. H., Wellen, S. L., Klimstra, D., Lenczowski, J. M., Tichelaar, J. W., Lizak, M. J., Whitsett, J. A., Koretsky, A., and Varmus, H. E. (2001). Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. *Genes and Development*, 15(24):3249–3262.

Fodde, R., Smits, R., and Clevers, H. (2001). APC, signal transduction and genetic instability in colorectal cancer. *Nature Reviews Cancer*, 1(1):55–67.

Franzén, O., Gan, L.-M., and Björkegren, J. L. M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database*, 2019.

Fujii, M., Shimokawa, M., Date, S., Takano, A., Matano, M., Nanki, K., Ohta, Y., Toshimitsu, K., Nakazato, Y., Kawasaki, K., *et al.* (2016). A Colorectal Tumor Organoid Library

Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. *Cell Stem Cell*, 18(6):827–838.

Gao, D., Vela, I., Sboner, A., Iaquinta, P. J., Karthaus, W. R., Gopalan, A., Dowling, C., Wanjala, J. N., Undvall, E. A., Arora, V. K., *et al.* (2014). Organoid cultures derived from patients with advanced prostate cancer. *Cell*, 159(1):176–187.

Garcez, P. P., Loiola, E. C., Da Costa, R. M., Higa, L. M., Trindade, P., Delvecchio, R., Nascimento, J. M., Brindeiro, R., Tanuri, A., and Rehen, S. K. (2016). Zika virus: Zika virus impairs growth in human neurospheres and brain organoids. *Science*, 352(6287):816–818.

Ge, Y., Gomez, N. C., Adam, R. C., Nikolova, M., Yang, H., Verma, A., Lu, C. P.-J., Polak, L., Yuan, S., Elemento, O., *et al.* (2017). Stem Cell Lineage Infidelity Drives Wound Repair and Cancer. *Cell*, 169(4):636–650.e14.

Gehr, P., Bachofen, M., and Weibel, E. R. (1978). The normal human lung: ultrastructure and morphometric estimation of diffusion capacity. *Respiration Physiology*, 32(2):121–140.

Ghaleb, A. M. and Yang, V. W. (2017). Krüppel-like factor 4 (KLF4): What we currently know. *Gene*, 611:27–37.

Gomes, M., Teixeira, A. L., Coelho, A., Araújo, A., and Medeiros, R. (2014). The Role of Inflammation in Lung Cancer. In *Advances in experimental medicine and biology*, volume 816, pages 1–23. Adv Exp Med Biol.

Goss, A. M., Tian, Y., Tsukiyama, T., Cohen, E. D., Zhou, D., Lu, M. M., Yamaguchi, T. P., and Morrisey, E. E. (2009). Wnt2/2b and β-Catenin Signaling Are Necessary and Sufficient to Specify Lung Progenitors in the Foregut. *Developmental Cell*, 17(2):290–298.

Guerra, C., Mijimolle, N., Dhawahir, A., Dubus, P., Barradas, M., Serrano, M., Campuzano, V., and Barbacid, M. (2003). Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. *Cancer Cell*, 4(2):111–120.

Guggino, W. B. and Banks-Schlegel, S. P. (2004). Macromolecular interactions and ion

transport in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine*, 170(7):815–820.

Gupta, R., Leon, F., Thompson, C. M., Nimmakayala, R., Karmakar, S., Nallasamy, P., Chugh, S., Prajapati, D. R., Rachagani, S., Kumar, S., *et al.* (2020). Global analysis of human glycosyltransferases reveals novel targets for pancreatic cancer pathogenesis. *British Journal of Cancer*, 122(11):1661–1672.

Gupta, S., Ramjaun, A. R., Haiko, P., Wang, Y., Warne, P. H., Nicke, B., Nye, E., Stamp, G., Alitalo, K., and Downward, J. (2007). Binding of Ras to Phosphoinositide 3-Kinase p110α Is Required for Ras- Driven Tumorigenesis in Mice. *Cell*, 129(5):957–968.

Haigis, K. M. (2017). KRAS Alleles: The Devil Is in the Detail. *Trends in Cancer*, 3(10):686–697.

Han, H., Cho, J. W., Lee, S. S., Yun, A., Kim, H., Bae, D., Yang, S., Kim, C. Y., Lee, M., Kim, E., *et al.* (2018). TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Research*, 46(D1):D380–D386.

Harris-Johnson, K. S., Domyan, E. T., Vezina, C. M., and Sun, X. (2009). β-Catenin promotes respiratory progenitor identity in mouse foregut. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38):16287–16292.

Hasskarl, J. and Munger, K. (2002). Id Proteins - Tumor Markers or Oncogenes? *Cancer Biology & Therapy*, 1(2):91–96.

Herriges, M. and Morrisey, E. E. (2014). Lung development: orchestrating the generation and regeneration of a complex organ. *Development*, 141(3):502–513.

Heuvelmans, M. A., Groen, H. J., and Oudkerk, M. (2017). Early lung cancer detection by low-dose CT screening: therapeutic implications. *Expert Review of Respiratory Medicine*, 11(2):89–100.

Howlader, N., Noone, A., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D., *et al.* (2020). SEER Cancer Statistics Review, 1975-2017.

Hsia, C. C., Hyde, D. M., and Weibel, E. R. (2016). Lung structure and the intrinsic challenges of gas exchange. *Comprehensive Physiology*, 6(2):827–895.

Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., and Guo, A.-Y. (2019). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic acids research*, 47(D1):D33–D38.

Hu, W., Hofstetter, W. L., Li, H., Zhou, Y., He, Y., Pataer, A., Wang, L., Xie, K., Swisher, S. G., and Fang, B. (2009). Putative Tumor-Suppressive Function of Krüppel-Like Factor 4 in Primary Lung Carcinoma. *Clinical Cancer Research*, 15(18):5688–5695.

Huang, J. Q., Wei, F. K., Xu, X. L., Ye, S. X., Song, J. W., Ding, P. K., Zhu, J., Li, H. F., Luo, X. P., Gong, H., *et al.* (2019). SOX9 drives the epithelial-mesenchymal transition in non-small-cell lung cancer through the Wnt/β-catenin pathway. *Journal of Translational Medicine*, 17(1).

Hubert, C. G., Rivera, M., Spangler, L. C., Wu, Q., Mack, S. C., Prager, B. C., Couce, M., McLendon, R. E., Sloan, A. E., and Rich, J. N. (2016). A three-dimensional organoid culture system derived from human glioblastomas recapitulates the hypoxic gradients and cancer stem cell heterogeneity of tumors found in vivo. *Cancer Research*, 76(8):2465–2477.

Huch, M., Bonfanti, P., Boj, S. F., Sato, T., Loomans, C. J., Van De Wetering, M., Sojoodi, M., Li, V. S., Schuijers, J., Gracanin, A., *et al.* (2013a). Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. *EMBO Journal*, 32(20):2708–2721.

Huch, M., Dorrell, C., Boj, S. F., Van Es, J. H., Li, V. S., Van De Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M. J., *et al.* (2013b). In vitro expansion of single Lgr5 + liver stem cells induced by Wnt-driven regeneration. *Nature*, 494(7436):247–250.

Huch, M., Gehart, H., Van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M. M., Ellis, E., Van Wenum, M., Fuchs, S. A., De Ligt, J., *et al.* (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell*, 160(1-2):299–312.

Hunter, J. C., Manandhar, A., Carrasco, M. A., Gurbani, D., Gondi, S., and Westover, K. D.

(2015). Biochemical and Structural Analysis of Common Cancer-Associated KRAS Mutations. *Molecular Cancer Research*, 13(9):1325–1335.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(3):99–104.

Hyatt, B. A., Shangguan, X., and Shannon, J. M. (2002). BMP4 modulates fibroblast growth factor-mediated induction of proximal and distal lung differentiation in mouse embryonic tracheal epithelium in mesenchyme-free culture. *Developmental Dynamics*, 225(2):153–165.

Ieguchi, K. and Maru, Y. (2019). Roles of EphA1/A2 and ephrin-A1 in cancer. *Cancer Science*, 110(3):841–848.

Inamura, K. (2018). Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: Tumor initiation and progression. *International Journal of Molecular Sciences*, 19(4).

Ishikawa, M., Miyahara, R., Sonobe, M., Horiuchi, M., Mennju, T., Nakayama, E., Kobayashi, M., Kikuchi, R., Kitamura, J., Imamura, N., *et al.* (2012). Higher expression of EphA2 and ephrin-A1 is related to favorable clinicopathological features in pathological stage I non-small cell lung carcinoma. *Lung Cancer*, 76(3):431–438.

Iwanaga, K., Yang, Y., Raso, M. G., Ma, L., Hanna, A. E., Thilaganathan, N., Moghaddam, S., Evans, C. M., Li, H., Cai, W. W., *et al.* (2008). Pten inactivation accelerates oncogenic K-ras-initiated tumorigenesis in a mouse model of lung cancer. *Cancer Research*, 68(4):1119–1127.

Jackson, E. L., Olive, K. P., Tuveson, D. A., Bronson, R., Crowley, D., Brown, M., and Jacks, T. (2005). The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Research*, 65(22):10280–10288.

Jackson, E. L., Willis, N., Mercer, K., Bronson, R. T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D. A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes and Development*, 15(24):3243–3248.

Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE*, 9(6):e98679.

Jain, R., Barkauskas, C. E., Takeda, N., Bowie, E. J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L. J., Gupta, M., Li, D., *et al.* (2015). Plasticity of Hopx+ type I alveolar cells to regenerate type II cells in the lung. *Nature Communications*, 6.

Jiang, S. S., Fang, W. T., Hou, Y. H., Huang, S. F., Yen, B. L., Chang, J. L., Li, S. M., Liu, H. P., Liu, Y. L., Huang, C. T., *et al.* (2010). Upregulation of SOX9 in lung adenocarcinoma and its involvement in the regulation of cell growth and tumorigenicity. *Clinical Cancer Research*, 16(17):4363–4373.

Jin, L., Liu, Y., Wang, X., and Qi, X. (2018). Immunohistochemical analysis and comparison of napsin A, TTF1, SPA and CK7 expression in primary lung adenocarcinoma. *Biotechnic and Histochemistry*, 93(5):364–372.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R. T., Crowley, D., Tuveson, D. A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature*, 410(6832):1111–1116.

Karthaus, W. R., Iaquinta, P. J., Drost, J., Gracanin, A., Van Boxtel, R., Wongvipat, J., Dowling, C. M., Gao, D., Begthel, H., Sachs, N., *et al.* (2014). Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell*, 159(1):163–175.

Katsura, H., Kobayashi, Y., Tata, P. R., and Hogan, B. L. (2019). IL-1 and TNFα Contribute to the Inflammatory Niche to Enhance Alveolar Regeneration. *Stem Cell Reports*, 12(4):657–666.

Katsura, H., Sontake, V., Tata, A., Kobayashi, Y., Edwards, C. E., Heaton, B. E., Konkimalla, A., Asakura, T., Mikami, Y., Fritch, E. J., *et al.* (2020). Human Lung Stem Cell-Based Alveolospheres Provide Insights into SARS-CoV-2-Mediated Interferon Responses and Pneumocyte Dysfunction. *Cell Stem Cell*.

Khuder, S. A. (2001). Effect of cigarette smoking on major histological types of lung cancer:

A meta-analysis. *Lung Cancer*, 31(2-3):139–148.

Kim, C. F., Jackson, E. L., Woolfenden, A. E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R. T., and Jacks, T. (2005). Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell*, 121(6):823–835.

Kim, J., Koo, B. K., and Knoblich, J. A. (2020). Human organoids: model systems for human biology and medicine. *Nature Reviews Molecular Cell Biology*, 21(10):571–584.

