

Review



**Cite this article:** Hynds RE, Frese KK, Pearce DR, Grönroos E, Dive C, Swanton C. 2021 Progress towards non-small-cell lung cancer models that represent clinical evolutionary trajectories. *Open Biol.* **11**: 200247. <https://doi.org/10.1098/rsob.20.0247>

Received: 12 August 2020  
Accepted: 10 December 2020

**Subject Area:**  
biotechnology/genomics/cellular biology/  
molecular biology

**Keywords:**  
cell lines, organoids, patient-derived  
xenografts, genetically engineered mouse  
models, cancer evolution, model systems

**Author for correspondence:**  
Robert E. Hynds  
e-mail: [rob.hynds@ucl.ac.uk](mailto:rob.hynds@ucl.ac.uk)

# Progress towards non-small-cell lung cancer models that represent clinical evolutionary trajectories

Robert E. Hynds<sup>1,2</sup>, Kristopher K. Frese<sup>3,4</sup>, David R. Pearce<sup>1</sup>, Eva Grönroos<sup>2</sup>, Caroline Dive<sup>3,4</sup> and Charles Swanton<sup>1,2</sup>

<sup>1</sup>Cancer Research UK Lung Cancer Centre of Excellence, UCL Cancer Institute, University College London, London, UK

<sup>2</sup>Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK

<sup>3</sup>Cancer Research UK Lung Cancer Centre of Excellence, University of Manchester, Manchester, UK

<sup>4</sup>Cancer Research UK Manchester Institute Cancer Biomarker Centre, University of Manchester, Alderley Park, Macclesfield, UK

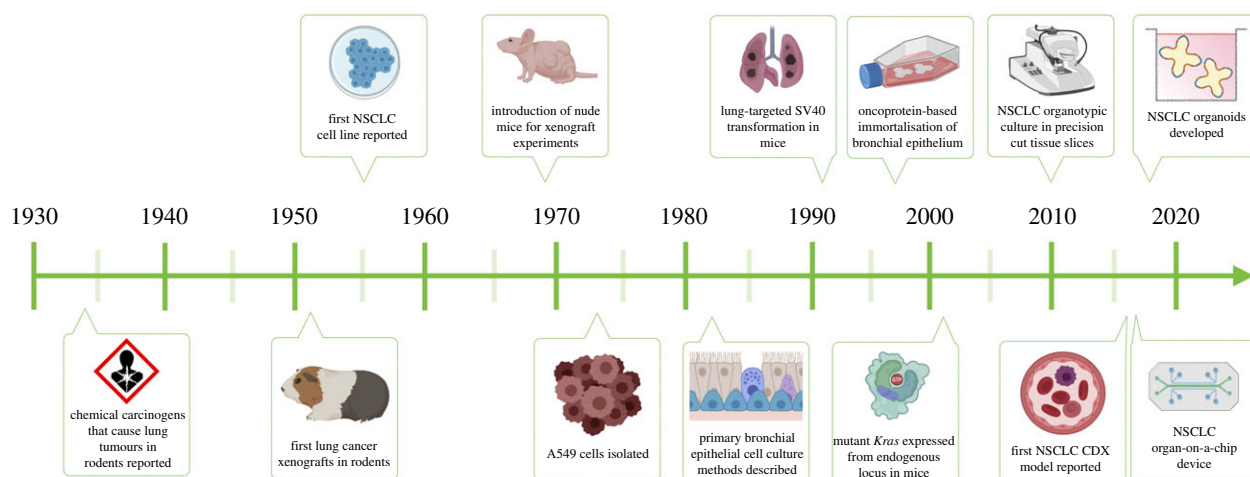
**id** REH, 0000-0002-2170-8791; KKF, 0000-0002-1369-0907; DRP, 0000-0001-8443-2924; EG, 0000-0001-8303-5409; CD, 0000-0002-1726-8850; CS, 0000-0002-4299-3018

Non-small-cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide. Although advances are being made towards earlier detection and the development of impactful targeted therapies and immunotherapies, the 5-year survival of patients with advanced disease is still below 20%. Effective cancer research relies on pre-clinical model systems that accurately reflect the evolutionary course of disease progression and mimic patient responses to therapy. Here, we review pre-clinical models, including genetically engineered mouse models and patient-derived materials, such as cell lines, primary cell cultures, explant cultures and xenografts, that are currently being used to interrogate NSCLC evolution from pre-invasive disease through locally invasive cancer to the metastatic colonization of distant organ sites.

## 1. Background

Lung cancer is the most commonly diagnosed cancer type globally for men and women, and constitutes almost one in five cancer deaths worldwide [1]. Lung cancers are classified as either small-cell lung cancer (SCLC; approx. 15%) or non-small-cell lung cancer (NSCLC; approx. 85%). NSCLC is, in turn, divisible into two main histological subtypes: lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), plus several less frequently observed subtypes, such as large cell carcinoma, adenosquamous carcinoma and carcinoid tumours. LUAD typically arises in the distal lung, whereas LUSC arises centrally, probably reflecting different cells-of-origin for these two lung cancer types [2]. It is widely thought that LUAD develops from alveolar type II (AT2) epithelial cells or cells within bronchioalveolar duct junctions, whereas LUSC develops from basal epithelial cells in airways, although data from animal models [3] and an increasing appreciation of the plasticity of lung epithelial cells [4] make this uncertain.

The two major NSCLC subtypes can be distinguished further based on cell morphology and histological staining: LUAD typically appears glandular, whereas LUSC harbours large polygonal cells with squamous differentiation. Different marker proteins also aid diagnosis as TTF-1/NKX2-1 and KRT7 expression are indicative of LUAD, whereas TP63 and KRT5/6 expression are indicative of LUSC [5]. At the genomic level, the mutational and copy number landscapes of LUAD and LUSC are distinct [6–8]. In LUAD, the occurrence of oncogenic driver mutations in *KRAS*, *EGFR*, *HER2*, *MET* and *FGFR1/2*, as well as oncogene fusions involving anaplastic lymphoma kinase (*ALK*), the *ROS1* receptor tyrosine kinase (RTK), neuregulin 1 (*NRG1*), neurotrophic tyrosine kinase receptor type 1 (*NTRK1*) and *RET*, offer possibilities for new targeted



**Figure 1.** A timeline of advances in pre-clinical models of non-small-cell lung cancer. Created with BioRender.com based on [16–28].

therapies. Recent progress has been made using small-molecule inhibitors to target difficult-to-drug mutated forms of KRAS, and clinical trials are on-going [9]. While LUSC is not characterized by the same mutations as LUAD and has fewer targetable oncogenic drivers, tumour suppressor alterations such as *TP53*, *CDKN2A* and *KEAP1* are common in both subtypes [6–8].

Lung cancer survival is highly stage dependent; in England between 2013 and 2017, diagnosis at stage I was associated with a greater than 50% 5-year survival, whereas if the diagnosis was at stage IV, i.e. metastatic disease, the equivalent figure was approximately 3% (Office of National Statistics, UK). In NSCLC, independent of histological subtype, the standard first-line treatment for patients with stage I–III tumours is surgical resection, with adjuvant chemotherapy offering a small benefit for those with locally advanced stage III disease [10]. If surgery is not possible or is declined then chemoradiotherapy is typically offered. Immune checkpoint inhibitors have revolutionized NSCLC treatment and emerging survival data from early phase clinical trials indicate a significant increase in median overall survival for a subset of patients. Anti PD-L1 and PD-1 therapies have been licenced for use in both locally advanced and advanced cases, respectively [11] and although treatment efficacy has been linked to tumour PD-L1 expression [12], patient stratification for immunotherapy agents requires further refinement [13]. Targeted therapy has predominantly focussed on inhibiting the constitutive activation of mutated forms of the epidermal growth factor receptor (EGFR). A majority of patients initially respond to treatment but eventually progress as therapy resistance develops [14]. The emergence of resistance coupled with a high number of unknown resistance mechanisms indicates the potential for rapid tumour evolution [15].

The recent advancements in cancer treatments outlined above would not have been achieved without experimental models to investigate the different aspects of disease initiation and progression. Pre-clinical models represent important tools that allow us to study tumour evolution in the absence of therapy in a manner that is not possible in patients. Along with enabling studies of early disease, these models also allow us to compare the efficacy of novel therapies with established treatments and to study mechanisms of therapy resistance. Such systems have the potential to identify biomarkers of response for patient stratification and to inform future personalized therapies. In this review, we describe the progress that has been made to diversify the tools available for NSCLC research, discuss their relative advantages and disadvantages for

particular research questions and reflect on some of the outstanding questions facing the field.

## 2. Pre-clinical NSCLC model systems

The study of NSCLC has progressed tremendously since the initial investigations identifying chemical carcinogens as a source of lung cancer (figure 1). The technical and scientific advancements in NSCLC research have included the establishment of immortal cell lines, primary cell cultures, xenografts and mouse models, which each have their relative merits and disadvantages (table 1).

### 2.1. Established cell lines

The most frequently reported NSCLC laboratory models are cell lines (table 2), which are inexpensive, scalable and widely available [29,30]. Cells, isolated from patients' tumours have been selected for growth, most often on plastic cultureware, in culture medium containing bovine serum. Lung cancer cells were among the first to be successfully cultivated in the laboratory [32] and, to date, over 200 NSCLC cell lines are available in cell line collections [33,34]. Sequencing efforts from the Wellcome Trust Sanger Institute's cell lines project (<https://cancer.sanger.ac.uk/cosmic>) and the Broad Institute's Cancer Cell Line Encyclopedia (<https://portals.broadinstitute.org/ccl>) have identified the mutational status for 110 and 136 NSCLC cell lines, respectively. Community resources such as Cellosaurus (<https://web.expasy.org/cellosaurus>) also collate publicly available cell line information. Most NSCLC cell lines are derived from adenocarcinomas with fewer LUSC cell lines available due to the lack of effective culturing methods. In general, LUSC cell lines tend to be less well-characterized and there are concerns that some cell lines are potentially misattributed due to their *in vitro* similarity to squamous cancers from other organs. NSCLC cell lines maintain some of the fundamental features of the tumours from which they were derived [35] but the most widely used NSCLC cell lines are now several decades post-establishment, limiting the availability of clinical data and modern genetic characterization of the parental tumour, including germline sequencing. It is important to recognize that, due to on-going mutational processes and genomic instability, the divergence of these long-term cultures from the original tumour occur during continued propagation. Additional

**Table 1.** An overview of the relative merits of NSCLC models.

| source material                             | advantages  | disadvantages   |
|---|---|---|
| cell lines                                  | large number available<br>inexpensive and widely available<br>many are well-characterized<br>permissive of genetic manipulation<br>permissive of biochemical studies and drug screens | cell selection based on growth on culture conditions<br>phenotypic changes caused by immortalization<br>divergence/genetic drift may affect reproducibility<br>few modern LUSC lines<br>germline data generally unavailable |
| primary patient samples                     | close-to-patient: recent clinical data and other matched  | difficult to maintain AT2 phenotype in culture  |
| normal tissue                               | patient samples   | difficult to establish pre-invasive or LUSC cultures  |
| tumour biopsy                               | permissive to the on-going development of novel models,   | contaminating normal airway cells common in tumour  |
| circulating tumour cells                    | e.g. those involving organoids and/or co-culture  | cultures  |
|   |   | establishment within a clinical timeframe is challenging  |
|   |   | limited availability  |
| genetically engineered mouse models (GEMMs) | reproducible within pure background strains<br><i>in situ</i> microenvironment<br>control over genetic alterations<br>immune competent animals  | strong oncogenic drive limits study of early tumour evolution<br>more rapid progression than in human patients<br>multiple allele generation can require many generations and be costly<br>species differences              |
| carcinogen-induced tumour models            | mimic pre-invasive disease stages<br><i>in situ</i> microenvironment<br>immune competent animals  | more rapid progression than in human patients<br>often develop the extensive multi-tumour disease<br>species differences  |

complexity and irreproducibility is introduced by the different selection pressures applied as multiple laboratories cultivate cell lines with variable tissue culture practices. Consequently, divergent cell growth behaviour [36] and response to therapies [37] have been reported.