Kim, M., Mun, H., Sung, C. O., Cho, E. J., Jeon, H. J., Chun, S. M., Jung, D. J., Shin, T. H., Jeong, G. S., Kim, D. K., *et al.* (2019). Patient-derived lung cancer organoids as in vitro cancer models for therapeutic screening. *Nature Communications*, 10(1).

Kleinman, H. K. and Martin, G. R. (2005). Matrigel: Basement membrane matrix with biological activity. *Seminars in Cancer Biology*, 15(5 SPEC. ISS.):378–386.

Kondo, J., Endo, H., Okuyama, H., Ishikawa, O., Iishi, H., Tsujii, M., Ohue, M., and Inoue, M. (2011). Retaining cell-cell contact enables preparation and culture of spheroids composed of pure primary cancer cells from colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 108(15):6235–6240.

Krasteva, G. and Kummer, W. (2012). "Tasting" the airway lining fluid. *Histochemistry and Cell Biology*, 138(3):365–383.

Kruspig, B., Monteverde, T., Neidler, S., Hock, A., Kerr, E., Nixon, C., Clark, W., Hedley, A., Laing, S., Coffelt, S. B., *et al.* (2018). The ERBB network facilitates KRAS-driven lung tumorigenesis. *Science Translational Medicine*, 10(44).

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., *et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*, 44(W1):W90–W97.

Kunii, R., Jiang, S., Hasegawa, G., Yamamoto, T., Umezu, H., Watanabe, T., Tsuchida, M., Hashimoto, T., Hamakubo, T., Kodama, T., *et al.* (2011). The predominant expression of

hepatocyte nuclear factor 4α (HNF4α) in thyroid transcription factor-1 (TTF-1)-negative pulmonary adenocarcinoma. *Histopathology*, 58(3):467–476.

Kurmann, A. A., Serra, M., Hawkins, F., Rankin, S. A., Mori, M., Astapova, I., Ullas, S., Lin, S., Bilodeau, M., Rossant, J., *et al.* (2015). Regeneration of Thyroid Function by Transplantation of Differentiated Pluripotent Stem Cells. *Cell Stem Cell*, 17(5):527–542.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., *et al.* (2018). RNA velocity of single cells. *Nature*, 560(7719):494–498.

LaFave, L. M., Kartha, V. K., Ma, S., Meli, K., Del Priore, I., Lareau, C., Naranjo, S., Westcott, P. M., Duarte, F. M., Sankar, V., *et al.* (2020). Epigenomic State Transitions Characterize Tumor Progression in Mouse Lung Adenocarcinoma. *Cancer Cell*, 38(2):212–228.e13.

Lamers, M. M., Beumer, J., Vaart, J. V. D., Knoops, K., Puschhof, J., Breugem, T. I., Ravelli, R. B., Schayck, J. P. V., Mykytyn, A. Z., Duimel, H. Q., *et al.* (2020). SARS-CoV-2 productively infects human gut enterocytes. *Science*, 369(6499):50–54.

Lancaster, M. A. and Knoblich, J. A. (2014). Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science*, 345(6194):1247125–1247125.

Lancaster, M. A., Renner, M., Martin, C. A., Wenzel, D., Bicknell, L. S., Hurles, M. E., Homfray, T., Penninger, J. M., Jackson, A. P., and Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467):373–379.

Laughney, A. M., Hu, J., Campbell, N. R., Bakhoum, S. F., Setty, M., Lavallée, V. P., Xie, Y., Masilionis, I., Carr, A. J., Kottapalli, S., *et al.* (2020). Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nature Medicine*, 26(2):259–269.

Lechner, A. J., Driver, I. H., Lee, J., Conroy, C. M., Nagle, A., Locksley, R. M., and Rock, J. R. (2017). Recruited Monocytes and Type 2 Immunity Promote Lung Regeneration following Pneumonectomy. *Cell Stem Cell*, 21(1):120–134.e7.

Lee, G. Y., Kenny, P. A., Lee, E. H., and Bissell, M. J. (2007). Three-dimensional culture models of normal and malignant breast epithelial cells. *Nature Methods*, 4(4):359–365.

Lee, J.-H., Bhang, D. H., Beede, A., Huang, T. L., Stripp, B. R., Bloch, K. D., Wagers, A. J., Tseng, Y.-H., Ryeom, S., and Kim, C. F. (2014). Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. *Cell*, 156(3):440–55.

Lee, J.-H., Tammela, T., Hofree, M., Jacks, T., Regev, A., Kim Correspondence, C. F., Choi, J., Marjanovic, N. D., Han, S., Canner, D., *et al.* (2017). Anatomically and Functionally Distinct Lung Mesenchymal Populations Marked by Lgr5 and Lgr6. *Cell*, 170(6):1149–1163.e12.

Lee, S. H., Hu, W., Matulay, J. T., Silva, M. V., Owczarek, T. B., Kim, K., Chua, C. W., Barlow, L. M. J., Kandoth, C., Williams, A. B., *et al.* (2018). Tumor Evolution and Drug Response in Patient-Derived Organoid Models of Bladder Cancer. *Cell*, 173(2):515–528.e17.

Li, L., Xu, B., Zhang, H., Wu, J., Song, Q., and Yu, J. (2020). Potentiality of forkhead box Q1 as a biomarker for monitoring tumor features and predicting prognosis in non-small cell lung cancer. *Journal of Clinical Laboratory Analysis*, 34(1).

Li, X., Nadauld, L., Ootani, A., Corney, D. C., Pai, R. K., Gevaert, O., Cantrell, M. A., Rack, P. G., Neal, J. T., Chan, C. W., *et al.* (2014). Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture. *Nature Medicine*, 20(7):769–777.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*, 1(6):417–425.

Liu, Q., Liu, K., Cui, G., Huang, X., Yao, S., Guo, W., Qin, Z., Li, Y., Yang, R., Pu, W., *et al.* (2019a). Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. *Nature Genetics*, 51(4):728–738.

Liu, Y., Jiang, H., Crawford, H. C., and Hogan, B. L. (2003). Role for ETS domain transcription factors Pea3/Erm in mouse lung development. *Developmental Biology*, 261(1):10–24.

Liu, Y., Mi, Y., Mueller, T., Kreibich, S., Williams, E. G., Van Drogen, A., Borel, C., Frank, M., Germain, P. L., Bludau, I., *et al.* (2019b). Multi-omic measurements of heterogeneity in HeLa cells across laboratories. *Nature Biotechnology*, 37(3):314–322.

Lourenço, A. R. and Coffer, P. J. (2017). SOX4: Joining the Master Regulators of Epithelial-to-Mesenchymal Transition? *Trends in Cancer*, 3(8):571–582.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550.

Lu, M. M., Yang, H., Zhang, L., Shu, W., Blair, D. G., and Morrisey, E. E. (2001). The bone morphogenic protein antagonist gremlin regulates proximal-distal patterning of the lung. *Developmental Dynamics*, 222(4):667–680.

Lyabin, D. N., Eliseeva, I. A., and Ovchinnikov, L. P. (2014). YB-1 protein: functions and regulation. *Wiley Interdisciplinary Reviews: RNA*, 5(1):95–110.

Macaulay, I. C., Haerty, W., Kumar, P., Li, Y. I., Hu, T. X., Teng, M. J., Goolam, M., Saurat, N., Coupland, P., Shirley, L. M., *et al.* (2015). G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nature Methods*, 12(6):519–522.

Mainardi, S., Mijimolle, N., Francoz, S., Vicente-Dueñas, C., Sánchez-García, I., and Barbacid, M. (2014). Identification of cancer initiating cells in K-Ras driven lung adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, 111(1):255–60.

Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., Amenduni, M., Szekely, A., Palejev, D., Wilson, M., *et al.* (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell*, 162(2):375–390.

Marjanovic, N. D., Hofree, M., Chan, J. E., Canner, D., Wu, K., Trakala, M., Hartmann, G. G., Smith, O. C., Kim, J. Y., Evans, K. V., *et al.* (2020). Emergence of a High-Plasticity Cell State during Lung Cancer Evolution. *Cancer Cell*, 38(2):229–246.e13.

Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nature Medicine*, 21(3):256–262.

Mazieres, J., He, B., You, L., Xu, Z., Lee, A. Y., Mikami, I., Reguart, N., Rosell, R., Mccormick, F., and Jablons, D. M. (2004). Wnt Inhibitory Factor-1 Is Silenced by Promoter Hypermethylation in Human Lung Cancer. *Cancer Research*, 64:4717–4720.

Mazzocchi, A., Devarasetty, M., Herberg, S., Petty, W. J., Marini, F., Miller, L., Kucera, G., Dukes, D. K., Ruiz, J., Skardal, A., *et al.* (2019). Pleural Effusion Aspirate for Use in 3D Lung Cancer Modeling and Chemotherapy Screening. *ACS Biomaterials Science and Engineering*, 5(4):1937–1943.

Mazzoleni, G., Di Lorenzo, D., and Steimberg, N. (2009). Modelling tissues in 3D: The next future of pharmaco-toxicology and food research? *Genes and Nutrition*, 4(1):13–22.

McCracken, K. W., Catá, E. M., Crawford, C. M., Sinagoga, K. L., Schumacher, M., Rockich, B. E., Tsai, Y. H., Mayhew, C. N., Spence, J. R., Zavros, Y., *et al.* (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature*, 516(7531):400–404.

McGranahan, N. and Swanton, C. (2017). Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell*, 168(4):613–628.

McQualter, J. L., Yuen, K., Williams, B., and Bertoncello, I. (2010). Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. *Proceedings of the National Academy of Sciences*, 107(4):1414–1419.

Mehregan, A. H. and Pinkus, H. (1965). Life History of Organoid Nevi: Special Reference to Nevus Sebaceus of Jadassohn. *Archives of Dermatology*, 91(6):574–588.

Mendoza, M. and Khanna, C. (2009). Revisiting the seed and soil in cancer metastasis. *The International Journal of Biochemistry & Cell Biology*, 41(7):1452–1462.

Metzger, D. E., Stahlman, M. T., and Shannon, J. M. (2008). Misexpression of ELF5

disrupts lung branching and inhibits epithelial differentiation. *Developmental Biology*, 320(1):149–160.

Meuwissen, R., Linn, S. C., Van Der Valk, M., Mooi, W. J., and Berns, A. (2001). Mouse model for lung tumorigenesis through Cre/lox controlled sporadic activation of the K-ras oncogene. *Oncogene*, 20(45):6551–6558.

Meylan, E., Dooley, A. L., Feldser, D. M., Shen, L., Turk, E., Ouyang, C., and Jacks, T. (2009). Requirement for NF-B signalling in a mouse model of lung adenocarcinoma. *Nature*, 462(7269):104–107.

Micke, P. and Östman, A. (2004). Tumour-stroma interaction: Cancer-associated fibroblasts as novel targets in anti-cancer therapy? *Lung Cancer*, 45(SUPPL. 2):S163–S175.

Micke, P. and Östman, A. (2005). Exploring the tumour environment: Cancer-associated fibroblasts as targets in cancer therapy. *Expert Opinion on Therapeutic Targets*, 9(6):1217–1233.

Mitchell, S., Vargas, J., and Hoffmann, A. (2016). Signaling via the NFκB system. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 8(3):227–241.

Mittal, V. (2018). Epithelial Mesenchymal Transition in Tumor Metastasis. *Annual Review of Pathology: Mechanisms of Disease*, 13(1):395–412.

Moffitt, J. R., Bambah-Mukku, D., Eichhorn, S. W., Vaughn, E., Shekhar, K., Perez, J. D., Rubinstein, N. D., Hao, J., Regev, A., Dulac, C., *et al.* (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, 362(6416):eaau5324.

Montoro, D. T., Haber, A. L., Biton, M., Vinarsky, V., Lin, B., Birket, S. E., Yuan, F., Chen, S., Leung, H. M., Villoria, J., *et al.* (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature*, 560(7718):319–324.

Moreno, C. S. (2020). SOX4: The unappreciated oncogene. *Seminars in Cancer Biology*, 67:57–64.

Morrisey, E. E. and Hogan, B. L. (2010). Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. *Developmental Cell*, 18(1):8–23.

Mucenski, M. L., Wert, S. E., Nation, J. M., Loudy, D. E., Huelsken, J., Birchmeier, W., Morrisey, E. E., and Whitsett, J. A. (2003). β-catenin is required for specification of proximal/distal cell fate during lung morphogenesis. *Journal of Biological Chemistry*, 278(41):40231–40238.

Muenst, S., Läubli, H., Soysal, S. D., Zippelius, A., Tzankov, A., and Hoeller, S. (2016). The immune system and cancer evasion strategies: Therapeutic concepts. *Journal of Internal Medicine*, 279(6):541–562.

Mullenders, J., de Jongh, E., Brousali, A., Roosen, M., Blom, J. P., Begthel, H., Korving, J., Jonges, T., Kranenburg, O., Meijer, R., *et al.* (2019). Mouse and human urothelial cancer organoids: A tool for bladder cancer research. *Proceedings of the National Academy of Sciences of the United States of America*, 116(10):4567–4574.

Nabhan, A. N., Brownfield, D. G., Harbury, P. B., Krasnow, M. A., and Desai, T. J. (2018). Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science*, 359(6380):1118–1123.

Nakanishi, H., Matsumoto, S., Iwakawa, R., Kohno, T., Suzuki, K., Tsuta, K., Matsuno, Y., Noguchi, M., Shimizu, E., and Yokota, J. (2009). Whole genome comparison of allelic imbalance between noninvasive and invasive small-sized lung adenocarcinomas. *Cancer Research*, 69(4):1615–1623.

Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell*, 10(6):771–785.

Nguyen, D. X., Chiang, A. C., Zhang, X. H., Kim, J. Y., Kris, M. G., Ladanyi, M., Gerald, W. L., and Massagué, J. (2009). WNT/TCF Signaling through LEF1 and HOXB9 Mediates Lung Adenocarcinoma Metastasis. *Cell*, 138(1):51–62.

Nikolić, M. Z., Caritg, O., Jeng, Q., Johnson, J. A., Sun, D., Howell, K. J., Brady, J. L.,

Laresgoiti, U., Allen, G., Butler, R., *et al.* (2017). Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *eLife*, 6.

Ohuchi, H., Hori, Y., Yamasaki, M., Harada, H., Sekine, K., Kato, S., and Itoh, N. (2000). FGF10 acts as a major ligand for FGF receptor 2 IIIb in mouse multi-organ development. *Biochemical and Biophysical Research Communications*, 277(3):643–649.

Ouadah, Y., Rojas, E. R., Riordan, D. P., Capostagno, S., Kuo, C. S., and Krasnow, M. A. (2019). Rare Pulmonary Neuroendocrine Cells Are Stem Cells Regulated by Rb, p53, and Notch. *Cell*, 179(2):403–416.e23.

Pacheco-Pinedo, E. C., Durham, A. C., Stewart, K. M., Goss, A. M., Lu, M. M., DeMayo, F. J., and Morrisey, E. E. (2011). Wnt/β-catenin signaling accelerates mouse lung tumorigenesis by imposing an embryonic distal progenitor phenotype on lung epithelium. *Journal of Clinical Investigation*, 121(5):1935–1945.

Pasquale, E. B. (2010). Eph receptors and ephrins in cancer: Bidirectional signalling and beyond. *Nature Reviews Cancer*, 10(3):165–180.

Pauli, C., Hopkins, B. D., Prandi, D., Shaw, R., Fedrizzi, T., Sboner, A., Sailer, V., Augello, M., Puca, L., Rosati, R., *et al.* (2017). Personalized in vitro and in vivo cancer models to guide precision medicine. *Cancer Discovery*, 7(5):462–477.

Paulsson, J. and Micke, P. (2014). Prognostic relevance of cancer-associated fibroblasts in human cancer. *Seminars in Cancer Biology*, 25:61–68.

Pepicelli, C. V., Lewis, P. M., and McMahon, A. P. (1998). Sonic hedgehog regulates branching morphogenesis in the mammalian lung. *Current Biology*, 8(19):1083–1086.

Petty, A., Myshkin, E., Qin, H., Guo, H., Miao, H., Tochtrop, G. P., Hsieh, J.-T., Page, P., Liu, L., Lindner, D. J., *et al.* (2012). A Small Molecule Agonist of EphA2 Receptor Tyrosine Kinase Inhibits Tumor Cell Migration In Vitro and Prostate Cancer Metastasis In Vivo. *PLoS ONE*, 7(8):e42120.

Pickl, M. and Ries, C. H. (2009). Comparison of 3D and 2D tumor models reveals enhanced HER2 activation in 3D associated with an increased response to trastuzumab. *Oncogene*, 28(3):461–468.

Pietras, K. and Östman, A. (2010). Hallmarks of cancer: Interactions with the tumor stroma. *Experimental Cell Research*, 316(8):1324–1331.

Pillai, S., Rizwani, W., Li, X., Rawal, B., Nair, S., Schell, M. J., Bepler, G., Haura, E., Coppola, D., and Chellappan, S. (2011). ID1 Facilitates the Growth and Metastasis of Non-Small Cell Lung Cancer in Response to Nicotinic Acetylcholine Receptor and Epidermal Growth Factor Receptor Signaling. *Molecular and Cellular Biology*, 31(14):3052–3067.

Plasschaert, L. W., Žilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., Klein, A. M., and Jaffe, A. B. (2018). A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature*, 560(7718):377–381.

Politi, K., Zakowski, M. F., Fan, P. D., Schonfeld, E. A., Pao, W., and Varmus, H. E. (2006). Lung adenocarcinomas induced in mice by mutant EGF receptors found in human lung cancers respond to a tyrosine kinase inhibitor or to down-regulation of the receptors. *Genes and Development*, 20(11):1496–1510.

Post, Y., Puschhof, J., Beumer, J., Kerkkamp, H. M., de Bakker, M. A., Slagboom, J., de Barbanson, B., Wevers, N. R., Spijkers, X. M., Olivier, T., *et al.* (2020). Snake Venom Gland Organoids. *Cell*, 180(2):233–247.e21.

Price, J. E. (2014). Spontaneous and Experimental Metastasis Models: Nude Mice. In *Dwek M., Schumacher U., Brooks S. (eds) Metastasis Research Protocols. Methods in Molecular Biology (Methods and Protocols), vol 1070. Humana Press, New York, NY*, pages 223–233.

Qian, X., Nguyen, H. N., Song, M. M., Hadiono, C., Ogden, S. C., Hammack, C., Yao, B., Hamersky, G. R., Jacob, F., Zhong, C., *et al.* (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell*, 165(5):1238–1254.

Que, J., Luo, X., Schwartz, R. J., and Hogan, B. L. (2009). Multiple roles for Sox2 in the developing and adult mouse trachea. *Development*, 136(11):1899–1907.

Raj, B., Wagner, D. E., McKenna, A., Pandey, S., Klein, A. M., Shendure, J., Gagnon, J. A., and Schier, A. F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nature Biotechnology*, 36(5):442–450.

Ramamoorthy, P., Thomas, S. M., Kaushik, G., Subramaniam, D., Chastain, K. M., Dhar, A., Tawfik, O., Kasi, A., Sun, W., Ramalingam, S., *et al.* (2019). Metastatic tumor-in-A-Dish, a novel multicellular organoid to study lung colonization and predict therapeutic response. *Cancer Research*, 79(7):1681–1695.

Rami-Porta, R., Asamura, H., Travis, W. D., and Rusch, V. W. (2017). Lung cancer - major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA: A Cancer Journal for Clinicians*, 67(2):138–155.

Ramkumar, K., Samanta, S., Kyani, A., Yang, S., Tamura, S., Ziemke, E., Stuckey, J. A., Li, S., Chinnaswamy, K., Otake, H., *et al.* (2016). Mechanistic evaluation and transcriptional signature of a glutathione S-transferase omega 1 inhibitor. *Nature Communications*, 7.

Rao, C. V., Janakiram, N. B., Madka, V., Kumar, G., Scott, E. J., Pathuri, G., Bryant, T., Kutche, H., Zhang, Y., Biddick, L., *et al.* (2016). Small-molecule inhibition of GCNT3 disrupts mucin biosynthesis and malignant cellular behaviors in pancreatic cancer. *Cancer Research*, 76(7):1965–1974.

Rawlins, E. L., Clark, C. P., Xue, Y., and Hogan, B. L. M. (2009a). The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development*, 136(22):3741–3745.

Rawlins, E. L. and Hogan, B. L. (2008). Ciliated epithelial cell lifespan in the mouse trachea and lung. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 295(1):L231.

Rawlins, E. L., Okubo, T., Xue, Y., Brass, D. M., Auten, R. L., Hasegawa, H., Wang, F., and Hogan, B. L. (2009b). The Role of Scgb1a1+ Clara Cells in the Long-Term Maintenance and Repair of Lung Airway, but Not Alveolar, Epithelium. *Cell Stem Cell*, 4(6):525–534.

Riely, G. J., Marks, J., and Pao, W. (2009). KRAS mutations in non-small cell lung cancer.

Proceedings of the American Thoracic Society, 6(2):201–205.

Riordan, J. R., Rommens, J. M., Kerem, B. S., Alon, N. O., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S. I., Plavsic, N., Chou, J. L., *et al.* (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science*, 245(4922):1066–1073.

Roberts, P. J. and Der, C. J. (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene*, 26(22):3291–3310.

Rock, J. R., Barkauskas, C. E., Cronce, M. J., Xue, Y., Harris, J. R., Liang, J., Noble, P. W., and Hogan, B. L. (2011a). Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proceedings of the National Academy of Sciences of the United States of America*, 108(52):E1475.

Rock, J. R., Gao, X., Xue, Y., Randell, S. H., Kong, Y. Y., and Hogan, B. L. (2011b). Notch-dependent differentiation of adult airway basal stem cells. *Cell Stem Cell*, 8(6):639–648.

Rock, J. R. and Hogan, B. L. (2011). Epithelial Progenitor Cells in Lung Development, Maintenance, Repair, and Disease. *Annual Review of Cell and Developmental Biology*, 27(1):493–512.

Rock, J. R., Onaitis, M. W., Rawlins, E. L., Lu, Y., Clark, C. P., Xue, Y., Randell, S. H., and Hogan, B. L. (2009). Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31):12771–12775.

Rockich, B. E., Hrycaj, S. M., Shih, H. P., Nagy, M. S., Ferguson, M. A. H., Kopp, J. L., Sander, M., Wellik, D. M., and Spence, J. R. (2013). Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. *Proceedings of the National Academy of Sciences*, 110(47):E4456–E4464.

Rodríguez, C. I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A. F., and Dymecki, S. M. (2000). High-efficiency deleter mice show that FLPe is an alternative to Cre- loxP. *Nature Genetics*, 25(2):139–140.

Rogers, D. F. (1994). Airway goblet cells: responsive and adaptable front-line defenders. *Eur Respir J.*, 7(9):1690–706.

Rokicki, W., Rokicki, M., Wojtacha, J., and Dzeljijli, A. (2016). The role and importance of club cells (Clara cells) in the pathogenesis of some respiratory diseases. *Kardiochirurgia i Torakochirurgia Polska*, 13(1):26–30.

Sachs, N., de Ligt, J., Kopper, O., Gogola, E., Bounova, G., Weeber, F., Balgobind, A. V., Wind, K., Gracanin, A., Begthel, H., *et al.* (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell*, 172(1-2):373–386.e10.