## 2.2. Patient-derived tissue

### 2.2.1. *Ex vivo* explant cultures

Small fragments or slices of resected NSCLC tumours can be maintained in the cell culture medium, allowing short-term investigations. Explant cultures were pioneered as ‘histocultures’ in which tumour pieces were supported on collagen-based sponge gels that enabled short-term tumour cell survival and proliferation read-outs [38]. In a recent study, around 70% of NSCLCs were amenable to explant culture [39] and further optimization might be possible by customizing matrix protein composition and/or using the autologous serum in such cultures [40]. Precision-cut lung slices (PCLSs) can be generated by filling tissue with agarose before creating 100–500 µm sections using a tissue slicer or a vibratome. Sections are then cultured in a maintenance medium and, although survival times vary, some cell types can survive for more than 7 days. PCLSs have gained popularity in other chronic lung disease settings [41], although the difficulties of genetic manipulation within such a short timeframe make transformation studies challenging. The surgical resection of primary tumours lends itself to such studies [16] but, in practice, application of this technology in lung cancer studies remains limited.

### 2.2.2. Primary cell cultures

Primary cell cultures, in either 2D or 3D culture formats, can be derived from cells retrieved by lobectomies, brush or forcep biopsies, from resected or biopsied tumours, or from circulating tumour cells (CTCs) recovered from blood samples. These are distinguished from cell lines by the relative recency of their establishment and, in some cases, a lack of indefinite proliferative capacity, although the long-term culture of primary cells is now possible in some systems without immortalization. Despite the ease of establishing normal airway epithelial cell cultures, pre-invasive [42] and invasive LUSC cells [43–46] are difficult to culture by existing methods, perhaps due to the presence of widespread CNAs [6,47]. Increasing the culture success rate is hampered by a lack of clarity about what tumour (or cellular) features are selected for by *in vitro* culture conditions. Recently, 3D patient-derived organoids (PDOs) have been established [48,49]. Generating pure NSCLC organoid cultures without the presence of contaminating normal airway organoids is a challenge but can be partially resolved by selection for *TP53*-mutant cells using the small-molecule Nutlin-3a [50].

### 2.2.3. Patient-derived xenograft (PDX) models

In patient-derived xenograft (PDX) models, patient tumour tissue or CTCs are implanted or injected into an immunocompromised mouse host to achieve continued proliferation of tumour cells. This is most frequently done at a subcutaneous site but can also be orthotopic or injected into the circulation. PDX models are close-to-patient and have a 30–40% success rate [51]. Higher stage tumours engraft more readily in mice and

**Table 2.** A list of selected, commonly used NSCLC cell lines along with the driver mutations found in each. Oncogene driver information, TP53 status, sex and ethnicity was derived from COSMIC ([https://cancer.sanger.ac.uk/cell\\_lines](https://cancer.sanger.ac.uk/cell_lines)) and Cellosaurus (<https://web.expasy.org/cellosaurus>). LUDLU-1 is described as per a published report [31]. WT = wild type.

| name      | cancer type                         | tissue origin                      | driver mutation  | TP53 status                                       | sex | predominant ethnicity |
|-----------|-------------------------------------|------------------------------------|--|---|-----|-----------------------|
| A549      | LUAD/carcinoma                      | primary                            | <i>KRAS</i> p.Gly12Ser (Hom)<br><i>STK11</i> p.Gln37Ter (Hom)      | WT  | M   | Caucasian             |
| NCI-H322  | LUAD                                | primary                            | unknown  | p.Arg248Leu (Hom)                                 | M   | Caucasian             |
| NCI-H358  | LUAD                                | primary                            | <i>KRAS</i> p.Gly12Cys (Het)                                       | Loss (Hom)  | M   | Caucasian             |
| NCI-H522  | LUAD                                | primary                            | unknown  | p.Pro191fs*56 (Het)<br>Sanger/Hom<br>Cellosaurus) | M   | Caucasian             |
| NCI-H3255 | LUAD                                | primary                            | <i>EGFR</i> p.Leu858Arg (Hom)                                      | c.560-1G>A (Hom)                                  | F   | Caucasian             |
| HCC-4006  | LUAD                                | metastasis:<br>pleural<br>effusion | <i>EGFR</i> p.Leu747-Glu749del.                                    | WT  | M   | Caucasian             |
| PC9       | LUAD                                | metastasis:<br>lymph node          | <i>EGFR</i> amplified, <i>EGFR</i><br>ex19del                      | p.Arg248Gln (Hom)                                 | M   | unknown               |
| LUDLU-1   | LUSC                                | primary                            | <i>BRCA1</i> , <i>BRCA2</i>  | p.Arg248Trp (Hom)                                 | M   | Caucasian             |
| NCI-H520  | LUSC                                | primary                            | <i>ATM</i> p.Pro383Ala (Het)<br><i>CDKN2A</i> p.Gly45fs*8<br>(Hom) | p.Trp146Ter (Hom)                                 | M   | Caucasian             |
| NCI-H2170 | LUSC                                | primary                            | unknown  | p.Arg158Gly (Hom)                                 | M   | Caucasian             |
| SK-MES-1  | LUSC                                | metastasis:<br>pleural<br>effusion | unknown  | p.Glu298Ter (Hom)                                 | M   | Caucasian             |
| NCI-H647  | adenosquamous                       | primary                            | <i>KRAS</i> p.Gly13Asp (Hom)                                       | c.782+1G>T (Hom)                                  | M   | Caucasian             |
| NCI-H1299 | lung large cell<br>carcinoma        | metastasis:<br>lymph node          | <i>NRAS</i> p.Gln61Lys (Het)                                       | Loss (Hom)  | M   | Caucasian             |
| ChaGo-K1  | bronchogenic<br>carcinoma           | metastatic site:<br>subcutaneous   | <i>ARID1A</i> p.His684Asp (Het)<br><i>RB1</i> p.Glu837Lys (Hom)    | p.Cys275Phe (Het)                                 | M   | Caucasian             |
| NL20      | human bronchial<br>epithelial cells | normal bronchus                    | transformed; <i>SV40</i> , LargeT                                  | WT  | F   | unknown               |

successful engraftment is a negative prognostic indicator in early stage disease [52,53]. While most PDX models are established subcutaneously for convenient tumour burden monitoring, there is evidence that orthotopic [54] or renal capsule [55] transplantation might substantially increase engraftment. Genome editing technologies are also increasingly applicable to xenografts [56]. Nevertheless, the use of a mouse host is a limitation in terms of the stromal, vascular and immune microenvironments experienced by the tumour and, comparable to cell cultures, these models are susceptible to some degree of genomic divergence due to on-going evolution [37].

## 2.3. Model organism research

### 2.3.1. Genetically engineered mouse models

Most NSCLC model organism work has focussed on the mouse as a result of the powerful genetic tools available for tumour induction and lineage tracing. Genetically engineered mouse models (GEMMs) are typically inbred mouse strains

that have been genetically manipulated to express oncogenic alleles or delete tumour suppressor genes to generate autochthonous tumours.

Since smoking-induced lung cancer has a high mutational burden, it is advantageous to use reductionist GEMMs—which allow experimental control over a small number of genetic alterations—to establish which events drive cancer and which are passenger mutations. Simplified genetic models with alterations to a few typically strong oncogenes or tumour suppressor genes are useful to dissect complex pathological mechanisms and test putative therapies in controlled conditions. The conditional mutagenesis systems CRE-LOX and FLP-FRT allow temporal control of genetic events within specific lung cell populations [57]. Still, complex models containing five or more mutated alleles are costly and still lack pre-cancer evolutionary context. While GEMMs have traditionally taken a long time to derive, the emergence of CRISPR–Cas9 genome editing increasingly allows faster derivation [58]. Overall, GEMMs typically provide rapidly developing lung cancer models that generate multiple small

tumours in the lungs but, given that they lack the genomic damage caused by tobacco smoking, they often do not capture the mutational diversity of human tumours [59,60]. Incorporating additional mutagenesis, caused either by chemicals or the expression of proteins known to cause mutations, such as cytosine deaminases, might bring mouse models closer to human tumour mutational burden. Importantly, the short (approx. 2 year) lifespan of mice and differences in cell-intrinsic factors, such as in telomere biology [61], mean that GEMM tumours are not subject to the same evolutionary histories as human tumours. Nonetheless, the use of GEMMs has significantly increased our understanding of NSCLC, both with regards to disease progression and potential treatments.

Most NSCLC GEMMs result in adenocarcinoma and there has been a particular focus on those caused by *Kras* mutations. Multiple mutant alleles exist but the most widely used is *Kras*-G12D. Adenocarcinomas can be produced by *Kras*-G12D expression in AT2 cells expressing either *Sftpc* or *Scgb1a1* (*CCSP*; *CC10*) [62]. Although *Scgb1a1* is expressed in bronchiolar club cells in this model, they do not form invasive cancers [62] unless *Trp53* mutations are introduced [63,64]. Lineage tracing in a *Kras*-G12V model further showed that many alveolar cells expressing the mutant allele do not divide [65], suggesting heterogeneity among AT2 cells. Indeed, during homeostasis, only a subpopulation of AT2 cells are Wnt active stem cells within a fibroblastic niche [66,67] and Wnt signalling has been implicated in the progression [68] of *Kras*/*Trp53* adenocarcinoma models [69]. Multiple other *Kras* combinations have been generated, including those with activating PIK3CA mutations [70] or *MYC* overexpression [71]. Similarly, mouse models of targetable mutations have been developed by expressing human EGFR [72], EML4-ALK fusion kinase [73] or ROS1 fusion kinases [74,75].

LUSC GEMM development has been hindered by the rarity of activating oncogenes, the lack of lung basal cell-specific Cre-drivers and the absence of basal cells from airway epithelium distal to the main bronchi in mouse (compared to humans, where they are present throughout many airway generations). Combining *Sox2* overexpression with tumour suppressive *Pten* and *Cdkn2a* mutations leads to LUSC-like tumours regardless of whether the Cre-driver gene is expressed by basal, club or AT2 cells [76]. *Sox2* overexpression in mouse club cells leads to the proximalization of the bronchiolar epithelium and adenocarcinomas expressing the squamous marker TP63 eventually form [77]. Hybrid approaches using mouse genetics and *in vitro* organoids have also allowed the development of tumorigenic mouse organoids that overexpress *Sox2* and harbour deletions in the key LUSC tumour suppressor genes *Trp53*, *Cdkn2a* and *Pten* [78]. *Lkb1* loss, a rare event in human LUSC, together with *Sox2* overexpression lead to mouse LUSC [79]. In LUAD, the addition of *Lkb1* mutations to *Kras*-driven GEMMs has confirmed that they modify histology, as dual mutants give rise to LUAD, LUSC and mixed adenosquamous lesions [80]. In established *Kras*-driven tumours, *Lkb1* loss promotes the transition to squamous histology [81] with redox balance [82] and epigenetic mechanisms involving polycomb repressive complex 2 (PRC2) [81] both implicated mechanistically in LUSC formation.