Sachs, N., Papaspyropoulos, A., Zomer-van Ommen, D. D., Heo, I., Böttinger, L., Klay, D., Weeber, F., Huelsz-Prince, G., Iakobachvili, N., Amatngalim, G. D., *et al.* (2019). Long-term expanding human airway organoids for disease modeling. *The EMBO Journal*, 38(4).

Sainz de Aja, J., Dost, A. F. M., and Kim, C. F. (2020). Alveolar progenitor cells and the origin of lung cancer. *Journal of Internal Medicine*, page joim.13201.

Salahudeen, A. A., Choi, S. S., Rustagi, A., Zhu, J., van Unen, V., de la O, S. M., Flynn, R. A., Margalef-Català, M., Santos, A. J., Ju, J., *et al.* (2020). Progenitor identification and SARS-CoV-2 infection in human distal lung organoids. *Nature*, 588(7839):670–675.

Salwig, I., Spitznagel, B., Vazquez-Armendariz, A. I., Khalooghi, K., Guenther, S., Herold, S., Szibor, M., and Braun, T. (2019). Bronchioalveolar stem cells are a main source for regeneration of distal lung epithelia in vivo . *The EMBO Journal*, 38(12).

Sato, T., Stange, D. E., Ferrante, M., Vries, R. G., van Es, J. H., Van Den Brink, S., Van Houdt, W. J., Pronk, A., Van Gorp, J., Siersema, P. D., *et al.* (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*, 141(5):1762–1772.

Sato, T., Vries, R. G., Snippert, H. J., van de Wetering, M., Barker, N., Stange, D. E., van Es, J. H., Abo, A., Kujala, P., Peters, P. J., *et al.* (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459(7244):262–265.

Schlieve, C. R., Fowler, K. L., Thornton, M., Huang, S., Hajjali, I., Hou, X., Grubbs, B., Spence, J. R., and Grikscheit, T. C. (2017). Neural Crest Cell Implantation Restores Enteric Nervous System Function and Alters the Gastrointestinal Transcriptome in Human Tissue-Engineered Small Intestine. *Stem Cell Reports*, 9(3):883–896.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675.

Schutgens, F. and Clevers, H. (2020). Human Organoids: Tools for Understanding Biology and Treating Diseases. *Annual Review of Pathology: Mechanisms of Disease*, 15(1):211–234.

Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., Yagishita, N., Matsui, D., Koga, Y., Itoh, N., *et al.* (1999). Fgf10 is essential for limb and lung formation. *Nature Genetics*, 21(1):138–141.

Shintani, Y., Fujiwara, A., Kimura, T., Kawamura, T., Funaki, S., Minami, M., and Okumura, M. (2016). IL-6 secreted from Cancer-Associated fibroblasts mediates chemoresistance in NSCLC by increasing epithelial-mesenchymal transition signaling. *Journal of Thoracic Oncology*, 11(9):1482–1492.

Silva, D. M. G., Nardiello, C., Pozarska, A., and Morty, R. E. (2015). Recent advances in the mechanisms of lung alveolarization and the pathogenesis of bronchopulmonary dysplasia. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 309(11):L1239–L1272.

Singh, I., Mehta, A., Contreras, A., Boettger, T., Carraro, G., Wheeler, M., Cabrera-Fuentes, H. A., Bellusci, S., Seeger, W., Braun, T., *et al.* (2014). Hmga2 is required for canonical WNT signaling during lung development. *BMC Biology*, 12.

Smith, E. and Cochrane, W. J. (1946). CYSTIC ORGANOID TERATOMA: (Report of a Case). *Canadian Medical Association journal*, 55(2):151–2.

Snyder, E. L., Watanabe, H., Magendantz, M., Hoersch, S., Chen, T. A., Wang, D. G., Crowley, D., Whittaker, C. A., Meyerson, M., Kimura, S., *et al.* (2013). Nkx2-1 Represses a Latent Gastric Differentiation Program in Lung Adenocarcinoma. *Molecular Cell*, 50(2):185–199.

Song, W., Hwang, Y., Youngblood, V. M., Cook, R. S., Balko, J. M., Chen, J., and Brantley-Sieders, D. M. (2017). Targeting EphA2 impairs cell cycle progression and growth of basal-like/triple-negative breast cancers. *Oncogene*, 36(40):5620–5630.

Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J. P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. *Nature Biotechnology*, 36(5):469–473.

Stenhouse, G., Fyfe, N., King, G., Chapman, A., and Kerr, K. M. (2004). Thyroid transcription factor 1 in pulmonary adenocarcinoma. *Journal of Clinical Pathology*, 57(4):383–387.

Stewart, D. J. (2014). Wnt Signaling Pathway in Non-Small Cell Lung Cancer. *JNCI Journal of the National Cancer Institute*, 106(1):djt356–djt356.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, 177(7):1888–1902.e21.

Stuart, T. and Satija, R. (2019). Integrative single-cell analysis. *Nature Reviews Genetics*, 20(5):257–272.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., *et al.* (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43):15545–15550.

Sugimoto, H., Mundel, T. M., Kieran, M. W., and Kalluri, R. (2006). Identification of fibroblast heterogeneity in the tumor microenvironment. *Cancer Biology and Therapy*, 5(12):1640–1646.

Sullivan, K. D., Galbraith, M. D., Andrysik, Z., and Espinosa, J. M. (2018). Mechanisms of transcriptional regulation by p53. *Cell Death and Differentiation*, 25(1):133–143.

Sung, P. J., Rama, N., Imbach, J., Fiore, S., Ducarouge, B., Neves, D., Chen, H. W., Bernard,

D., Yang, P. C., Bernet, A., *et al.* (2019). Cancer-associated fibroblasts produce netrin-1 to control cancer cell plasticity. *Cancer Research*, 79(14):3651–3661.

Sutherland, K. D., Song, J.-Y., Kwon, M. C., Proost, N., Zevenhoven, J., and Berns, A. (2014). Multiple cells-of-origin of mutant K-Ras-induced mouse lung adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, 111(13):4952–7.

Swarm, R. (1963). TRANSPLANTATION OF A MURINE CHONDROSARCOMA IN MICE OF DIFFERENT INBRED STRAINS. *Journal of the National Cancer Institute*, 31:953–975.

Takahashi, K. and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4):663–676.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R. R., Ueno, Y., Zheng, Y. W., Koike, N., *et al.* (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature*, 499(7459):481–484.

Tammela, T., Sanchez-Rivera, F. J., Cetinbas, N. M., Wu, K., Joshi, N. S., Helenius, K., Park, Y., Azimi, R., Kerper, N. R., Wesselhoeft, R. A., *et al.* (2017). A Wnt-producing niche drives proliferative potential and progression in lung adenocarcinoma. *Nature*, 545(7654):355–359.

Tang, X., Liu, D., Shishodia, S., Ozburn, N., Behrens, C., Lee, J. J., Waun, K. H., Aggarwal, B. B., and Wistuba, I. I. (2006). Nuclear factor-κB (NF-κB) is frequently expressed in lung cancer and preneoplastic lesions. *Cancer*, 107(11):2637–2646.

Tata, P. R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B. M., Vinarsky, V., Cho, J. L., Breton, S., Sahay, A., *et al.* (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature*, 503(7475):218–223.

Tompkins, D. H., Besnard, V., Lange, A. W., Keiser, A. R., Wert, S. E., Bruno, M. D., and Whitsett, J. A. (2011). Sox2 activates cell proliferation and differentiation in the respiratory epithelium. *American Journal of Respiratory Cell and Molecular Biology*, 45(1):101–110.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg,

S. L., Wold, B. J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5):511–515.

Travaglini, K. J., Nabhan, A. N., Penland, L., Sinha, R., Gillich, A., Sit, R. V., Chang, S., Conley, S. D., Mori, Y., Seita, J., *et al.* (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature*, 587(7835):619–625.

Treutlein, B., Brownfield, D. G., Wu, A. R., Neff, N. F., Mantalas, G. L., Espinoza, F. H., Desai, T. J., Krasnow, M. A., and Quake, S. R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*, 509(7500):371–375.

Van De Wetering, M., Francies, H. E., Francis, J. M., Bounova, G., Iorio, F., Pronk, A., Van Houdt, W., Van Gorp, J., Taylor-Weiner, A., Kester, L., *et al.* (2015). Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*, 161(4):933–945.

Van der Velden, J. L., Bertoncello, I., and McQualter, J. L. (2013). LysoTracker is a marker of differentiated alveolar type II cells. *Respiratory Research*, 14(1).

van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A. J., Burdziak, C., Moon, K. R., Chaffer, C. L., Pattabiraman, D., *et al.* (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell*, 174(3):716–729.e27.

Vaughan, A. E., Brumwell, A. N., Xi, Y., Gotts, J. E., Brownfield, D. G., Treutlein, B., Tan, K., Tan, V., Liu, F. C., Looney, M. R., *et al.* (2015). Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. *Nature*, 517(7536):621–625.

Vaz, M., Hwang, S. Y., Kagiampakis, I., Phallen, J., Patil, A., O'Hagan, H. M., Murphy, L., Zahnow, C. A., Gabrielson, E., Velculescu, V. E., *et al.* (2017). Chronic Cigarette Smoke-Induced Epigenomic Changes Precede Sensitization of Bronchial Epithelial Cells to Single-Step Transformation by KRAS Mutations. *Cancer Cell*, 32(3):360–376.e6.

Vlachogiannis, G., Hedayat, S., Vatsiou, A., Jamin, Y., Fernández-Mateos, J., Khan, K., Lampis, A., Eason, K., Huntingford, I., Burke, R., *et al.* (2018). Patient-derived

organoids model treatment response of metastatic gastrointestinal cancers. *Science*, 359(6378):920–926.

Wadosky, K. M., Wang, Y., Zhang, X., and Goodrich, D. W. (2019). Generation of tumor organoids from genetically engineered mouse models of prostate cancer. *Journal of Visualized Experiments*, 2019(148).

Wang, L., Li, X., Ren, Y., Geng, H., Zhang, Q., Cao, L., Meng, Z., Wu, X., Xu, M., and Xu, K. (2019). Cancer-associated fibroblasts contribute to cisplatin resistance by modulating ANXA3 in lung cancer cells. *Cancer Science*, 110(5):1609–1620.

Wang, X., Allen, W. E., Wright, M. A., Sylwestrak, E. L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., *et al.* (2018a). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, 361(6400):eaat5691.

Wang, Y., Tang, Z., Huang, H., Li, J., Wang, Z., Yu, Y., Zhang, C., Li, J., Dai, H., Wang, F., *et al.* (2018b). Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. *Proceedings of the National Academy of Sciences*, 115(10):2407–2412.

Weaver, T. E. and Whitsett, J. A. (1991). Function and regulation of expression of pulmonary surfactant-associated proteins. *Biochemical Journal*, 273(2):249–264.

Weeber, F., Van De Wetering, M., Hoogstraat, M., Dijkstra, K. K., Krijgsman, O., Kuilman, T., Gadellaa-Van Hooijdonk, C. G., Van Der Velden, D. L., Peeper, D. S., Cuppen, E. P., *et al.* (2015). Preserved genetic diversity in organoids cultured from biopsies of human colorectal cancer metastases. *Proceedings of the National Academy of Sciences of the United States of America*, 112(43):13308–13311.

Weibel, E. R. (2015). On the tricks alveolar epithelial cells play to make a good lung. *American Journal of Respiratory and Critical Care Medicine*, 191(5):504–513.

Weibel, E. R., Federspiel, W. J., Fryder-Doffey, F., Hsia, C. C., König, M., Stalder-Navarro, V., and Vock, R. (1993). Morphometric model for pulmonary diffusing capacity I. Membrane diffusing capacity. *Respiration Physiology*, 93(2):125–149.