### 2.3.2. Carcinogen-induced NSCLC

Exposure of rodents to particulate matter, whole tobacco smoke or e-cigarette vapour produces physiologically

relevant changes in the lung epithelium [83,84] and accelerates tumour development [85–87]. Cigarette smoke condensate or pure chemical carcinogens induce lung cancers that arise through relevant pre-invasive disease processes. The most commonly used, urethane, generates predominantly *Kras*-driven LUAD which transitions through adenoma precursor lesions [88]. Multiple carcinogen-driven LUSC models are also available: benzo[a]pyrene causes squamous bronchial lesions in hamsters [89]; repeated intratracheal injection of 3-methylcholanthrene (MCA) in mice causes metaplastic lesions throughout the bronchial tree which progress to invasive and metastatic LUSC [90]; and N-nitroso-tris-chloroethylurea (NTCU) produces a LUSC-like lung cancer when applied to the back skin of mice [91–93]. Susceptibility to chemical carcinogens varies by mouse strain and correlates with the prevalence of spontaneous tumour formation.

## 3. Cancer origins in ‘normal’ tissue

The use of clonal cell culture to expand single progenitor cells has allowed whole-genome sequencing of histologically normal airway epithelium [42], revealing the presence of somatic mutations, including known cancer driver mutations. As expected, the mutational burden increased with age and was significantly higher in adults with a tobacco smoking history [42]. The detected cancer driver mutations were those typical of LUSC, with *TP53*, *NOTCH1* and *FAT1* mutations being the most frequent. Individual cells rarely contained multiple driver mutations but, unlike in LUSC tumours, copy number alterations were uncommon [42]. It remains unclear whether primary ‘normal’ airway epithelial cell culture imposes selective pressures similar to those observed in established cell lines [36,37]. Despite these concerns, primary airway cultures might facilitate studies on differences in cancer susceptibility among the four recently identified subpopulations of cultured normal airway basal epithelial cells [94].

Airway basal cells have also been immortalized, for example by using overexpression of CDK4 and TERT [95], and region-specific differences in airway biology are retained in cell lines from different locations in the proximal–distal axis of the bronchial tree [96]. To more closely resemble fully differentiated airway epithelium, it is possible to direct basal cell differentiation towards multiciliated and mucosecretory cell fate in either air-liquid interface cultures [97] or as 3D ‘tracheospheres’ [98]. More recent advances in organoid culture enable expanding cultures of airway epithelial cells containing all three major cell lineages as 3D organoids [48].

Unfortunately, *in vitro* studies of primary alveolar epithelial cells—and therefore efforts to map mutations in the normal alveolar epithelium—have been hindered by the short time that it is possible to maintain proliferating AT2 progenitor cells in culture before they differentiate to non-proliferative AT1-like cells [99]. Immortalized alveolar epithelial cells have been developed [100] but the narrow timeframe available for transduction with immortalization factors means that, similar to primary alveolar cultures, these cells more closely resemble AT1 cells. Use of Rho-associated protein kinase (ROCK) inhibitors during AT2 cell culture derivation might enable AT2 cell-like phenotype preservation for studies of step-wise carcinogenesis [99]. Similar to 2D cultures, most reported alveolar organoid systems have not

allowed serial passaging [67,101], limiting their utility for cancer modelling. Recently, long-term culture methods for human organoids containing AT2 cells that are capable of AT1-like differentiation have been described [102,103], further expanding the repertoire of models available.

Robust and scalable primary airway and alveolar culture methodologies have the potential to improve our understanding of the effect of specific somatic mutations on cellular dynamics in human epithelia and are compatible with CRISPR–Cas9-mediated gene editing approaches that might allow temporal reconstruction of NSCLC molecular events, as has been possible in studies of colorectal carcinogenesis [104].

## 4. Modelling early tumorigenesis

Early detection of NSCLC is a priority for improving clinical outcomes, particularly in cases where intervention could occur before invasive disease occurs. Achieving this will require a better understanding of pre-invasive cancer biology but the required laboratory studies are currently challenging. GEMMs typically introduce strong oncogenic and/or tumour suppressor alterations in many cells simultaneously, thus limiting the extent to which their progression mimics early tumour evolution. Nevertheless, mouse models allow a high degree of genetic control and provide a much greater range of tumour microenvironmental cues than culture systems. Primary or immortalized cell cultures from carcinoma-in-situ or adenomatous lesions have not yet been derived but progression to malignancy can be investigated through the introduction of cancer mutations into immortalized ‘normal’ basal cell lines [105]. This approach has identified a subpopulation of basal epithelial cells with enhanced motility [106] and has been used as an organotypic model of dysplasia by introducing *TP53* inactivation and *SOX2* overexpression [107]. It is highly likely that both GEMMs and cell cultures will be invaluable tools for establishing the physiological order of events that drive early oncogenesis and the dependencies of pre-malignant cells above and beyond those of normal tissues.

Chemical carcinogen models are particularly relevant in early tumorigenesis research because the tumours undergo a histological transition similar to that seen in human patients. In the urethane-induced LUAD model, the tumours are also morphologically similar to spontaneous tumours in aged mice, with a robust immune infiltrate organized in tertiary lymphoid structures. This observation, together with the fact that the growth pattern of LUAD tumours varies between chemicals, suggests that carcinogens accelerate physiologically relevant processes [108,109]. Consistent with patient data [110], *Kras* mutations are identified as an early event, *Cdkn2a* can be epigenetically downregulated in early foci and later deleted in both adenomas and adenocarcinomas, while *Trp53* mutations are found in adenocarcinomas but not hyperplasias, supporting a role for *Trp53* in invasion [60]. However, it is noteworthy that *Kras* G12C mutations are not found in either urethane- (Q61R/Q61 L) or nitrosomethylurea-induced (NMU; G12D) [60] mouse tumours as these carcinogens do not induce the requisite base substitution.

Squamous chemical carcinogen models also mimic patient pre-cancerous lesions, albeit on an accelerated time frame. In NTCU-treated mice, tracheal dysplasia precedes proximalization of the bronchial epithelium and progression

to invasive LUSC [111]. In spite of the differences in cellular composition between mouse and human airways [112], the model recapitulates key aspects of LUSC natural history, with KRT5-expressing lesions and apparent PI3 K signalling [113]. RNA sequencing analysis suggests similarities in the immune response of LUSC patients and NTCU-treated mice [114], although full genomic characterization to compare mouse to human lesions is not yet available and differences emerge between mouse strains and sexes [115]. Despite some systemic and local side effects, NTCU-treated mice rarely develop other squamous cell carcinomas, such as of the skin or oesophagus, and are suitable for chemoprevention studies. Therefore, deeper characterization and comparison of carcinogen-induced NSCLC models using modern genomic, epigenomic and transcriptomic tools are warranted.

## 5. Drug response

Cell lines are extensively used for assaying drug efficacy as clonal genetic alterations are maintained and they are amenable to high-throughput assays. There are concerns about genetic and functional differences between different sublines of established cancer cell lines, including A549 cells [37], that might result in divergent responses in compound screening experiments [116]. The use of cell lines to predict the efficacy of treatment in relation to particular tumour features, such as histology or particular mutations, will probably require investigating a large number of cell lines in order to generate robust data. Efforts such as the Cancer Cell Line Encyclopaedia [117] that have deeply characterized NSCLC cell lines enable the integration of genetic characteristics with functional drug sensitivity assays (e.g. those that can be interrogated via DepMap; <https://depmap.org/portal>). Patient-derived models such as primary cell cultures, PDOs and PDX models are expected to display greater fidelity to the behaviour of the patient tumour as a result of the recency of establishment. Indeed tumour explants [38] and PDX models [118] can be predictive of the efficacy of chemotherapy, while PCLSs have also been used to test novel therapeutics [119]. Comparisons of pre-clinical model efficacy with phase II clinical trial results support the predictive value of cell lines and human xenografts but not mouse allografts in NSCLC [120]. Overall, these studies show that selecting a range of cell lines (or patient-derived models) to address specific questions (e.g. within mutation status or histology) can improve predictivity and argue the need for comparing multiple well-characterized cell lines with other disease models.

Pre-clinical models are not only important tools for predicting the efficacy of targeted therapies, but also for determining mechanisms of resistance, for example to chemotherapies or targeted therapies. That cell lines can develop resistance to therapies through the expansion of pre-existing subclones argues that they preserve some tumour heterogeneity [121]. A concern is that selection pressure upon cell line establishment or during expansion might not fully reflect the primary cancer, thus preventing the discovery of possible resistance mechanisms. These concerns are also relevant in patient-derived models, reiterating the importance of using a wide range of models and maximizing comparisons with patient datasets where these are available.

In addition to predicting cohort-level responsiveness to therapy, it has also been proposed that patient-derived models might predict drug efficacy for individuals. However, for precision oncology to become a reality, the scalability and

time frame must fit the clinical need. Currently, these models take several weeks or months to establish, but in one study reductions in the time required to get clinically actionable information from PDX models was achieved by implantation of tumour pieces in the subrenal capsule followed by testing of alternative chemotherapy regimens with results known within eight weeks [55]. Although not yet applied to NSCLC, patient-derived tumour cell clusters have been developed which allow short-term cell expansions from primary tumours, including immune and fibroblast populations, allowing the investigation of hundreds of therapeutic options per patient in a manner that correlated with the clinical performance [122]. Of course, a caveat of these strategies might be tumour complexity with regards to intra-tumour heterogeneity and sampling bias. This concern might be minimized by targeting truncal mutations and perhaps by using as much as possible of the tumour material remaining after clinical diagnostics to generate more representative models [123].

## 6. Tumour microenvironment

An ideal model of the tumour microenvironment would comprise patient-matched tumour, stromal and immune cell populations with their native architecture preserved. Tumour explant cultures and PCLSs attempt to achieve this by maintaining living, surgically resected tumour tissue in the laboratory. Although existing studies largely focus on tumour cells in these systems, such cultures also offer an opportunity to study tumour–stroma interactions and the effects of immunomodulatory drugs. T cell populations can be imaged within PCLSs [124] suggesting the opportunity to study cellular localization, for example, the exclusion of T cells from tumours [125], in a manner that is not possible in other human *in vitro* systems. Nevertheless, the limited timeframe of such experiments and the differential sensitivity of different cell types to culture mean that mouse models and reductionist human models recapitulating specific cellular interactions remain the mainstay of tumour microenvironment research in NSCLC.

### 6.1. Tumour–stromal interactions

Increasingly, we recognize the importance of cancer-associated fibroblasts (CAFs) as regulators of lung cancer growth, immunogenicity and metastasis. CAFs are fibroblasts that have been activated by a variety of signals in the tumour microenvironment and single-cell RNA sequencing studies suggest the heterogenous nature of lung CAFs [126].

Genetic and syngeneic mouse models feature authentic stromal–epithelial interaction after tumour induction allowing the study of the interplay between the tumour and stroma. In xenotransplantation experiments, host–tumour interactions can be limited by cell signalling incompatibilities between species. A prominent example is stromal mouse hepatocyte growth factor (HGF), which cannot fully activate the human MET receptor in epithelial cells [127], potentially limiting the growth of HGF-dependent tumours. However, human fibroblasts are readily expanded in cell culture on plastic, allowing comparisons between healthy control, matched adjacent lung and tumour-associated fibroblasts. These can be introduced into cell lines, primary cell culture or xenograft assays to determine phenotypic changes and

their effects on epithelial cells. Experiments to date support the notion that (at least a subset of) CAFs act as a supportive niche that maintains NSCLC cells in a de-differentiated state [128–130]. Fully describing CAF subsets, the epigenetic stability of the CAF state and the extent to which targeting CAFs can be therapeutically beneficial are active areas of research [131].