Weiner, A. I., Jackson, S. R., Zhao, G., Quansah, K. K., Farshchian, J. N., Neupauer, K. M., Littauer, E. Q., Paris, A. J., Liberti, D. C., Scott Worthen, G., *et al.* (2019). Mesenchyme-free expansion and transplantation of adult alveolar progenitor cells: steps toward cell-based regenerative therapies. *npj Regenerative Medicine*, 4(1).

Weir, B. A., Woo, M. S., Getz, G., Perner, S., Ding, L., Beroukhim, R., Lin, W. M., Province, M. A., Kraja, A., Johnson, L. A., *et al.* (2007). Characterizing the cancer genome in lung adenocarcinoma. *Nature*, 450(7171):893–898.

Wheeler, D. A. and Wang, L. (2013). From human genome to cancer genome: The first decade. *Genome Research*, 23(7):1054–1062.

Whitsett, J. A., Kalin, T. V., Xu, Y., and Kalinichenko, V. V. (2019). Building and Regenerating the Lung Cell by Cell. *Physiological Reviews*, 99(1):513–554.

Winslow, M. M., Dayton, T. L., Verhaak, R. G., Kim-Kiselak, C., Snyder, E. L., Feldser, D. M., Hubbard, D. D., Dupage, M. J., Whittaker, C. A., Hoersch, S., *et al.* (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature*, 473(7345):101–104.

Wolf, F. A., Angerer, P., and Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*, 19(1):15.

Wong, A. P., Bear, C. E., Chin, S., Pasceri, P., Thompson, T. O., Huan, L. J., Ratjen, F., Ellis, J., and Rossant, J. (2012). Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. *Nature Biotechnology*, 30(9):876–882.

Workman, M. J., Mahe, M. M., Trisno, S., Poling, H. M., Watson, C. L., Sundaram, N., Chang, C. F., Schiesser, J., Aubert, P., Stanley, E. G., *et al.* (2017). Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. *Nature Medicine*, 23(1):49–59.

World Health Organization (2018). World Health Organization Cancer Fact Sheet.

Xiang, H., Ramil, C. P., Hai, J., Zhang, C., Wang, H., Watkins, A. A., Afshar, R., Georgiev,

P., Sze, M. A., Song, X. S., *et al.* (2020). Cancer-Associated Fibroblasts Promote Immunosuppression by Inducing ROS-Generating Monocytic MDSCs in Lung Squamous Cell Carcinoma. *Cancer immunology research*, 8(4):436–450.

Xu, J., Gong, T., Heng, B. C., and Zhang, C. F. (2017). A systematic review: differentiation of stem cells into functional pericytes. *The FASEB Journal*, 31(5):1775–1786.

Xu, X., Rock, J. R., Lu, Y., Futtner, C., Schwab, B., Guinney, J., Hogan, B. L. M., and Onaitis, M. W. (2012). Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13):4910–5.

Xue, W., Meylan, E., Oliver, T. G., Feldser, D. M., Winslow, M. M., Bronson, R., and Jacks, T. (2011). Response and resistance to NF-κB inhibitors in mouse models of lung adenocarcinoma. *Cancer Discovery*, 1(3):236–247.

Yan, H. H., Siu, H. C., Law, S., Ho, S. L., Yue, S. S., Tsui, W. Y., Chan, D., Chan, A. S., Ma, S., Lam, K. O., *et al.* (2018). A Comprehensive Human Gastric Cancer Organoid Biobank Captures Tumor Subtype Heterogeneity and Enables Therapeutic Screening. *Cell Stem Cell*, 23(6):882–897.e11.

Yang, J. and Weinberg, R. A. (2008). Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis. *Developmental Cell*, 14(6):818–829.

Yates, P. R., Atherton, G. T., Deed, R. W., Norton, J. D., and Sharrocks, A. D. (1999). Id helix–loop–helix proteins inhibit nucleoprotein complex formation by the TCF ETS-domain transcription factors. *The EMBO Journal*, 18(4):968–976.

Ye, J., Findeis-Hosey, J. J., Yang, Q., McMahon, L. A., Yao, J. L., Li, F., and Xu, H. (2011). Combination of napsin A and TTF-1 immunohistochemistry helps in differentiating primary lung adenocarcinoma from metastatic carcinoma in the lung. *Applied Immunohistochemistry and Molecular Morphology*, 19(4):313–317.

Yin, X., Mead, B. E., Safaee, H., Langer, R., Karp, J. M., and Levy, O. (2016). Engineering Stem Cell Organoids. *Cell Stem Cell*, 18(1):25–38.

Young, J. J., Roffers, S., Ries, L., Fritz, A., and Hurlbut, A. (2001). SEER Summary Staging Manual - 2000: Codes and Coding Instructions .

Yu, T., Chen, X., Zhang, W., Liu, J., Avdiushko, R., Napier, D. L., Liu, A. X., Neltner, J. M., Wang, C., Cohen, D., *et al.* (2016). KLF4 regulates adult lung tumor-initiating cells and represses K-Ras-mediated lung cancer. *Cell Death & Differentiation*, 23(2):207–215.

Zacharias, W. J., Frank, D. B., Zepp, J. A., Morley, M. P., Alkhaleel, F. A., Kong, J., Zhou, S., Cantu, E., and Morrisey, E. E. (2018). Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature*, 555(7695):251–255.

Zepp, J. A., Zacharias, W. J., Frank, D. B., Cavanaugh, C. A., Zhou, S., Morley, M. P., and Morrisey, E. E. (2017). Distinct Mesenchymal Lineages and Niches Promote Epithelial Self-Renewal and Myofibrogenesis in the Lung. *Cell*, 170(6):1134–1148.e10.

Zhan, C., Yan, L., Wang, L., Sun, Y., Wang, X., Lin, Z., Zhang, Y., Shi, Y., Jiang, W., and Wang, Q. (2015). Identification of immunohistochemical markers for distinguishing lung adenocarcinoma from squamous cell carcinoma. *Journal of Thoracic Disease*, 7(8):1398–1405.

Zhang, H., Brainson, C. F., Koyama, S., Redig, A. J., Chen, T., Li, S., Gupta, M., Garcia-de alba, C., Paschini, M., Herter-sprie, G. S., *et al.* (2017a). Lkb1 inactivation drives lung cancer lineage switching governed by Polycomb Repressive Complex 2. *Nature Communications*, 8:1–14.

Zhang, Y., Rath, N., Hannenhalli, S., Wang, Z., Cappola, T., Kimura, S., Atochina-Vasserman, E., Lu, M. M., Beers, M. F., and Morrisey, E. E. (2007). GATA and Nkx factors synergistically regulate tissue-specific gene expression and development in vivo. *Development*, 134(1):189–198.

Zhang, Z., Newton, K., Kummerfeld, S. K., Webster, J., Kirkpatrick, D. S., Phu, L., Eastham-Anderson, J., Liu, J., Lee, W. P., Wu, J., *et al.* (2017b). Transcription factor Etv5 is essential for the maintenance of alveolar type II cells. *Proceedings of the National Academy of Sciences of the United States of America*, 114(15):3903–3908.

Zhao, S., Guo, W., Li, J., Yu, W., Guo, T., Deng, W., and Gu, C. (2016). High expression

of Y-box-binding protein 1 correlates with poor prognosis and early recurrence in patients with small invasive lung adenocarcinoma. *OncoTargets and Therapy*, page 2683.

Zheng, H., Zhan, Y., Zhang, Y., Liu, S., Lu, J., Yang, Y., Wen, Q., and Fan, S. (2019). Elevated expression of G3BP1 associates with YB1 and p-AKT and predicts poor prognosis in nonsmall cell lung cancer patients after surgical resection. *Cancer Medicine*, 8(16):6894–6903.

Zhou, C. H., Ye, L. P., Ye, S. X., Li, Y., Zhang, X. Y., Xu, X. Y., and Gong, L. Y. (2012). Clinical significance of SOX9 in human non-small cell lung cancer progression and overall patient survival. *Journal of Experimental and Clinical Cancer Research*, 31(1).

Zuo, W., Zhang, T., Wu, D. Z., Guan, S. P., Liew, A.-A., Yamamoto, Y., Wang, X., Lim, S. J., Vincent, M., Lessard, M., *et al.* (2015). p63+Krt5+ distal airway stem cells are essential for lung regeneration. *Nature*, 517(7536):616–620.

Appendix

Table 14: Commonly down- and upregulated genes in RNA-Seq data
Genes that were within the top 100 down- and upregulated gene lists of KY-Dif and KPY-Dif and shared between the two, as determined by RNA-Seq.

Commonly downregulated	Commonly upregulated
Sec14l3	Procr
Gsta3	Ecm1
Kcnk2	F2r
Nkd1	Hnf4a
Cd74	Emp1
Nrep	B4galt6
Gstt1	Anxa2
Ldhb	Phgdh
Lyz2	Prss22
S100g	Slc2a1
Ces1g	Tpi1
Cyp4b1	Pfkl
Mlc1	St8sia6
Gas6	Ldha
Aox3	Errfi1
Itm2a	Ero1l
Mettl7a1	Pglyrp1
2610028H24Rik	Fabp5
Creg1	Spp1
H2-Eb1	Ptges
H2-Aa	Msln
Cytip	Tnfrsf23
Serpinb9	Hmga2
Tppp3	Sox9
Galnt18	Ly6a
Gsap	F3
Sepp1	Kcnq1
Scnn1b	Lad1
Ak1	Pthlh

Hnmt	Itga6
Adcy7	Aqp3
H2-Ab1	Tigit
Tspan11	Krt7
Ptn	Slc16a3
Cyp4v3	Adss
Tmem116	B3gnt3
Pmp22	Fam107b
Ces1d	Tubb6
Cyp2b10	Mcpt2
Lgi3	
Mme	
Scnn1g	
Dynlrb2	
Akr1c14	
H2-DMb1	
Ppp2r2b	
Fgf1	
Ddo	
Col6a1	

Table 15: Top 100 upregulated genes in KY scRNA-Seq clusters
$C_{KY}1$ represents the control cluster, $C_{KY}0$ and $C_{KY}2$ the Cre clusters. Genes are ranked by p values with lower values ranked higher.

rank	$C_{KY}0$	$C_{KY}1$	$C_{KY}2$
1	Fn1	Sftpc	Ldha
2	Rtn4	Cd74	Psca
3	Krt8	Scd1	Fabp5
4	Clu	Ager	Pkm
5	Krt18	Lpcat1	Ybx1
6	Lgals1	Napsa	Tpi1
7	Tmsb10	H2-Aa	Gapdh
8	Nupr1	H2-Ab1	Krt7
9	Cldn4	Sftpa1	Pglyrp1

10	Actn1	Wbp5	Tff1
11	Rdh10	Slc34a2	2200002D01Rik
12	S100a10	H2-Eb1	Phgr1
13	Pdlim7	Sftpd	Aqp5
14	Ly6e	Ppp1r14c	Sec61g
15	Ctsl	Dram1	Mal
16	S100a11	Gde1	Anxa2
17	Ptgs2	Cadm1	Il33
18	Tnfrsf12a	Sfta2	Rps8
19	S100a6	Ctsc	Mif
20	Cdkn2a	Cxcl15	Ly6d
21	Plin2	Chchd10	Rpl27a
22	Thbs1	Sparc	Tma7
23	Fam107b	Ptprf	Rps2
24	Tpm1	Tmem243	Pgam1
25	Morc4	Lamp3	Rps21
26	Cnn3	Fth1	Glrx
27	Msn	Scd2	Wfdc2
28	Marcks	Mgst1	Rpl36al
29	Sox4	Lyz2	Areg
30	Pmepa1	Npc2	Psat1
31	F3	Egfl6	Gsto1
32	Cd44	Elovl1	Rps27a
33	Basp1	Sftpb	Rpl14
34	Cdkn2b	Mpc1	Rpl41
35	Ktn1	Lgi3	Pthlh
36	Tgif1	Sepp1	Sec61b
37	Msln	Scp2	Cltb
38	Sfn	Tgoln1	Ly6a
39	Tnfrsf23	Fasn	Rpl18
40	Btg1	S100g	Lgals3
41	Epcam	Brd7	Aldoa
42	Spp1	Dpysl2	Rpsa