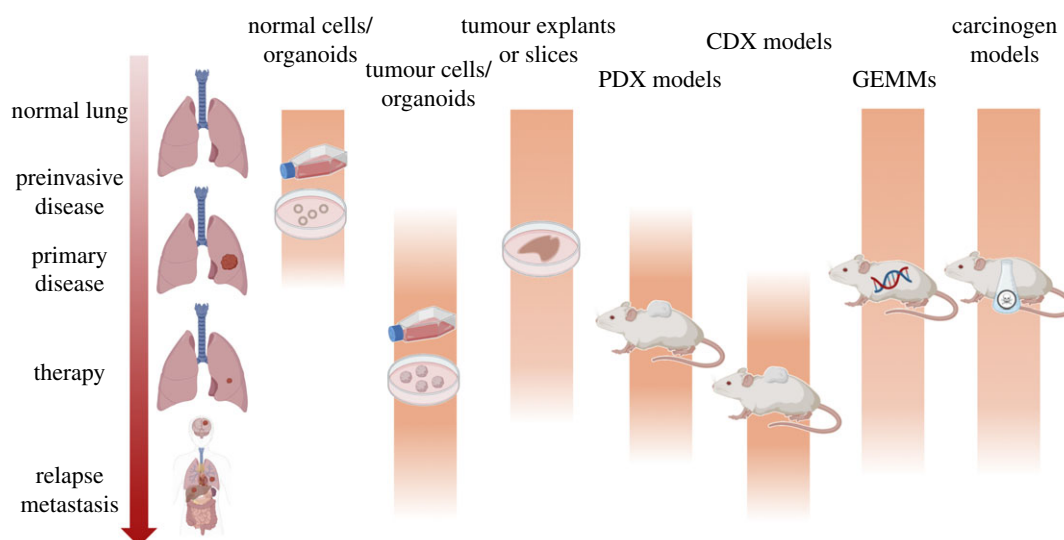
### 6.2. Mechanical force

Advancements in microfluidic technologies have enabled the development of NSCLC cancer-on-a-chip devices. Originally these devices were designed as a ‘lung-on-a-chip’, combining airway or alveolar epithelial cells with endothelial cells in distinct channels [132]. When the EGFR-mutant LUAD cell line H1975 was seeded in this system among normal epithelial cells, it grew faster among alveolar than airway epithelial cells [17], consistent with its alveolar origin. Application of mechanical force to mimic breathing led to reduced proliferation of the cancer cell line, which also became more resistant to the EGFR inhibitor rociletinib [17], demonstrating the potential for changes in cell behaviour upon incorporating physiological parameters into cell culture experiments.

### 6.3. Tumour-immune interactions

In lung cancer immunology research, intact immune surveillance is available in GEMMs as well as carcinogen-induced and syngeneic mouse models. Since immune evasion occurs early in LUSC pathogenesis [133,134], carcinogen-based models might offer an opportunity for immunotherapy development. Three main syngeneic models have been described: Madison 109 (MAD109) was derived from a spontaneous BALB/c mouse lung tumour in 1964 [135], the Lewis lung carcinoma (LLC) cell line was derived from a C57BL mouse in 1951 [136] and KLN-205 was derived from a DBA/2 mouse treated with MCA in the 1970s [90,137]. These models produce very rapidly growing tumours in the lung either following direct implantation or as a result of metastases from subcutaneous xenografts. As discussed above, GEMMs are frequently poorly immunogenic, presumably due to their low mutational burden. Nevertheless, both syngeneic models and GEMMs can be modified to present model antigens such as ovalbumin in order to study antigen presentation and recognition by transgenic T cells engineered with reactivity against ovalbumin-derived peptides [138,139]. More recently, the inversion-induced joined neoantigen (NINJA) system has been developed, which allows spatially and temporally controlled expression of defined neoantigens in mouse lung cancer cell lines, lung tissue and in GEMMs [140]. Crucially the system avoids central tolerance and as such will allow the investigation of endogenous T cells in anti-tumour immunity [140].

Nevertheless, there are substantial differences between the composition, protein production and function of the human and mouse immune systems. Of particular relevance given that neutrophils are the most abundant immune cell population in the NSCLC tumour microenvironment [141], human blood contains more neutrophils than lymphocytes, whereas the mouse lymphocytes substantially outnumber neutrophils [142]. As such, a range of model systems now exist to explore human immune cells in the context of NSCLC.



**Figure 2.** Pre-clinical models across the disease course of non-small-cell lung cancer. Created with BioRender.com.

Firstly, human immune cells can be introduced into mouse models. In adoptive cell therapy experiments, injection of tumour-bearing mice with engineered CAR-T cells has expanded studies beyond *in vitro* reactivity assays [143] and could similarly be applied to tumour-infiltrating lymphocytes (TILs) [144], which have entered early phase clinical trials for NSCLC. To investigate checkpoint inhibitors, TILs within initial biopsies can be studied by co-administering checkpoint inhibitors at the time of tumour implantation and then sacrificing the mice at early time points [145]. Although the long-term effects of therapy on tumour burden cannot be established in this model, it is nevertheless possible to study the phenotypes of both the tumour and immune cell populations.

In xenotransplantation experiments, the required use of severely immunocompromised mice with deficiencies in both innate and adaptive immune responses has drawbacks for studying pathogenesis and testing potential therapies, particularly emerging immunomodulators. More complete reconstitution of the human blood and immune cell lineages has been achieved in ‘humanized’ mice where haematopoietic stem cells (HSCs) are engrafted in NSG mice following irradiation-induced clearance of host progenitor cells [146]. Humanized mice can live for more than 30 weeks post-transplantation [147], which has allowed the subsequent application of a human tumour model. The response to anti-PD-1 therapy was dependent upon the presence of human T cells and variable depending on the donor cells used for reconstitution [148]. Unfortunately, it is rarely possible to isolate HSCs and tumour cells from the same individual and, even in donor-matched experiments, the T cells that develop in humanized mice are educated in the mouse thymus and thus selected on mouse MHC, resulting in an allogeneic response which limits studies of tumour-specific antigen responses.

In studies of newly derived cell lines or PDOs, it is possible to establish simultaneous patient-matched cultures of fibroblasts or immune cells to increase model complexity and better represent the tumour microenvironment. For example, T cell-tumour interactions have been investigated in entirely *in vitro* systems with expansion of tumour-reactive T cells from PBMCs and co-culture with autologous NSCLC organoids [18].

## 7. Metastasis and relapse

NSCLCs commonly metastasize to the brain, bone, liver and adrenal glands [149]. Compared to primary disease, tissue acquisition in metastatic, progression and relapse settings are more challenging; clinical diagnosis typically relies on imaging and biopsies, but metastases are rarely surgically resected. Some NSCLC cell lines have been derived from metastases but most lung PDX models are primary tumour-derived even though metastases have a high engraftment rate (figure 2) [150].

CTCs, which are more abundant in patients with late-stage cancers [151], might help to overcome the primary disease bias of PDX collections. Methods to derive CTC-derived explant (CDX) models have been developed in SCLC [152] where researchers faced similar sampling difficulties. CTCs are more frequent in SCLC than in NSCLC patients, and while NSCLC CDX models have been reported that mimic donor patient treatment response [19,153], the success rate is considerably lower than for SCLC using current methods. A marker-independent methodology to isolate CTCs for engraftment is especially important in NSCLC as the partial epithelial-to-mesenchymal transition may result in loss of the epithelial CTC capture and detection markers commonly used in some CTC platforms. Short-term CTC cultures are emerging as alternatives to study the biology of metastatic disease [154]. Another method has been to culture metastatic NSCLC cells from malignant pleural effusions before injection as xenografts [155]. Lung cancer metastasis can also be comprehensively characterized in autopsy settings but there are obvious challenges in maintaining cell viability for the establishment of pre-clinical models, with key parameters including the intervals to both refrigeration and to sampling. Nevertheless, early reports suggest that model derivation is possible through rapid autopsies (ideally performed within 24 h post-mortem) in both primary NSCLC and in the metastasis of cancers from other organ sites to the lungs [156].

The injection of cancer cell lines into immunocompromised mice is an extensively reported *in vivo* cancer metastasis model [157]. These experiments are particularly relevant in lung cancer because cells injected into the tail vein circulate to the lungs, providing a model of metastasis when cells become lodged in small alveolar capillaries. However, such experiments are limited by the drawbacks of prior



**Table 3.** Outstanding questions in the application of pre-clinical model systems to non-small-cell lung cancer.

| questions regarding the use of pre-clinical models   | discussion points   |
|--|---|
| why is there a low success rate in translatability of pre-clinical data to clinical applications?  | do we do enough orthogonal validation of pre-clinical model findings in patient samples?<br>statistically significant differences observed <i>in vitro/in vivo</i> might not translate into biological relevant differences in patients<br>pharmacokinetic differences between species  |
| how does the fact that genetically complex GEMMs traditionally lose or gain all of their modifications at the same time point impact their tumour evolution? | in human disease, subclonal modifications occur later during disease progression<br>better models with temporal modifications are needed in order to model acquired metastatic potential<br>we typically induce tumours in young mice, does age affect temporal evolution?  |
| how much does pre-clinical model choice skew research outcome?   | do we fully understand the extent to which drug responses <i>in vitro</i> are shaped by cell culture artefacts (e.g. genetic drift, culture conditions)?<br>how would existing NSCLC therapies perform across an unbiased array of pre-clinical models?<br>how should we weight evidence from different models where data are contradictory?<br>are patient sex and ethnic background sufficiently explored as variables? |
| can we model copy number-driven tumours in non-human species?  | genome organization across species is not conserved so tumour copy number evolution will be dissimilar  |
| are the molecular mechanisms leading to drug resistance shared between pre-clinical models and patients?   | species differences and non-physiological conditions might lead to resistance mechanisms not observed in the clinic   |
| how should we use knowledge from big datasets to generate hypotheses that can be investigated in pre-clinical models?  | do existing training programs sufficiently emphasize computational skills?  |
| are the clinical parameters investigated during drug screening in agreement between species?   | tumour burden monitoring and definitions of progression vary between models and patients<br>species differences might exist in the side effects profiles of new therapies   |

culture and additional issues caused by a second bottleneck imposed by an introduction to the mouse circulation and the fast-growing and homogeneous nature of the resulting tumours. Brain metastases can be initiated through intracranial or intracardiac delivery of cells in mice and more advanced models—in which human brain metastasis-initiating cells are injected via the intrathoracic route, reform lung tumours and subsequently re-seed brain metastases [158]—have focussed on the earliest steps of brain metastatic colonization with the rationale that future therapies might prevent the seeding of metastases before they form.

Neither *Kras* G12D- or *Egfr* L858R-driven LUAD GEMMs generate robust spontaneous metastasis [59], although the introduction of *Trp53* loss into the *Kras* G12D model is able to generate metastatic lesions [69,159,160]. It was recently shown that mice with additional E-Cadherin loss suffer an increased incidence of metastasis to the chest wall, lymph nodes, liver and kidney [161]. The development of new barcoding technologies, such as high-throughput barcode sequencing (Tuba-seq) [162,163], coupled with CRISPR–Cas9-mediated gene editing, will facilitate studies to determine the effects of other additional mutations on tumour evolution and to identify pro-metastatic factors [164].

It is noteworthy that the *Kras* G12D-driven GEMM also releases tumour-derived cell-free DNA (cfDNA) into circulation [165] which might provide a tractable model to study mechanistic questions about cfDNA release as liquid biopsies move towards clinical application in the early detection and relapse settings. Recently, zebrafish have emerged as a powerful alternative to mouse metastasis and drug sensitivity models as they require less patient material, are more rapidly established and their near transparency as embryos and larvae facilitates imaging studies [166]. However, concerns regarding species differences increase with evolutionary distance.

## 8. Future directions

Despite extensive progress in established and new NSCLC pre-clinical model systems, there remain a number of questions for the field to reflect upon (table 3). Pre-clinical models are unevenly distributed across the disease course, with an overwhelming majority representative of primary NSCLC and relatively few available in the pre-invasive, metastatic, progression or relapse settings. Efforts to prioritize early

detection in lung cancer are likely to lead to increased diagnosis of pre-invasive and early stage disease, making new and improved experimental systems to model and predict the trajectory of pre-cancerous lesions invaluable. An expanding knowledge of the mutations and copy number events that are tolerated and positively selected within physiologically functional epithelium will also help to guide more sophisticated models of early NSCLC.

In addition to deriving new models, there are opportunities to improve the efficiency and reproducibility of existing approaches. Creating and distributing cell lines, GEMMs and xenograft models that generate reproducible results is a priority. In doing so, effort should be made to ensure that patient-derived models are from well-characterized patients with in-depth molecular and genetic characterization and high-quality control standards. For example, human epithelial cell markers should be monitored in PDX models as xenografts can form human lymphomas or mouse sarcomas [167], while primary cell cultures should be monitored for contamination with non-tumour stromal or epithelial cells.

Care should also be taken to address biases in our development of patient-derived models relating to patient sex, ethnicity and cancer genomics. Most pre-clinical models in Western countries are derived from male Caucasian or Hispanic patients, with few from patients of other ethnicities, and they typically represent smoking-associated NSCLC. For example, patients of East Asian descent constitute only 2% of NSCLC cell lines in US-based collections [33], meaning that never-smoker, EGFR-driven NSCLC is underrepresented. Effective international efforts to share pre-clinical models are required to help to address these issues and allow the collection of sufficient numbers of well-described models for more diverse patient groups and those with rare mutations.

Finally, the design of NSCLC animal experiments might also benefit from mimicking treatment regimens and study designs that are used in clinical practice. Commonly used endpoints in clinical studies—such as overall survival, progression-free survival, time to progression and overall response rate—map poorly to studies in mice. This is particularly true when using subcutaneous tumour/cell lines xenografts whose growth have limited systemic impact. GEMMs might be amenable to more comparable dosing, dose regimens and sequential therapeutics [168]. Mouse studies might also more closely mimic RECIST guidelines in which a partial response is defined as a 30% loss and progressive disease as a 20% increase in the sum of lesion diameters against baseline [169]. The study of progression, metastasis and relapse is also complicated by the divergent approaches taken in mouse and human studies. Surgical resection of primary tumours in rodent models is challenging, particularly due to multifocal tumour growth. Indeed, even when this is less problematic, such as in breast cancer models [170], the long and variable latency of metastatic lesion appearance creates challenges for timing therapeutic interventions in most current mouse models.

## 9. Conclusion

In the past several decades, our understanding of NSCLC has moved from histopathological depictions, through an increased comprehension of molecular and genetic causes towards

understanding the effects of the tumour microenvironment and the molecular dynamics of tumour evolution. Pre-clinical lung cancer research continues to become more multidisciplinary with contributions from the fields of developmental biology, stem cell biology and immunology helping to decipher interactions between tumour cells and their environment. In addition, it is vital that new clinical knowledge is fed back to improve pre-clinical models. While no models are (or are ever likely to be) able to fully recapitulate the phenotypes and responses of patient tumours, the application of multiple approaches with an awareness of their limitations is driving progress in the field of NSCLC.

**Data accessibility.** This article has no additional data.

**Authors' contributions.** R.E.H., K.K.F., D.R.P. and E.G. wrote the draft manuscript. All authors critically revised the manuscript and approved the final version.

**Competing interests.** C.S. acknowledges grant support from Pfizer, AstraZeneca, Bristol Myers Squibb, Roche-Ventana, Boehringer-Ingelheim, Archer Dx Inc (collaboration in minimal residual disease sequencing technologies) and Ono Pharmaceutical, is an AstraZeneca Advisory Board member and Chief Investigator for the MeRmaiD1 clinical trial, has consulted for Pfizer, Novartis, GlaxoSmithKline, MSD, Bristol Myers Squibb, Celgene, AstraZeneca, Illumina, Genentech, Roche-Ventana, GRAIL, Medixi, Bicycle Therapeutics and the Sarah Cannon Research Institute, has stock options in Apogen Biotechnologies, Epic Bioscience and GRAIL, and has stock options and is co-founder of Achilles Therapeutics. C.S. holds European patents relating to assay technology to detect tumour recurrence (PCT/GB2017/053289), targeting neoantigens (PCT/EP2016/059401), identifying patent response to immune checkpoint blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221) and identifying patients who respond to cancer treatment (PCT/GB2018/051912), a US patent relating to detecting tumour mutations (PCT/US2017/28013), and both a European and US patent related to identifying insertion/deletion mutation targets (PCT/GB2018/051892). C.D. acknowledges research funding/grants received from AstraZeneca, Astex Pharmaceuticals, Bioven, Amgen, Carrick Therapeutics, Merck AG, Taiho Oncology, GSK, Bayer, Boehringer Ingelheim, Roche, BMS, Novartis, Celgene, Epigene Therapeutics Inc., Angle PLC, Menarini and Clearbridge Biomedics. C.D. has received honoraria for consultancy/advisory board from Biocartis, Merck, AstraZeneca and Illumina.

**Funding.** R.E.H. is a Wellcome Trust Sir Henry Wellcome Fellow (WT209199/Z/17/Z). R.E.H. and D.R.P. are supported by the Cancer Research UK (CRUK) Lung Cancer Centre of Excellence, the Roy Castle Lung Cancer Foundation and the James Tudor Foundation. E.G. is funded by an ERC Advanced Grant (PROTEUS, Grant Agreement no. 835297). C.S. is Royal Society Napier Research Professor (grant no. RP150154). His work is supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (grant no. FC001169), the UK Medical Research Council (grant no. FC001169) and the Wellcome Trust (grant no. FC001169). C.S. is funded by Cancer Research UK (TRACERx, PEACE and CRUK Cancer Immunotherapy Catalyst Network), Cancer Research UK Lung Cancer Centre of Excellence, the Rosetrees Trust, Butterfield and Stonegate Trusts, NovoNordisk Foundation (ID16584), Royal Society Research Professorship Enhancement Award (RP/EA/180007), the NIHR BRC at University College London Hospitals, the CRUK-UCL Centre, Experimental Cancer Medicine Centre and the Breast Cancer Research Foundation, (BCRF). His research is supported by a Stand Up To Cancer–LUNGeVity–American Lung Association Lung Cancer Interception Dream Team Translational Research grant no. (SU2C-AACR-DT23-17). Stand Up To Cancer is a programme of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C. C.S. also receives funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) Consolidator grant no. (FP7-THESEUS-617844), European Commission ITN (FP7-PloidyNet 607722), an ERC Advanced grant no. (PROTEUS) from the European Research Council under the European Union's

Horizon 2020 research and innovation programme (835297) and Chromavision from the European Union's Horizon 2020 research and innovation programme (665233). C.D. and K.K.F. are funded by CRUK core funding to the CRUK Manchester Institute (A27412). Work in C.D.'s Biomarker Centre is funded by the CRUK Manchester Centre Award (A25254), the CRUK Lung Cancer Centre of Excellence, the Manchester Experimental Cancer Medicine

Centre and is supported by the NIHR Manchester Biomedical Research Centre.

**Acknowledgements.** The authors thank Dr Doraid Alrifai, Dr Adam Pennycuik and Dr Kyren Lazarus (UCL Respiratory, University College London), as well as members of the Cancer Evolution and Genome Instability Laboratory (UCL Cancer Institute and the Francis Crick Institute) for helpful discussions and feedback on a draft manuscript.

## References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018 Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. (doi:10.3322/caac.21492)
- Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. 2014 Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat. Rev. Cancer* **14**, 535–546. (doi:10.1038/nrc3775)
- Ferone G, Lee MC, Sage J, Berns A. 2020 Cells of origin of lung cancers: lessons from mouse studies. *Genes Dev.* **34**, 1017–1032. (doi:10.1101/gad.338228.120)
- Tata PR, Rajagopal J. 2017 Plasticity in the lung: making and breaking cell identity. *Development* **144**, 755–766. (doi:10.1242/dev.143784)
- Inamura K. 2018 Update on immunohistochemistry for the diagnosis of lung cancer. *Cancers (Basel)* **10**, 72. (doi:10.3390/cancers10030072)
- Cancer Genome Atlas Research, N. 2012 Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525. (doi:10.1038/nature11404)
- Cancer Genome Atlas Research, N. 2014 Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550. (doi:10.1038/nature13385)
- Jamal-Hanjani M *et al.* 2017 Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* **376**, 2109–2121. (doi:10.1056/NEJMoa1616288)
- Moore AR, Rosenberg SC, McCormick F, Malek S. 2020 RAS-targeted therapies: is the undruggable drugged? *Nat. Rev. Drug Discov.* **19**, 533–552. (doi:10.1038/s41573-020-0068-6)
- Pignon JP *et al.* 2008 Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *J. Clin. Oncol.* **26**, 3552–3559. (doi:10.1200/JCO.2007.13.9030)
- Garon EB *et al.* 2019 Five-year overall survival for patients with advanced non small-cell lung cancer treated with pembrolizumab: results from the phase I KEYNOTE-001 study. *J. Clin. Oncol.* **37**, 2518–2527. (doi:10.1200/JCO.19.00934)
- Reck M *et al.* 2016 Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N. Engl. J. Med.* **375**, 1823–1833. (doi:10.1056/NEJMoa1606774)
- Havel JJ, Chowell D, Chan TA. 2019 The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat. Rev. Cancer* **19**, 133–150. (doi:10.1038/s41568-019-0116-x)
- Lim SM, Syn NL, Cho BC, Soo RA. 2018 Acquired resistance to EGFR targeted therapy in non-small cell lung cancer: mechanisms and therapeutic strategies. *Cancer Treat Rev.* **65**, 1–10. (doi:10.1016/j.ctrv.2018.02.006)
- Gelatti ACZ, Dylon A, Santini FC. 2019 Optimizing the sequencing of tyrosine kinase inhibitors (TKIs) in epidermal growth factor receptor (EGFR) mutation-positive non-small cell lung cancer (NSCLC). *Lung Cancer* **137**, 113–122. (doi:10.1016/j.lungcan.2019.09.017)
- Vaira V *et al.* 2010 Preclinical model of organotypic culture for pharmacodynamic profiling of human tumors. *Proc. Natl Acad. Sci. USA* **107**, 8352–8356. (doi:10.1073/pnas.0907676107)
- Hassell BA, Goyal G, Lee E, Sontheimer-Phelps A, Levy O, Chen CS, Ingber DE. 2017 Human organ chip models recapitulate orthotopic lung cancer growth, therapeutic responses, and tumor dormancy in vitro. *Cell Rep.* **21**, 508–516. (doi:10.1016/j.celrep.2017.09.043)
- Dijkstra KK *et al.* 2018 Generation of tumor-reactive T cells by co-culture of peripheral blood lymphocytes and tumor organoids. *Cell* **174**, 1586–1598.e1512. (doi:10.1016/j.cell.2018.07.009)
- Morrow CJ *et al.* 2016 Tumorigenic non-small-cell lung cancer mesenchymal circulating tumour cells: a clinical case study. *Ann. Oncol.* **27**, 1155–1160. (doi:10.1093/annonc/mdw122)
- Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP. 1973 In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J. Natl Cancer Inst.* **51**, 1417–1423. (doi:10.1093/jnci/51.5.1417)
- DeMayo FJ, Finegold MJ, Hansen TN, Stanley LA, Smith B, Bullock DW. 1991 Expression of SV40 T antigen under control of rabbit uteroglobin promoter in transgenic mice. *Am. J. Physiol.* **261**, L70–L76. (doi:10.1152/ajplung.1991.261.2.L70)
- Johnson L, Mercer K, Greenbaum D, Bronson RT, Crowley D, Tuveson DA, Jacks T. 2001 Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116. (doi:10.1038/35074129)
- Jackson EL, Willis N, Mercer K, Bronson RT, Crowley D, Montoya R, Jacks T, Tuveson DA. 2001 Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248. (doi:10.1101/gad.943001)
- Coursen JD, Bennett WP, Gollahon L, Shay JW, Harris CC. 1997 Genomic instability and telomerase activity in human bronchial epithelial cells during immortalization by human papillomavirus-16 E6 and E7 genes. *Exp. Cell Res.* **235**, 245–253. (doi:10.1006/excr.1997.3670)
- Lechner JF, Haugen A, McClendon IA, Pettis EW. 1982 Clonal growth of normal adult human bronchial epithelial cells in a serum-free medium. *In Vitro* **18**, 633–642. (doi:10.1007/BF02796396)
- Andrus SB, Meissner GF, Whorton CM. 1951 Heterologous transplantation of human neoplasms including lymphomas. *Cancer* **4**, 1015–1024. (doi:10.1002/1097-0142(195109)4:5<1015::aid-cncr2820040517>3.0.co;2-z)
- Andervont HB. 1934 The production of dibenzanthracene tumors in pure strain mice. *Public Health Rep. (1896–1970)* **49**, 620–624. (doi:10.2307/4581185)
- Frisch AW, Jentoft V, Barger R, Losli EJ. 1955 A human epithelium-like cell (maben) derived from an adenocarcinoma of lung: isolation, continuous propagation, and effects of selected viruses. *Am. J. Clin. Pathol.* **25**, 1107–1112. (doi:10.1093/ajcp/25.10.1107)
- Geraghty RJ *et al.* 2014 Guidelines for the use of cell lines in biomedical research. *Br. J. Cancer* **111**, 1021–1046. (doi:10.1038/bjc.2014.166)
- Wilding JL, Bodmer WF. 2014 Cancer cell lines for drug discovery and development. *Cancer Res.* **74**, 2377–2384. (doi:10.1158/0008-5472.CAN-13-2971)
- Stead LF *et al.* 2013 An integrated inspection of the somatic mutations in a lung squamous cell carcinoma using next-generation sequencing. *PLoS ONE* **8**, e78823. (doi:10.1371/journal.pone.0078823)
- Reed MV, Gey GO. 1962 Cultivation of normal and malignant human lung tissue. I. The establishment of three adenocarcinoma cell strains. *Lab Invest.* **11**, 638–652.
- Gazdar AF, Hirsch FR, Minna JD. 2016 From mice to men and back: an assessment of preclinical model systems for the study of lung cancers. *J. Thorac. Oncol.* **11**, 287–299. (doi:10.1016/j.jtho.2015.10.009)
- Gazdar AF, Girard L, Lockwood WW, Lam WL, Minna JD. 2010 Lung cancer cell lines as tools for biomedical discovery and research. *J. Natl Cancer Inst.* **102**, 1310–1321. (doi:10.1093/jnci/djq279)
- Wistuba II, Bryant D, Behrens C, Milchgrub S, Virmani AK, Ashfaq R, Minna JD, Gazdar AF. 1999 Comparison of features of human lung cancer cell lines and their corresponding tumors. *Clin. Cancer Res.* **5**, 991–1000.
- Liu Y *et al.* 2019 Multi-omic measurements of heterogeneity in HeLa cells across laboratories. *Nat.*