43	Myl12a	Cd36	Fgfbp1
44	Ccnd2	Creg1	Fam162a
45	Gadd45a	Nrp1	Gm8730
46	Ddah1	Tmem30a	Rpl8
47	F2r	Nkx2-1	Rps6
48	Litaf	Oat	Tspan8
49	Ly6a	Cldn18	Sprr2a3
50	Emp1	H2-K1	Rpl31
51	Ier3	Abca3	2810417H13Rik
52	Flna	Col4a1	Ptges
53	Perp	Slco2a1	Tmsb4x
54	Arl4c	Aldh2	Hmgb2
55	S100a14	Tmbim6	Krt19
56	St13	Alcam	Golm1
57	Tes	Npw	Rps13
58	Cdk6	Etfb	Ywhaz
59	Igf1r	H2-DMb1	Synpr
60	Lurap1l	Atp11a	Id1
61	Marcksl1	Acsl4	H2afz
62	Cald1	Bex4	Higd1a
63	Hsp90ab1	Timp3	Ereg
64	Arpc1b	H2-D1	Anxa10
65	Rps5	Ndufa3	Birc5
66	Fhl2	Ngfrap1	Nhp2
67	Ano1	Mylk	Cks1b
68	Ctgf	Chil1	Rpl11
69	Clic1	Emb	Bnip3
70	Serpine2	Zdhhc3	Tuba1b
71	Anxa1	Lcn2	S100a14
72	Mtpn	Cyb5r3	Rps18
73	Epb41l2	Cd200	Ffar4
74	Myof	Cbr2	Manf
75	Igfbp7	Oxct1	Ube2c

76	Chic2	App	Rps10
77	Vmp1	Retnla	Rps23
78	Mast4	H3f3a	Hsd17b11
79	Zak	Psme1	Slc7a11
80	Krt19	Cmtm8	Il4ra
81	Rhoc	Ifi27	Rps17
82	Sat1	Nckap5	Rps11
83	Cd81	Mettl7a1	Npm1
84	Cdkn1a	Lamp1	Slc35d1
85	Ctse	Rnaset2a	Rpl26
86	Arpc5	Bex2	Rpl29
87	Ecm1	Icam1	Rpl13
88	Hn1	Atp1b1	Rplp1
89	Ltbp1	Abcd3	Rps15
90	Rps27l	Cebpa	Rps12
91	Phactr4	Ucp2	Eif5a
92	Gnai2	Ndufb11	Ran
93	Flnb	Sepw1	Sdf2l1
94	Phgdh	Ces1d	Rplp2
95	Cd2ap	Prnp	Ube2s
96	Cfl1	Maob	Mgst3
97	Rpl7	Gstm1	Eef1g
98	Ddit4	Matn4	Rpl28
99	Rpl22l1	Soat1	Sftpb
100	Rpl18a	Gstt1	Cenpa

Table 16: Top 50 upregulated Gene Ontology terms in KY scRNA-Seq clusters
$C_{KY}1$ represents the control cluster, $C_{KY}0$ and $C_{KY}2$ the Cre clusters. GO terms are ranked by p values with lower values ranked higher.

rank	$C_{KY}0$	$C_{KY}1$	$C_{KY}2$
1	regulation of apoptotic process (GO:0042981)	mitochondrial ATP synthesis coupled electron transport (GO:0042775)	translational termination (GO:0006415)

2	regulation of cell migration (GO:0030334)	respiratory electron transport chain (GO:0022904)	translational elongation (GO:0006414)
3	negative regulation of apoptotic process (GO:0043066)	mitochondrial electron transport, NADH to ubiquinone (GO:0006120)	mitochondrial translational elongation (GO:0070125)
4	regulation of cell proliferation (GO:0042127)	mitochondrial respiratory chain complex assembly (GO:0033108)	mitochondrial translational termination (GO:0070126)
5	positive regulation of transcription, DNA-templated (GO:0045893)	NADH dehydrogenase complex assembly (GO:0010257)	mitochondrial translation (GO:0032543)
6	positive regulation of gene expression (GO:0010628)	mitochondrial respiratory chain complex I biogenesis (GO:0097031)	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile (GO:0000377)
7	positive regulation of angiogenesis (GO:0045766)	mitochondrial respiratory chain complex I assembly (GO:0032981)	mRNA processing (GO:0006397)
8	positive regulation of vasculature development (GO:1904018)	mitochondrial ATP synthesis coupled proton transport (GO:0042776)	mRNA splicing, via spliceosome (GO:0000398)
9	vascular endothelial growth factor receptor signaling pathway (GO:0048010)	fatty acid beta-oxidation (GO:0006635)	DNA metabolic process (GO:0006259)
10	positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	cellular respiration (GO:0045333)	establishment of protein localization to mitochondrion (GO:0072655)

11	cellular protein modification process (GO:0006464)	ATP synthesis coupled proton transport (GO:0015986)	regulation of ubiquitin protein ligase activity (GO:1904666)
12	negative regulation of protein ubiquitination (GO:0031397)	ATP biosynthetic process (GO:0006754)	ribosomal large subunit assembly (GO:0000027)
13	actin filament organization (GO:0007015)	cristae formation (GO:0042407)	glycolytic process through glucose-6-phosphate (GO:0061620)
14	regulation of transcription from RNA polymerase II promoter (GO:0006357)	purine ribonucleoside triphosphate biosynthetic process (GO:0009206)	canonical glycolysis (GO:0061621)
15	extracellular matrix organization (GO:0030198)	aerobic respiration (GO:0009060)	glucose catabolic process to pyruvate (GO:0061718)
16	regulation of angiogenesis (GO:0045765)	inner mitochondrial membrane organization (GO:0007007)	nucleobase-containing small molecule interconversion (GO:0015949)
17	positive regulation of cellular amide metabolic process (GO:0034250)	peptide metabolic process (GO:0006518)	nuclear transport (GO:0051169)
18	positive regulation of cell aging (GO:0090343)	secondary alcohol biosynthetic process (GO:1902653)	nucleocytoplasmic transport (GO:0006913)
19	regulation of I-kappaB kinase/NF-kappaB signaling (GO:0043122)	iron ion homeostasis (GO:0055072)	histone exchange (GO:0043486)
20	regulation of protein ubiquitination (GO:0031396)	fatty acid biosynthetic process (GO:0006633)	centromere complex assembly (GO:0034508)
21	negative regulation of cellular process (GO:0048523)	cellular iron ion homeostasis (GO:0006879)	negative regulation of ubiquitin protein ligase activity (GO:1904667)

22	positive regulation of apoptotic signaling pathway (GO:2001235)	regulation of cholesterol biosynthetic process (GO:0045540)	G1/S transition of mitotic cell cycle (GO:0000082)
23	positive regulation of viral life cycle (GO:1903902)	cellular transition metal ion homeostasis (GO:0046916)	RNA export from nucleus (GO:0006405)
24	regulation of cell-matrix adhesion (GO:0001952)	mitochondrial electron transport, cytochrome c to oxygen (GO:0006123)	regulation of mitotic cell cycle phase transition (GO:1901990)
25	positive regulation of intracellular signal transduction (GO:1902533)	glutathione metabolic process (GO:0006749)	mitotic sister chromatid segregation (GO:0000070)
26	epiboly involved in wound healing (GO:0090505)	fatty acid metabolic process (GO:0006631)	nuclear export (GO:0051168)
27	positive regulation of apoptotic process (GO:0043065)	glutathione derivative biosynthetic process (GO:1901687)	glycolytic process (GO:0006096)
28	protein stabilization (GO:0050821)	glutathione derivative metabolic process (GO:1901685)	anaphase-promoting complex-dependent catabolic process (GO:0031145)
29	cell-matrix adhesion (GO:0007160)	cholesterol metabolic process (GO:0008203)	ATP generation from ADP (GO:0006757)
30	negative regulation of protein modification by small protein conjugation or removal (GO:1903321)	organonitrogen compound biosynthetic process (GO:1901566)	protein targeting to mitochondrion (GO:0006626)
31	platelet aggregation (GO:0070527)	fatty-acyl-CoA biosynthetic process (GO:0046949)	DNA replication-independent nucleosome assembly (GO:0006336)

32	positive regulation of cellular biosynthetic process (GO:0031328)	energy coupled proton transport, down electrochemical gradient (GO:0015985)	chromatin remodeling at centromere (GO:0031055)
33	regulated exocytosis (GO:0045055)	purine ribonucleotide biosynthetic process (GO:0009152)	cell cycle G1/S phase transition (GO:0044843)
34	positive regulation of cellular senescence (GO:2000774)	regulation of cholesterol metabolic process (GO:0090181)	termination of RNA polymerase II transcription (GO:0006369)
35	negative regulation of intracellular signal transduction (GO:1902532)	acetyl-CoA metabolic process (GO:0006084)	nucleic acid metabolic process (GO:0090304)
36	positive regulation of cell migration (GO:0030335)	fatty acid oxidation (GO:0019395)	regulation of cell cycle process (GO:0010564)
37	negative regulation of cell proliferation (GO:0008285)	regulation of lipid metabolic process (GO:0019216)	kinetochore organization (GO:0051383)
38	cell junction assembly (GO:0034329)	regulation of protein binding (GO:0043393)	regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle (GO:0051439)
39	Ras protein signal transduction (GO:0007265)	fatty acid beta-oxidation using acyl-CoA dehydrogenase (GO:0033539)	positive regulation of ubiquitin protein ligase activity (GO:1904668)
40	transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)	lipid biosynthetic process (GO:0008610)	CENP-A containing nucleosome assembly (GO:0034080)

41	actin cytoskeleton reorganization (GO:0031532)	regulation of steroid biosynthetic process (GO:0050810)	CENP-A containing chromatin organization (GO:0061641)
42	positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway (GO:1901030)	lytic vacuole organization (GO:0080171)	pyruvate metabolic process (GO:0006090)
43	Rho protein signal transduction (GO:0007266)	very long-chain fatty acid metabolic process (GO:0000038)	regulation of mitotic nuclear division (GO:0007088)
44	positive regulation of response to stimulus (GO:0048584)	chemical homeostasis within a tissue (GO:0048875)	positive regulation of ubiquitin-protein transferase activity (GO:0051443)
45	release of cytochrome c from mitochondria (GO:0001836)	surfactant homeostasis (GO:0043129)	mRNA export from nucleus (GO:0006406)
46	homotypic cell-cell adhesion (GO:0034109)	fatty acid catabolic process (GO:0009062)	DNA-templated transcription, termination (GO:0006353)
47	negative regulation of proteolysis involved in cellular protein catabolic process (GO:1903051)	negative regulation of protein binding (GO:0032091)	protein import into mitochondrial matrix (GO:0030150)
48	positive regulation of JNK cascade (GO:0046330)	lysosome organization (GO:0007040)	positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition (GO:0051437)

49	viral life cycle (GO:0019058)	positive regulation of peptidase activity (GO:0010952)	positive regulation of cell cycle process (GO:0090068)
50	regulation of spindle assembly (GO:0090169)	glycogen metabolic process (GO:0005977)	mRNA-containing ribonucleoprotein complex export from nucleus (GO:0071427)

Table 17: Top 100 upregulated genes in Y control scRNA-Seq time course clusters
C_Y3 represents the root cluster, C_Y5 the AT1 transitioning cluster, C_Y6 the AT1-like cluster, C_Y7 the AT2 transitioning cluster, and C_Y1 the AT2-like cluster. Genes are ranked by p values with lower values ranked higher.