- Biotechnol.* **37**, 314–322. (doi:10.1038/s41587-019-0037-y)
37. Ben-David U *et al.* 2018 Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**, 325–330. (doi:10.1038/s41586-018-0409-3)
  38. Yoshimasu T *et al.* 2007 Data acquisition for the histoculture drug response assay in lung cancer. *J. Thorac. Cardiovasc. Surg.* **133**, 303–308. (doi:10.1016/j.jtcvs.2006.06.030)
  39. Karekla E *et al.* 2017 Ex vivo explant cultures of non-small cell lung carcinoma enable evaluation of primary tumor responses to anticancer therapy. *Cancer Res.* **77**, 2029–2039. (doi:10.1158/0008-5472.CAN-16-1121)
  40. Majumder B *et al.* 2015 Predicting clinical response to anticancer drugs using an ex vivo platform that captures tumour heterogeneity. *Nat. Commun.* **6**, 6169. (doi:10.1038/ncomms7169)
  41. Alsafadi HN, Uhl FE, Pineda RH, Bailey KE, Rojas M, Wagner DE, Konigshoff M. 2020 Applications and approaches for three-dimensional precision-cut lung slices: disease modeling and drug discovery. *Am. J. Respir. Cell Mol. Biol.* **62**, 681–691. (doi:10.1165/rcmb.2019-0276TR)
  42. Yoshida K *et al.* 2020 Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* **578**, 266–272. (doi:10.1038/s41586-020-1961-1)
  43. Gao B *et al.* 2017 Non-malignant respiratory epithelial cells preferentially proliferate from resected non-small cell lung cancer specimens cultured under conditionally reprogrammed conditions. *Oncotarget* **8**, 11 114–11 126. (doi:10.18632/oncotarget.14366)
  44. Hynds RE *et al.* 2018 Expansion of airway basal epithelial cells from primary human non-small cell lung cancer tumors. *Int. J. Cancer* **143**, 160–166. (doi:10.1002/ijc.31383)
  45. Sette G *et al.* 2018 Conditionally reprogrammed cells (CRC) methodology does not allow the *in vitro* expansion of patient-derived primary and metastatic lung cancer cells. *Int. J. Cancer* **143**, 88–99. (doi:10.1002/ijc.31260)
  46. Kodack DP *et al.* 2017 Primary patient-derived cancer cells and their potential for personalized cancer patient care. *Cell Rep.* **21**, 3298–3309. (doi:10.1016/j.celrep.2017.11.051)
  47. Teixeira VH *et al.* 2019 Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat. Med.* **25**, 517–525. (doi:10.1038/s41591-018-0323-0)
  48. Sachs N *et al.* 2019 Long-term expanding human airway organoids for disease modeling. *EMBO J.* **38**, e100300. (doi:10.15252/embj.2018100300)
  49. Kim M *et al.* 2019 Patient-derived lung cancer organoids as in vitro cancer models for therapeutic screening. *Nat. Commun.* **10**, 3991. (doi:10.1038/s41467-019-11867-6)
  50. Dijkstra KK *et al.* 2020 Challenges in establishing pure lung cancer organoids limit their utility for personalized medicine. *Cell Rep.* **31**, 107588. (doi:10.1016/j.celrep.2020.107588)
  51. Wang D *et al.* 2017 Molecular heterogeneity of non-small cell lung carcinoma patient-derived xenografts closely reflect their primary tumors. *Int. J. Cancer* **140**, 662–673. (doi:10.1002/ijc.30472)
  52. John T *et al.* 2011 The ability to form primary tumor xenografts is predictive of increased risk of disease recurrence in early-stage non-small cell lung cancer. *Clin. Cancer Res.* **17**, 134–141. (doi:10.1158/1078-0432.CCR-10-2224)
  53. Moro M *et al.* 2017 Establishment of patient derived xenografts as functional testing of lung cancer aggressiveness. *Sci. Rep.* **7**, 6689. (doi:10.1038/s41598-017-06912-7)
  54. McLemore TL *et al.* 1987 Novel intrapulmonary model for orthotopic propagation of human lung cancers in athymic nude mice. *Cancer Res.* **47**, 5132–5140.
  55. Dong X *et al.* 2010 Patient-derived first generation xenografts of non-small cell lung cancers: promising tools for predicting drug responses for personalized chemotherapy. *Clin. Cancer Res.* **16**, 1442–1451. (doi:10.1158/1078-0432.CCR-09-2878)
  56. Hulton CH, Costa EA, Shah NS, Quintanal-Villalonga A, Heller G, de Stanchina E, Rudin CM, Poirier JT. 2020 Direct genome editing of patient-derived xenografts using CRISPR-Cas9 enables rapid *in vivo* functional genomics. *Nat. Cancer* **1**, 359–369. (doi:10.1038/s43018-020-0040-8)
  57. Rawlins EL, Perl AK. 2012 The a'MAZE'ing world of lung-specific transgenic mice. *Am. J. Respir. Cell Mol. Biol.* **46**, 269–282. (doi:10.1165/rcmb.2011-0372PS)
  58. Weber J, Rad R. 2019 Engineering CRISPR mouse models of cancer. *Curr. Opin. Genet. Dev.* **54**, 88–96. (doi:10.1016/j.gde.2019.04.001)
  59. McFadden DG *et al.* 2016 Mutational landscape of EGFR-, MYC-, and Kras-driven genetically engineered mouse models of lung adenocarcinoma. *Proc. Natl Acad. Sci. USA* **113**, E6409–E6417. (doi:10.1073/pnas.1613601113)
  60. Westcott PM *et al.* 2015 The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature* **517**, 489–492. (doi:10.1038/nature13898)
  61. Varela E, Munoz-Lorente MA, Tejera AM, Ortega S, Blasco MA. 2016 Generation of mice with longer and better preserved telomeres in the absence of genetic manipulations. *Nat. Commun.* **7**, 11739. (doi:10.1038/ncomms11739)
  62. Xu X, Rock JR, Lu Y, Futtner C, Schwab B, Guinney J, Hogan BL, Onaitis MW. 2012 Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. *Proc. Natl Acad. Sci. USA* **109**, 4910–4915. (doi:10.1073/pnas.1112499109)
  63. Sutherland KD, Song JY, Kwon MC, Proost N, Zevenhoven J, Berns A. 2014 Multiple cells-of-origin of mutant K-Ras-induced mouse lung adenocarcinoma. *Proc. Natl Acad. Sci. USA* **111**, 4952–4957. (doi:10.1073/pnas.1319963111)
  64. Marjanovic ND *et al.* 2020 Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* **38**, 229–246.e213. (doi:10.1016/j.ccell.2020.06.012)
  65. Guerra C, Mijimolle N, Dhawahir A, Dubus P, Barradas M, Serrano M, Campuzano V, Barbacid M. 2003 Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. *Cancer Cell* **4**, 111–120. (doi:10.1016/s1535-6108(03)00191-0)
  66. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ. 2018 Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* **359**, 1118–1123. (doi:10.1126/science.aam6603)
  67. Zacharias WJ, Frank DB, Zepp JA, Morley MP, Alkhaleel FA, Kong J, Zhou S, Cantu E, Morrissey EE. 2018 Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* **555**, 251–255. (doi:10.1038/nature25786)
  68. Tammela T *et al.* 2017 A Wnt-producing niche drives proliferative potential and progression in lung adenocarcinoma. *Nature* **545**, 355–359. (doi:10.1038/nature22334)
  69. Jackson EL, Olive KP, Tuveson DA, Bronson R, Crowley D, Brown M, Jacks T. 2005 The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res.* **65**, 10 280–10 288. (doi:10.1158/0008-5472.CAN-05-2193)
  70. Green S, Trejo CL, McMahon M. 2015 PIK3CA(H1047R) accelerates and enhances KRAS(G12D)-driven lung tumorigenesis. *Cancer Res.* **75**, 5378–5391. (doi:10.1158/0008-5472.CAN-15-1249)
  71. Tran PT *et al.* 2008 Combined Inactivation of MYC and K-Ras oncogenes reverses tumorigenesis in lung adenocarcinomas and lymphomas. *PLoS ONE* **3**, e2125. (doi:10.1371/journal.pone.0002125)
  72. Politi K, Zakowski MF, Fan PD, Schonfeld EA, Pao W, Varmus HE. 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 Dev.* **20**, 1496–1510. (doi:10.1101/gad.1417406)
  73. Soda M *et al.* 2008 A mouse model for EML4-ALK-positive lung cancer. *Proc. Natl Acad. Sci. USA* **105**, 19 893–19 897. (doi:10.1073/pnas.0805381105)
  74. Arai Y *et al.* 2013 Mouse model for ROS1-rearranged lung cancer. *PLoS ONE* **8**, e56010. (doi:10.1371/journal.pone.0056010)
  75. Inoue M *et al.* 2016 Mouse models for ROS1-fusion-positive lung cancers and their application to the analysis of multikinase inhibitor efficiency. *Carcinogenesis* **37**, 452–460. (doi:10.1093/carcin/bgw028)
  76. Ferone G, Song JY, Sutherland KD, Bhaskaran R, Monkhorst K, Lambooi JP, Proost N, Gargiulo G, Berns A. 2016 SOX2 is the determining oncogenic switch in promoting lung squamous cell carcinoma from different cells of origin. *Cancer Cell* **30**, 519–532. (doi:10.1016/j.ccell.2016.09.001)
  77. Lu Y, Futtner C, Rock JR, Xu X, Whitworth W, Hogan BL, Onaitis MW. 2010 Evidence that SOX2 overexpression is oncogenic in the lung. *PLoS ONE* **5**, e11022. (doi:10.1371/journal.pone.0011022)
  78. Hai J *et al.* 2020 Generation of genetically engineered mouse lung organoid models for squamous cell lung cancers allows for the study of combinatorial immunotherapy. *Clin. Cancer Res.*