rank	C_Y1	C_Y3	C_Y5	C_Y6	C_Y7
1	Scd1	Rps2	Clu	Ager	Cks1b
2	Sftpc	Ybx1	Zfp36l1	Sparc	Ptma
3	Lyz2	Npm1	Tmem176b	Cldn18	Rpl13a
4	Fasn	Rplp1	Sat1	Emp2	Prdx4
5	Slc34a2	Ly6a	Tspo	Hopx	Ran
6	Chil1	Rpl14	Tmem176a	Clic5	Tubb5
7	Lamp3	Rpsa	Btg2	Cryab	Rpl5
8	Hc	Rps23	Jund	Prdx6	H2afz
9	Rnase4	Rpl12	H3f3a	Akap5	Stmn1
10	Sepp1	Ran	S100a11	Ahnak	Calr
11	Sfta2	Krt7	Cd24a	Crip2	Dtymk
12	Cxcl15	Rplp2	Cldn4	Hs2st1	Ybx1
13	Elovl1	Rps14	Ctsl	Col4a4	Rpl12
14	Lrg1	S100a6	Ubb	Col4a3	Rps6
15	Lcn2	Rpl7	Ier3	Fbln5	Hspd1
16	Scp2	Rpl27a	Ly6e	Ndnf	Rplp1
17	Abca3	Rps18	Qsox1	Pmp22	Hsp90aa1
18	Acoxl	Ybx3	Ffar4	Bcam	Ranbp1
19	Sftpa1	Rps17	Sox4	Fam174b	Hsp90b1
20	Napsa	Erh	Arl4c	Igfbp7	Tuba1b
21	Lrp2	Rps12	Fn1	Clic3	Txnrd1

22	Tmem243	Rpl5	Anxa1	Myl12a	Prdx1
23	Npc2	Rps6	Tspan1	Timp3	Spc24
24	S100g	Krt19	Btg1	Sec14l3	Siva1
25	Lgi3	Ptma	Myh9	Sep15	Rps2
26	Soat1	Eif5a	Lurap1l	Qk	Tpm4
27	Dram1	Rpl17	Myl12b	Dag1	Rps25
28	Fabp5	Snrpd1	Pmepa1	Cd81	Rps15
29	Sftpb	Homer2	Gadd45b	Msn	Hnrnpab
30	Lpcat1	Rbm3	Krt18	Pdgfa	Acot7
31	Nrp1	Ddx39	Mfge8	Icam1	Cbr3
32	Acly	Ppp1r14b	Neat1	Vegfa	Ppp1r14b
33	Rps28	Rpl13	Malat1	Myh14	Hmgb1
34	Etv5	Rps13	Epcam	Gsn	Rpl23a
35	Rps29	Gnb2l1	Basp1	Sema3e	Rps18
36	Scd2	Rps25	Fxyd3	Myo1b	Rps14
37	Tmprss4	Rps10	Por	Ppp3ca	Eif4a1
38	Slco4c1	Rps15	AW112010	Malat1	Manf
39	Bex2	Set	Junb	Aqp5	Rps17
40	Atp6v1g1	Wfdc2	Krt19	Fam189a2	Nme1
41	Muc1	H2afz	Laptm4a	Ppic	Gsr
42	Car8	Rps16	Slc4a4	Gprc5a	Rps8
43	Cpm	Rps11	Itgb6	Pdlim2	Rpl17
44	Mid1ip1	Snrpe	Vill	Slco3a1	Sec61b
45	Gm10076	Emp1	Prdx2	Gramd2	Erh
46	Tgoln1	Ranbp1	Runx1	Lrpap1	Cdc20
47	Ppp1r14c	Areg	Cdkn1a	Samhd1	Ckap4
48	Lbp	Lsm6	Vmp1	Vamp8	Snrpd1
49	Epas1	S100a14	Cyba	Slc44a2	Sdf2l1
50	Trf	Rpl22	Pnrc1	Myl12b	Rpsa
51	Mien1	Foxq1	Serpinb9	Lmo7	Tma7
52	Rpl21	Rpl32	Rhoc	Ehd2	Arpc2
53	Mtch1	Ppa1	Cd81	Scnn1a	Cks2
54	Rps27	Rps8	Srsf5	Serpinh1	Ube2s

55	Acsl4	Psca	Dusp1	Fads3	Pdia6
56	Mt1	Ffar4	Mmp14	mt-Cytb	Mylk
57	Lrrk2	Krt18	Thbs1	Sptbn1	Hint1
58	Acsl5	Rps20	Calm1	Rtkn2	Lsm6
59	Ifi30	Hsp90ab1	Foxq1	Cnn2	Tagln2
60	H2afj	Nap1l1	H3f3b	Myl6	Cdk1
61	Ang	Nme1	Ifitm3	Tmem37	2810417H13Rik
62	Rpl38	Ywhaz	Hp	Anxa3	Hsp90ab1
63	Col6a1	Rpl26	Tnfaip8	Agrn	Smc2
64	Tmem30a	Ncl	Ube2b	Asah1	Kdelr2
65	Atox1	Msln	Psca	Hck	Pfn1
66	Hdc	Rps15a	Serpinb6b	Scnn1g	Anp32b
67	Tspan11	Rpl13a	Gsto1	Gng5	Nucks1
68	Adam19	Pfn1	Ly6c1	Dpysl2	Slc25a4
69	Tc2n	Gsto1	Cd44	Serpinb9	Rpl14
70	Atp11a	Btf3	Anxa5	Ano1	Ak2
71	H2-Aa	Hspd1	Ctnnb1	Mmp11	mt-Nd1
72	Rps21	Gpx2	Mmp23	Cd9	Ccnd1
73	Id2	Uchl3	Sdc1	Laptm4a	Mki67
74	Secisbp2l	Mal	Rras	Magi3	Ppa1
75	Cat	Fam107b	S100a6	Tspan15	Cd36
76	Rbms3	Hnrnpab	Sorcs2	Cav2	Fen1
77	Egfl6	Rpl9	Luzp1	App	Set
78	Abcd3	Esd	Cxcl17	Galnt18	Gclc
79	Tmem164	Eif4a1	Msn	Cdkn2b	Cldn3
80	Ank3	Mgst3	Hsp90ab1	Anxa5	Ptgr1
81	Uba52	Utp11l	Ptprs	Npnt	Hnrnpa2b1
82	Itpr2	Cct8	Spint2	Tmem59	Ywhaz
83	Aldh2	Rpl23	Pdgfa	Rhoa	Atp2a2
84	Cebpa	Txn1	Glul	Serf2	Smc1a
85	Azin1	Krt8	Arpc3	Limch1	Rps16
86	Cd74	Hmgb1	Il18r1	Crlf1	Banf1
87	Mt2	Eef1b2	Gnas	Scnn1b	Rplp2

88	Fgfr2	Rpl23a	Ahr	Bsg	Ndufa4
89	Cd36	Odc1	Aqp4	Dapk2	Gpx2
90	Pi4k2b	Rpl7a	Ier5	Cav1	Prc1
91	Dpp4	Srsf3	Ano1	Ypel3	Dut
92	Nucb2	Serbp1	Map1lc3b	Mthfd1	Ptges3
93	Spg21	Anxa1	Msln	Tmbim6	Tyms
94	Scamp1	Eif1ax	mt-Atp6	Nedd9	Birc5
95	Fdps	Nhp2	Vamp5	Zfyve21	Tubb4b
96	Lcp1	Cct2	Ppic	Ndst1	Knstrn
97	Cd302	Rpl4	Mtus1	Cadm1	Dnmt1
98	Rbpjl	Snrpf	Nupr1	Timp2	Rangap1
99	Atp6v0d1	Phgdh	mt-Co3	Tead1	Rpl32
100	Zdhhc3	Anxa8	Rtn4	Epb41l5	Lsm2

Table 18: Top 100 upregulated genes in KPY scRNA-Seq time course clusters
$C_{KPY}4$ represents the root cluster, $C_{KPY}0$ and $C_{KPY}8$ transitioning clusters, and $C_{KPY}2$ and $C_{KPY}5$ late-stage clusters. Genes are ranked by p values with lower values ranked higher.

rank	$C_{KPY}0$	$C_{KPY}2$	$C_{KPY}4$	$C_{KPY}5$	$C_{KPY}8$
1	Clu	Malat1	Ybx1	Scd1	Cxcl15
2	Ppia	Bsg	Ppp1r14b	Slc34a2	Sftpd
3	Spp1	Fau	Plaur	Lcn2	Sftpc
4	Chchd2	Rps27	Rplp2	Sftpc	Prdx4
5	Prdx2	Rpl38	Pfn1	Ppp1r14c	Phgr1
6	Igfbp7	Rps28	Rplp1	Lrg1	Wbp5
7	Tspo	Bnip3	Hdgf	Dram1	Ngfrap1
8	Basp1	Aldoa	Hnrnpab	Lpcat1	Dbi
9	Krt7	Vegfa	Ywhaz	Lamp3	Ctsh
10	Krt18	Ero1l	Hspd1	Fasn	Ager
11	Hn1	Rpl37	S100a6	Napsa	Napsa
12	Cdkn2a	Emp2	Snrpf	Atp11a	Slc34a2
13	Pmepa1	Rbm39	Rpl5	Cxcl15	Siva1
14	Erh	Epb41l4aos	Ran	Elovl1	Ppia
15	Lgals1	Rpl36	Set	Scp2	Trf
16	Tmsb4x	Hcfc1r1	Pcbp1	Acsl4	Lpcat1

17	Tmsb10	Rpl39	Mal	Tgoln1	Sftpa1
18	Igfbp4	Neat1	Ybx3	Lbp	Tmem213
19	Fn1	Ypel3	Eif4a1	Lyz2	Dram1
20	Epcam	P4ha2	Ddx39	Abca3	Pdia6
21	Rpsa	Eif4a2	Slc7a11	Sftpa1	Hdc
22	Prss22	Rps21	Ncl	Npc2	Lamp3
23	Park7	Rpl37a	mt-Nd1	Fgfr2	Rgcc
24	Ctse	Gpi1	Lmna	mt-Co1	Fasn
25	Arl4c	F3	Eif4g2	Tmem243	Cldn3
26	Pfdn1	Rps29	Akr1b8	Map1lc3a	Icam1
27	Hint1	Bhlhe40	Rps12	Nrp1	Atp5g1
28	Ly6a	Npepps	Actb	Tmem30a	Elovl1
29	Nedd8	Lamp2	Ptma	Trf	Scd1
30	Msln	Psap	Alyref	Mgst1	Sftpb
31	Sox4	Rpl35a	mt-Nd3	Ptprf	Ssr2
32	Edf1	Itm2b	Tuba4a	Cd74	Sec61b
33	Pdlim7	Rpl24	Nap1l1	Secisbp2l	Sparc
34	Tnfrsf12a	Bnip3l	Ranbp1	Dbi	Psmb6
35	Hmgn1	Txnip	Hsp90aa1	Gstt1	Rpl36al
36	Psmb5	Ctsd	Rpl14	Atp5e	Bex4
37	Fam213a	Nupr1	Mgst3	Lgi3	Hsp90b1
38	Rpl19	Tpm2	Rps6	Brd7	Avpi1
39	Cfl1	Ndufa6	Npm1	Cebpa	Snx7
40	Arpc1b	Cd81	Ccnd1	Soat1	Tmem37
41	Eif5a	Rras	Odc1	Cox6c	Hmgn1
42	H3f3b	Cystm1	Ppa1	Mpc1	Ppib
43	Ctsl	Ube2b	Txn1	Aox3	Atp6v1g1
44	Rpl10a	Ndrg1	Emp1	Cpm	Scd2
45	Dad1	Egln3	Serbp1	Lrrk2	Abca3
46	Rps5	mt-Co1	Srsf2	Atp6v1c2	Pmvk
47	Cdc42ep5	Rpl28	Cltb	Nkd1	Etfb
48	Ly6c1	Gapdh	Rps27l	Tfrc	H2afz
49	Cks1b	Sec62	Rps15a	Trp53inp2	Atox1