- 26, 3431–3442. (doi:10.1158/1078-0432.CCR-19-1627)
79. Mukhopadhyay A, Berrett KC, Kc U, Clair PM, Pop SM, Carr SR, Witt BL, Oliver TG. 2014 Sox2 cooperates with Lkb1 loss in a mouse model of squamous cell lung cancer. *Cell Rep.* **8**, 40–49. (doi:10.1016/j.celrep.2014.05.036)
  80. Ji H *et al.* 2007 LKB1 modulates lung cancer differentiation and metastasis. *Nature* **448**, 807–810. (doi:10.1038/nature06030)
  81. Zhang H *et al.* 2017 Lkb1 inactivation drives lung cancer lineage switching governed by polycomb repressive complex 2. *Nat. Commun.* **8**, 14922. (doi:10.1038/ncomms14922)
  82. Li F *et al.* 2015 LKB1 inactivation elicits a redox imbalance to modulate non-small cell lung cancer plasticity and therapeutic response. *Cancer Cell* **27**, 698–711. (doi:10.1016/j.ccell.2015.04.001)
  83. Shu J *et al.* 2017 Comparison and evaluation of two different methods to establish the cigarette smoke exposure mouse model of COPD. *Sci. Rep.* **7**, 15454. (doi:10.1038/s41598-017-15685-y)
  84. Chan YL *et al.* 2019 Pulmonary inflammation induced by low-dose particulate matter exposure in mice. *Am. J. Physiol. Lung Cell Mol. Physiol.* **317**, L424–L430. (doi:10.1152/ajplung.00232.2019)
  85. Witschi H, Espiritu I, Peake JL, Wu K, Maronpot RR, Pinkerton KE. 1997 The carcinogenicity of environmental tobacco smoke. *Carcinogenesis* **18**, 575–586. (doi:10.1093/carcin/18.3.575)
  86. Reymao MS *et al.* 1997 Urban air pollution enhances the formation of urethane-induced lung tumors in mice. *Environ. Res.* **74**, 150–158. (doi:10.1006/enrs.1997.3740)
  87. Tang MS, Wu XR, Lee HW, Xia Y, Deng FM, Moreira AL, Chen LC, Huang WC, Lepor H. 2019 Electronic-cigarette smoke induces lung adenocarcinoma and bladder urothelial hyperplasia in mice. *Proc. Natl Acad. Sci. USA* **116**, 21 727–21 731. (doi:10.1073/pnas.1911321116)
  88. Shimkin MB, Stoner GD. 1975 Lung tumors in mice: application to carcinogenesis bioassay. *Adv. Cancer Res.* **21**, 1–58. (doi:10.1016/s0065-230x(08)60970-7)
  89. Saffiotti U, Cefis F, Kolb LH. 1968 A method for the experimental induction of bronchogenic carcinoma. *Cancer Res.* **28**, 104–124.
  90. Nettesheim P, Hammons AS. 1971 Induction of squamous cell carcinoma in the respiratory tract of mice. *J. Natl Cancer Inst.* **47**, 697–701.
  91. Rehms S, Lijinsky W, Singh G, Katyal SL. 1991 Mouse bronchiolar cell carcinogenesis: histologic characterization and expression of Clara cell antigen in lesions induced by N-nitrosobis-(2-chloroethyl) ureas. *Am. J. Pathol.* **139**, 413–422.
  92. Hudish TM, Opincariu LI, Mozer AB, Johnson MS, Cleaver TG, Malkoski SP, Merrick DT, Keith RL. 2012 N-nitroso-tris-chloroethylurea induces premalignant squamous dysplasia in mice. *Cancer Prev. Res (Phila).* **5**, 283–289. (doi:10.1158/1940-6207.CAPR-11-0257)
  93. Tago Y, Yamano S, Wei M, Kakehashi A, Kitano M, Fujioka M, Ishii N, Wanibuchi H. 2013 Novel medium-term carcinogenesis model for lung squamous cell carcinoma induced by N-nitroso-tris-chloroethylurea in mice. *Cancer Sci.* **104**, 1560–1566. (doi:10.1111/cas.12289)
  94. Rao W *et al.* 2020 Regenerative metaplastic clones in COPD lung drive inflammation and fibrosis. *Cell* **181**, 848–864.e818. (doi:10.1016/j.cell.2020.03.047)
  95. Sato M, Shay JW, Minna JD. 2020 Immortalized normal human lung epithelial cell models for studying lung cancer biology. *Respir. Investig.* **58**, 344–354. (doi:10.1016/j.resinv.2020.04.005)
  96. Wang G *et al.* 2019 Characterization of an immortalized human small airway basal stem/progenitor cell line with airway region-specific differentiation capacity. *Respir. Res.* **20**, 196. (doi:10.1186/s12931-019-1140-9)
  97. Hiemstra PS, Grootaers G, van der Does AM, Krul CAM, Kooter IM. 2018 Human lung epithelial cell cultures for analysis of inhaled toxicants: lessons learned and future directions. *Toxicol. In Vitro* **47**, 137–146. (doi:10.1016/j.tiv.2017.11.005)
  98. Danahay H *et al.* 2015 Notch2 is required for inflammatory cytokine-driven goblet cell metaplasia in the lung. *Cell Rep.* **10**, 239–252. (doi:10.1016/j.celrep.2014.12.017)
  99. Bove PF, Dang H, Chelvaraju C, Jones LC, Liu X, O'Neal WK, Randell SH, Schlegel R, Boucher RC. 2014 Breaking the in vitro alveolar type II cell proliferation barrier while retaining ion transport properties. *Am. J. Respir. Cell Mol. Biol.* **50**, 767–776. (doi:10.1165/rcmb.2013-00710C)
  100. Kemp SJ, Thorley AJ, Gorelik J, Seckl MJ, O'Hare MJ, Arcaro A, Korchev Y, Goldstraw P, Tetley TD. 2008 Immortalization of human alveolar epithelial cells to investigate nanoparticle uptake. *Am. J. Respir. Cell Mol. Biol.* **39**, 591–597. (doi:10.1165/rcmb.2007-03340C)
  101. Barkauskas CE, Cronce MJ, Rackley CR, Bowie EJ, Keene DR, Stripp BR, Randell SH, Noble PW, Hogan BL. 2013 Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036. (doi:10.1172/JCI68782)
  102. Salahudeen AA *et al.* 2020 Progenitor identification and SARS-CoV-2 infection in human distal lung organoid. *Nature* **588**, 670–675. (doi:10.1038/s41586-020-3014-1)
  103. Youk J *et al.* 2020 Three-dimensional human alveolar stem cell culture models reveal infection response to SARS-CoV-2. *Cell Stem Cell.* **27**, 905–919. (doi:10.1016/j.stem.2020.10.004)
  104. Matano M, Date S, Shimokawa M, Takano A, Fujii M, Ohta Y, Watanabe T, Kanai T, Sato T. 2015 Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262. (doi:10.1038/nm.3802)
  105. Sato M *et al.* 2013 Human lung epithelial cells progressed to malignancy through specific oncogenic manipulations. *Mol. Cancer Res.* **11**, 638–650. (doi:10.1158/1541-7786.MCR-12-0634-T)
  106. Pagano PC *et al.* 2017 Identification of a human airway epithelial cell subpopulation with altered biophysical, molecular, and metastatic properties. *Cancer Prev. Res (Phila).* **10**, 514–524. (doi:10.1158/1940-6207.CAPR-16-0335)
  107. Correia LL *et al.* 2017 SOX2 drives bronchial dysplasia in a novel organotypic model of early human squamous lung cancer. *Am. J. Respir. Crit. Care Med.* **195**, 1494–1508. (doi:10.1164/rccm.201510-20840C)
  108. Gunning WT, Castonguay A, Goldblatt PJ, Stoner GD. 1991 Strain A/J mouse lung adenoma growth patterns vary when induced by different carcinogens. *Toxicol. Pathol.* **19**, 168–175. (doi:10.1177/019262339101900212)
  109. Spella M *et al.* 2019 Club cells form lung adenocarcinomas and maintain the alveoli of adult mice. *Elife* **8**, e45571. (doi:10.7554/eLife.45571)
  110. Chen H *et al.* 2019 Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nat. Commun.* **10**, 5472. (doi:10.1038/s41467-019-13460-3)
  111. Ghosh M, Dwyer-Nield LD, Kwon JB, Barthel L, Janssen WJ, Merrick DT, Keith RL. 2015 Tracheal dysplasia precedes bronchial dysplasia in mouse model of N-nitroso trischloroethylurea induced squamous cell lung cancer. *PLoS ONE* **10**, e0122823. (doi:10.1371/journal.pone.0122823)
  112. Basil MC *et al.* 2020 The cellular and physiological basis for lung repair and regeneration: past, present, and future. *Cell Stem Cell* **26**, 482–502. (doi:10.1016/j.stem.2020.03.009)
  113. Xiong D, Pan J, Zhang Q, Szabo E, Miller MS, Lubet RA, You M, Wang Y. 2017 Bronchial airway gene expression signatures in mouse lung squamous cell carcinoma and their modulation by cancer chemopreventive agents. *Oncotarget* **8**, 18 885–18 900. (doi:10.18632/oncotarget.13806)
  114. Beane JE *et al.* 2019 Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. *Nat. Commun.* **10**, 1856. (doi:10.1038/s41467-019-09834-2)
  115. Riobobos L, Gad EA, Treuting PM, Timms AE, Hershberg EA, Corulli LR, Rodmaker E, Disis ML. 2019 The effect of mouse strain, sex, and carcinogen dose on toxicity and the development of lung dysplasia and squamous cell carcinomas in mice. *Cancer Prev. Res (Phila).* **12**, 507–516. (doi:10.1158/1940-6207.CAPR-18-0442)
  116. Quevedo R, Smirnov P, Tkachuk D, Ho C, El-Hachem N, Safikhani Z, Pugh TJ, Haibe-Kains B. 2020 Assessment of genetic drift in large pharmacogenomic studies. *Cell Syst.* **11**, 393–401.e2. (doi:10.1016/j.cels.2020.08.012)
  117. Ghandi M *et al.* 2019 Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508. (doi:10.1038/s41586-019-1186-3)
  118. Shorthouse AJ, Peckham MJ, Smyth JF, Steel GG. 1980 The therapeutic response of bronchial carcinoma xenografts: a direct patient-xenograft comparison. *Br. J. Cancer Suppl.* **4**, 142–145.
  119. van Rijt SH, Bolukbas DA, Argyo C, Datz S, Lindner M, Eickelberg O, Konigshoff M, Bein T, Meiners S. 2015 Protease-mediated release of

- chemotherapeutics from mesoporous silica nanoparticles to ex vivo human and mouse lung tumors. *ACS Nano* **9**, 2377–2389. (doi:10.1021/nl5070343)
120. Voskoglou-Nomikos T, Pater JL, Seymour L. 2003 Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models. *Clin. Cancer Res.* **9**, 4227–4239.
121. Hata AN *et al.* 2016 Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* **22**, 262–269. (doi:10.1038/nm.4040)
122. Yin S *et al.* 2020 Patient-derived tumor-like cell clusters for drug testing in cancer therapy. *Sci. Transl. Med.* **12**, eaaz1723. (doi:10.1126/scitranslmed.aaz1723)
123. Litchfield K *et al.* 2020 Representative sequencing: unbiased sampling of solid tumor tissue. *Cell Rep.* **31**, 107550. (doi:10.1016/j.celrep.2020.107550)
124. Salmon H, Franciszkiewicz K, Damotte D, Dieu-Nosjean MC, Validire P, Trautmann A, Mami-Chouaib F, Donnadiou E. 2012 Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* **122**, 899–910. (doi:10.1172/JCI45817)
125. Mariathasan S *et al.* 2018 TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548. (doi:10.1038/nature25501)
126. Lambrechts D *et al.* 2018 Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289. (doi:10.1038/s41591-018-0096-5)
127. Hynds RE, Gowers KHC, Nigro E, Butler CR, Bonfanti P, Giangreco A, Prele CM, Janes SM. 2018 Cross-talk between human airway epithelial cells and 3T3-J2 feeder cells involves partial activation of human MET by murine HGF. *PLoS ONE* **13**, e0197129. (doi:10.1371/journal.pone.0197129)
128. Su S *et al.* 2018 CD10<sup>+</sup>GPR77<sup>+</sup> cancer-associated fibroblasts promote cancer formation and chemoresistance by sustaining cancer stemness. *Cell* **172**, 841–856.e816. (doi:10.1016/j.cell.2018.01.009)
129. Chen WJ *et al.* 2014 Cancer-associated fibroblasts regulate the plasticity of lung cancer stemness via paracrine signalling. *Nat. Commun.* **5**, 3472. (doi:10.1038/ncomms4472)
130. Navab R *et al.* 2011 Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc. Natl Acad. Sci. USA* **108**, 7160–7165. (doi:10.1073/pnas.1014506108)
131. Sahai E *et al.* 2020 A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186. (doi:10.1038/s41568-019-0238-1)
132. Huh D, Matthews BD, Mammoto A, Montoya-Zavala M, Hsin HY, Ingber DE. 2010 Reconstituting organ-level lung functions on a chip. *Science* **328**, 1662–1668. (doi:10.1126/science.1188302)
133. Mascaux C *et al.* 2019 Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature* **571**, 570–575. (doi:10.1038/s41586-019-1330-0)
134. Pennycook A *et al.* 2020 Immune surveillance in clinical regression of preinvasive squamous cell lung cancer. *Cancer Discov.* **10**, 1489–1499. (doi:10.1158/2159-8290.CD-19-1366)
135. Marks TA, Woodman RJ, Geran RI, Billups LH, Madison RM. 1977 Characterization and responsiveness of the Madison 109 lung carcinoma to various antitumor agents. *Cancer Treat. Rep.* **61**, 1459–1470.
136. Sugiura K, Stock CC. 1955 Studies in a tumor spectrum. III. The effect of phosphoramides on the growth of a variety of mouse and rat tumors. *Cancer Res.* **15**, 38–51.
137. Kaneko T, LePage GA. 1978 Growth characteristics and drug responses of a murine lung carcinoma *in vitro* and *in vivo*. *Cancer Res.* **38**, 2084–2090.
138. Yi KH *et al.* 2007 Adoptively transferred tumor-specific T cells stimulated ex vivo using herpes simplex virus amplicons encoding 4-1BBL persist in the host and show antitumor activity in vivo. *Cancer Res.* **67**, 10 027–10 037. (doi:10.1158/0008-5472.CAN-06-2391)
139. DuPage M, Cheung AF, Mazumdar C, Winslow MM, Bronson R, Schmidt LM, Crowley D, Chen J, Jacks T. 2011 Endogenous T cell responses to antigens expressed in lung adenocarcinomas delay malignant tumor progression. *Cancer Cell* **19**, 72–85. (doi:10.1016/j.ccr.2010.11.011)
140. Damo M *et al.* In press. Inducible de novo expression of neoantigens in tumor cells and mice. *Nat. Biotechnol.* (doi:10.1038/s41587-020-0613-1)
141. Kargl J *et al.* 2017 Neutrophils dominate the immune cell composition in non-small cell lung cancer. *Nat. Commun.* **8**, 14381. (doi:10.1038/ncomms14381)
142. Mestas J, Hughes CC. 2004 Of mice and not men: differences between mouse and human immunology. *J. Immunol.* **172**, 2731–2738. (doi:10.4049/jimmunol.172.5.2731)
143. Li H *et al.* 2018 Antitumor activity of EGFR-specific CAR T cells against non-small-cell lung cancer cells *in vitro* and in mice. *Cell Death Dis.* **9**, 177. (doi:10.1038/s41419-017-0238-6)
144. De Groot R *et al.* 2019 Polyfunctional tumor-reactive T cells are effectively expanded from non-small cell lung cancers, and correlate with an immune-engaged T cell profile. *Oncoimmunology* **8**, e1648170. (doi:10.1080/2162402X.2019.1648170)
145. Gettinger SN *et al.* 2018 A dormant TIL phenotype defines non-small cell lung carcinomas sensitive to immune checkpoint blockers. *Nat. Commun.* **9**, 3196. (doi:10.1038/s41467-018-05032-8)
146. Ishikawa F *et al.* 2005 Development of functional human blood and immune systems in NOD/SCID/IL2 receptor γ chain<sup>null</sup> mice. *Blood* **106**, 1565–1573. (doi:10.1182/blood-2005-02-0516)
147. Yao LC, Aryee KE, Cheng M, Kaur P, Keck JG, Brehm MA. 2019 Creation of PDX-bearing humanized mice to study immuno-oncology. *Methods Mol. Biol.* **1953**, 241–252. (doi:10.1007/978-1-4939-9145-7\_15)
148. Wang M *et al.* 2018 Humanized mice in studying efficacy and mechanisms of PD-1-targeted cancer immunotherapy. *FASEB J.* **32**, 1537–1549. (doi:10.1096/fj.201700740R)
149. Quint LE, Tummala S, Brisson LJ, Francis IR, Krupnick AS, Kazerooni EA, Iannettoni MD, Whyte RI, Orringer MB. 1996 Distribution of distant metastases from newly diagnosed non-small cell lung cancer. *Ann. Thorac. Surg.* **62**, 246–250. (doi:10.1016/0003-4975(96)00220-2)
150. Lee HW *et al.* 2015 Patient-derived xenografts from non-small cell lung cancer brain metastases are valuable translational platforms for the development of personalized targeted therapy. *Clin. Cancer Res.* **21**, 1172–1182. (doi:10.1158/1078-0432.CCR-14-1589)
151. Krebs MG *et al.* 2011 Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. *J. Clin. Oncol.* **29**, 1556–1563. (doi:10.1200/JCO.2010.28.7045)
152. Simpson KL *et al.* 2020 A biobank of small cell lung cancer CDX models elucidates inter- and intratumoral phenotypic heterogeneity. *Nat. Cancer* **1**, 437–451. (doi:10.1038/s43018-020-0046-2)
153. Que Z, Luo B, Zhou Z, Dong C, Jiang Y, Wang L, Shi Q, Tian J. 2019 Establishment and characterization of a patient-derived circulating lung tumor cell line *in vitro* and *in vivo*. *Cancer Cell Int.* **19**, 21. (doi:10.1186/s12935-019-0735-z)
154. Zhang Z, Shiratsuchi H, Palanisamy N, Nagrath S, Ramnath N. 2017 Expanded circulating tumor cells from a patient with ALK-positive lung cancer present with EML4-ALK rearrangement along with resistance mutation and enable drug sensitivity testing: a case study. *J. Thorac. Oncol.* **12**, 397–402. (doi:10.1016/j.jtho.2016.07.027)
155. Roscilli G *et al.* 2016 Human lung adenocarcinoma cell cultures derived from malignant pleural effusions as model system to predict patients chemosensitivity. *J. Transl. Med.* **14**, 61. (doi:10.1186/s12967-016-0816-x)
156. Pisapia DJ *et al.* 2017 Next-generation rapid autopsies enable tumor evolution tracking and generation of preclinical models. *JCO Precis. Oncol.* **1**. (doi:10.1200/PO.16.00038)
157. Gomez-Cuadrado L, Tracey N, Ma R, Qian B, Brunton VG. 2017 Mouse models of metastasis: progress and prospects. *Dis. Model Mech.* **10**, 1061–1074. (doi:10.1242/dmm.030403)
158. Singh M *et al.* 2018 Therapeutic targeting of the premetastatic stage in human lung-to-brain metastasis. *Cancer Res.* **78**, 5124–5134. (doi:10.1158/0008-5472.CAN-18-1022)
159. Xiao J, Glasgow E, Agarwal S. 2020 Zebrafish xenografts for drug discovery and personalized medicine. *Trends Cancer* **6**, 569–579. (doi:10.1016/j.trecan.2020.03.012)
160. Winslow MM *et al.* 2011 Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* **473**, 101–104. (doi:10.1038/nature09881)

161. LaFave LM *et al.* 2020 Epigenomic state transitions characterize tumor progression in mouse lung adenocarcinoma. *Cancer Cell* **38**, 212–228 e213. (doi:10.1016/j.ccell.2020.06.006)
162. Sinkevicius KW, Bellaria KJ, Barrios J, Pessina P, Gupta M, Brainson CF, Bronson RT, Kim CF. 2018 E-cadherin loss accelerates tumor progression and metastasis in a mouse model of lung adenocarcinoma. *Am. J. Respir. Cell Mol. Biol.* **59**, 237–245. (doi:10.1165/rcmb.2017-02100C)
163. Rogers ZN, McFarland CD, Winters IP, Seoane JA, Brady JJ, Yoon S, Curtis C, Petrov DA, Winslow MM. 2018 Mapping the in vivo fitness landscape of lung adenocarcinoma tumor suppression in mice. *Nat. Genet.* **50**, 483–486. (doi:10.1038/s41588-018-0083-2)
164. Rogers ZN, McFarland CD, Winters IP, Naranjo S, Chuang CH, Petrov D, Winslow MM. 2017 A quantitative and multiplexed approach to uncover the fitness landscape of tumor suppression in vivo. *Nat. Methods* **14**, 737–742. (doi:10.1038/nmeth.4297)
165. Chuang CH *et al.* 2017 Molecular definition of a metastatic lung cancer state reveals a targetable CD109-Janus kinase-Stat axis. *Nat. Med.* **23**, 291–300. (doi:10.1038/nm.4285)
166. Rakhit CP, Trigg RM, Le Quesne J, Kelly M, Shaw JA, Pritchard C, Martins LM. 2019 Early detection of pre-malignant lesions in a KRAS(G12D)-driven mouse lung cancer model by monitoring circulating free DNA. *Dis. Model Mech.* **12**, dmm036863. (doi:10.1242/dmm.036863)
167. Bondarenko G *et al.* 2015 Patient-derived tumor xenografts are susceptible to formation of human lymphocytic tumors. *Neoplasia* **17**, 735–741. (doi:10.1016/j.neo.2015.09.004)
168. Singh M *et al.* 2010 Assessing therapeutic responses in Kras mutant cancers using genetically engineered mouse models. *Nat. Biotechnol.* **28**, 585–593. (doi:10.1038/nbt.1640)
169. Schwartz LH *et al.* 2016 RECIST 1.1-update and clarification: from the RECIST committee. *Eur. J. Cancer* **62**, 132–137. (doi:10.1016/j.ejca.2016.03.081)
170. Coffelt SB *et al.* 2015 IL-17-producing gammadelta T cells and neutrophils conspire to promote breast cancer metastasis. *Nature* **522**, 345–348. (doi:10.1038/nature14282)