50	Chic2	Nsa2	Tubb5	Rps27	Sec11c
51	Calm1	Fndc3a	Lsm6	Lrp2	Lbp
52	Snrpd1	Cdkn1a	Gsn	Ptgs1	Ifi30
53	2700094K13Rik	Rabac1	S100a10	Sfta2	Ostc
54	Rpl3	Maff	Mbnl1	Etfb	Id2
55	Cdk4	Junb	Sfpq	Sepp1	Sel1l
56	Prdx1	Rpl18a	Anxa2	Atp1b1	Fbp2
57	Ier3	Rpl35	Ier5	PISD	Atp5o
58	Swi5	Mif	Ptgs2	Muc1	Krtcap2
59	H2afz	Ndnf	Rps14	Sftpb	Chchd2
60	Cyba	Tmem59	Rpl13a	H2-D1	Oaz1
61	Clic1	Gng5	Itga6	Cd36	Lsm4
62	Nhp2	Pfkp	Procr	Rgcc	Lrg1
63	Sumo2	Gabarap	Ddx21	Chil1	Tuba1b
64	Abracl	Rpl22	Ptges3	Rnase4	Fkbp2
65	Fkbp1a	Ankrd37	Gls	Tcn2	Ctsc
66	Krt8	Mxi1	Rps17	Abcd3	Mydgf
67	Pomp	Serpine1	Ndufa4	Me1	Ptma
68	Anxa5	Eif3h	Rab21	Ifi27	1110008F13Rik
69	Taf10	Rbpms	Pgd	Atp6v1g1	Psmb3
70	Rnaseh2c	Clic5	Gpx2	Arg2	Atp5f1
71	Nme1	Btg1	Hnrnpa2b1	Tc2n	Nme1
72	Cox5a	Ptges	Psat1	Rab27a	Cd74
73	Psmb6	Aldoc	Epha2	Scd2	Ppp1r14c
74	Eif6	Morf4l1	Nasp	Dpp4	Cmtm8
75	Stmn1	P4ha1	Nme1	Slco4c1	Fkbp4
76	Rps3	Mfge8	Eef1g	Zdhhc3	Mgst1
77	Cnn3	Por	Cers2	Atp1a1	Tmem30a
78	Snrpd3	Rpl34	Prkar2a	Ces1d	Gm42418
79	Gsto1	Ctsb	S100a11	Dcxr	Ndufb9
80	Rbm3	Zfos1	Hsp90ab1	Epas1	Atp5b
81	Gnas	1810037I17Rik	Lmo7	Ndufa3	Tgoln1
82	Ndufab1	Fam162a	Hnrnph1	Lgals3bp	Ndufab1

83	Phgdh	Paip2	Cycs	Prnp	Pqlc1
84	Anxa3	Ly6e	Anp32b	Pi4k2b	Isyna1
85	Bok	Pdgfa	Eif5	Sftpd	Fabp5
86	Ran	Uba52	Rps15	Cat	Psmb5
87	Ranbp1	Aplp2	Bzw1	S100g	Prnp
88	Tmem176b	Lgals3	Eif4g1	Car8	Mpc1
89	Ssx2ip	Rps27rt	Nolc1	Psmb8	Spcs2
90	Sumo1	Rpl23	Car2	Tmem213	Rpl29
91	Dynll1	Rnf186	Slc35d1	Lamp1	Tubb2a
92	Rhoc	Mbnl2	Rps2	Tmbim6	Lcp1
93	Rps2	Pgk1	Tspan8	Pmvk	Npc2
94	Gzme	Phlda1	Cmpk1	Itga9	Lcn2
95	S100a11	Rpl36a	Rpl12	Chchd10	Atp5g3
96	Psmd8	Igf1r	Il33	Acly	Hspa5
97	Gtf3a	Pttg1ip	Cap1	Mtch1	Ssr4
98	Srsf3	Tspan2	Rrad	Rps29	Psmb4
99	Ckb	Hilpda	Mboat1	Ndufa1	Etv5
100	Pfn1	Hopx	Tomm20	Etv5	Cdc20

Table 19: Top 50 upregulated genes in mesenchymal scRNA-Seq time course clusters
Genes are ranked by p values with lower values ranked higher.

rank	C_M0	C_M1	C_M2	C_M3
1	mt-Co3	Fth1	Thbs1	Mgp
2	Mme	Cst3	Tmsb4x	Spp1
3	Ptn	C1ra	Tpm1	Cfh
4	mt-Nd1	Uba52	Tagln	Ctsl
5	Tpt1	Cp	Marcks	Mt2
6	Sned1	Gm10076	Fbln2	Rpl13a
7	mt-Nd4	C1s1	Lox	Mt1
8	Rps14	Nrp1	Ddah1	Timp1
9	mt-Atp6	C3	Fbln5	Pmepa1
10	H3f3a	Cd302	Col12a1	Rpl12
11	Malat1	Sparc	Mfap5	Rpl17

12	Zfp36l1	Igfbp7	Anxa3	Rbp1
13	Ptma	Rpl35	Myl12a	Mpp6
14	Igfbp4	Igfbp4	Sfrp1	Igfbp7
15	mt-Nd3	Cfh	S100a6	Tmem176a
16	Aplp2	Ptgfr	Acta2	Fam20c
17	Efemp1	Gas6	Tpm2	Gnb2l1
18	Hsp90ab1	Tmem176b	Crim1	Pi15
19	Id1	Lum	Cald1	Rps6
20	Gstm1	Tmem176a	S100a11	Gpi1
21	Tubb5	Ltbp2	Myl9	Tmem176b
22	Prdx1	Mmp2	Ogn	Colec12
23	Plac8	Itm2b	Timp1	Arpc1b
24	Fabp4	Cxcl5	Cxcl12	Gm12840
25	mt-Co2	Rpl22l1	mt-Cytb	Nupr1
26	Maf	Steap4	Col8a1	Por
27	Ifitm2	Ahr	Phldb2	Rpl8
28	Rpl13a	Slit2	Ptx3	Rpl10
29	Gsn	Col4a1	Timp3	Rps15
30	Tgfb2	Bgn	Serpine1	Rpl26
31	Ramp2	Igf1	Tpm4	Rps3
32	Rplp1	Eva1b	Wisp2	Gapdh
33	Eef1a1	Selm	Fam198b	Tpt1
34	Ptgfrn	Col3a1	Fn1	Dad1
35	Ftl1	Ifitm3	Myl6	Capg
36	Fos	Lbp	Ctps	Eef1g
37	Oaz1	Col1a2	Picalm	Pkm
38	mt-Nd2	Rpl15	Cryab	Rpl5
39	Enpp2	Saa3	Rspo3	Rpl4
40	Klf2	Srgn	Cemip	Rnf149
41	Fabp5	Nsg1	Fhl1	Slc36a2
42	Klf6	Nenf	Cyp1b1	Ybx3
43	mt-Co1	Fzd1	Myh10	Rpl28
44	Vcan	Serping1	Ccnd1	Fzd1

45	Rpl5	Ccdc80	Sdpr	Bgn
46	Nedd4	Hhip	Tnc	Rplp0
47	Ppp3ca	Gpr88	Hs6st2	Aldoa
48	Ddx5	Islr	Gyg	Ltbp3
49	Phlda1	Col1a1	Fst	Tm4sf1
50	S100a4	Sepw1	Actb	Pcdh9
rank	**C_M4**	**C_M5**	**C_M6**	**C_M7**
1	Ctsl	Gm42418	Col11a1	Dbi
2	Spp1	Rpl27	Matn4	Mpc2
3	Grem1	Ptn	Penk	Chchd2
4	Rpl38	Rps15a	Peg3	Ndufa4
5	Serpine2	Rps28	Dbi	Gas6
6	Npc2	Rpl23a	Col9a2	Cox8a
7	Tnfrsf9	Mme	Col9a3	Cox7c
8	Esd	H3f3a	Ptn	Cox6b1
9	Mif	Nedd4	Fbn2	Tgfb2
10	Rps29	Rps27	H19	Ndufb9
11	Psap	Gstm1	Col2a1	Uqcr11
12	Rpl37	Rpl37a	Fgfrl1	Uqcr10
13	Fau	Rps21	Serpinh1	Uqcrb
14	Rpl39	Ddx5	Vim	Nudt4
15	Rps21	Snrpg	Frzb	Ebf1
16	Ptgs2	Cyb5a	mt-Co3	Gpx1
17	Rpl23	Rplp2	Serf2	Cox5b
18	Rpl36	AY036118	Ndufa4	Tspo
19	Mgp	Dnaja1	Mest	Atp5l
20	Ctsb	S100a4	Col27a1	Atp5h
21	Rps28	Sned1	Fos	Ldhb
22	Cd9	Rpl37	Col9a1	Atp5j2
23	Rps27	Ptges3	Acan	Romo1
24	Gpi1	Rps26	Rplp1	Fabp4
25	Fndc3a	Rps23	Oaz1	Cox4i1
26	Ahr	Hnrnpa0	mt-Nd4	Atp5j

27	Rpl37a	Ptma	Ptma	Fabp5
28	Pi15	Mgst1	Nqo1	Cox7b
29	Rpl35a	Rpl34	Btf3	Uqcrh
30	B2m	mt-Nd4l	Rps12	Mpc1
31	Nrn1	Btf3	mt-Nd1	Uqcrq
32	Fn1	Rps12-ps3	Rps26	Ndufs5
33	Ero1l	Snrpe	Eef1a1	Hist1h2bc
34	P4ha1	Naca	Cnn3	Cox7a2
35	Ier3	Zfand5	Wwp2	Ndufb3
36	Bst2	Arl6ip1	Rps14	Atp5f1
37	Ctsc	Sumo2	Snai2	Atp5g1
38	Nupr1	Rps7	mt-Nd3	Hp
39	Mmp3	Npm1	Hsp90ab1	Ppia
40	Vegfa	Shfm1	Col11a2	Map1lc3a
41	Ptges	Tubb5	Pdia6	Ptgfr
42	Rpl22	Fam3c	Rps27l	Atp5g3
43	Igfbp7	Rps14	Shfm1	C3
44	Rpl30	Prdx4	Rps15a	2410015M20Rik
45	Pkm	Matn2	Ybx1	Rgs2
46	Ecm1	mt-Nd3	mt-Atp6	Atp6v0e
47	Ctsk	H2afz	Papss2	Sned1
48	Mt1	Rpl39	Lrrc17	Ndufc1
49	Gk	Clta	Tpt1	Cox6c
50	Rpl36a	Tpt1	Egr1	Ramp2

The curriculum vitae is not included for data protection reasons.

The curriculum vitae is not included for data protection reasons.

Declaration:

In accordance with § 6 (para. 2, clause g) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I represent the field to which the topic "*Development of an organoid model for lung adenocarcinoma*" is assigned in research and teaching and that I support the application of Antonella Dost.

Essen, 1.2.2021	Michael Ehrmann	
	Name of the scientific supervisor/member of the University of Duisburg-Essen	Signature of the scientific supervisor/member of the University of Duisburg-Essen

Declaration:

In accordance with § 7 (para. 2, clause d and f) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I have written the herewith submitted dissertation independently using only the materials listed, and have cited all sources taken over verbatim or in content as such.

Boston, 26.1.2021	Antonella Dost	
	Name of the doctoral candidate	Signature of the doctoral candidate

Declaration:

In accordance with § 7 (para. 2, clause e and g) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I have undertaken no previous attempts to attain a doctoral degree, that the current work has not been rejected by any other faculty, and that I am submitting the dissertation only in this procedure.

Boston, 26.1.2021	Antonella Dost	
	Name of the doctoral candidate	Signature of the doctoral candidate