Translating the biology of lung premalignancy to accelerate lung cancer interception.

Authors:

Sarah A. Mazzilli[1*], Zahraa Rahal[2], Maral J. Rouhani[3], Sam M. Janes[3], Humam Kadara[2], Steven Dubinett[4], Avrum E. Spira[1, 5]

Author Affiliations:

[1] Sectional Computational Biomedicine, Boston University Chobanian & Avedisian School of Medicine, Boston, MA

[2] Division of Pathology-Lab Medicine, Department of Translational Molecular Pathology, MD Anderson Cancer Center, Houston, TX

[3] Lungs for Living Research Centre, UCL Respiratory, University College London, London, United Kingdom.

[4] Division of Pulmonary and Critical Care, Department of Medicine, David Geffen School of Medicine at University of California, Los Angeles, Los Angeles, CA

[5] Johnson & Johnson Innovative Medicine

*Corresponding Author (mazzilli@bu.edu)

Abstract | Over the past decade, significant progress has been made in the development of targeted and immune-based therapies for advanced-stage non-small cell lung cancer (NSCLC) patients, driven by large-scale efforts to characterize the molecular and immunological underpinnings of established cancers. Emerging advances in the biology of earliest drivers of lung carcinogenesis are poised to address a growing unmet need for patients diagnosed with premalignant lesions and early-stage NSCLC. Multicenter consortia and Pre-Cancer Atlas efforts have begun to chart the temporal and spatial maps of the molecular events driving lung premalignant lesion progression to invasive carcinoma**,** which may be leveraged to develop biomarkers for early detection and prognostication to stratify high-risk premalignant lesions likely to progress. In addition, these efforts will drive toward identifying novel therapeutic targets for the interception of lung cancer preventing premalignant transition to or eradication of early invasive carcinomas. Novel strategies are needed for translating these large multimodal datasets into clinical tools that can transform the trajectory of those at risk for lung cancer. This review summaries the progress in identifying alterations in the epithelium of the lung and microenvironment associated with the progression of normal-appearing airway and alveoli that make up the lung field to preinvasive lesions and ultimately to invasive NSCLC. Additionally, we propose a path to build on the momentum established in advanced NSCLC to translate the growing advances into early detection and interception strategies.

Lung cancer remains the leading cause of cancer-related deaths among men and women worldwide, with more annual deaths than the second (colorectal), and third (liver) cancer-related deaths combined[1]. Lung cancer is classified into two main types including non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) which are distinguished by morphology and aggressive features that result in SCLC having and 5-year overall survival rate of 6-8%, compared 23-33% for all NSCLCs[2]. The majority of lung cancers diagnosed worldwide are NSCLC, which include the subtypes of adenocarcinoma (LUAD), squamous cell carcinoma (LUSC) and large cell carcinomas, with LUAD being the most prevalent in most countries among men and women[3]. LUAD develops deep in the lungs whereas LUSC develops in the conducting airway. These two predominant subtypes of lung cancer have been the focus of studies characterizing preneoplastic transitions, as each have well documented premalignant histology unlike SCLC[4]. Lung precancerous, premalignant, preneoplastic and preinvasive lesions [G] are all terms, often used interchangeably by pathologists and non-pathologists, to describe the histopathologic changes that precede invasive pathology indicative of carcinoma. In this Review, the concept of premalignancy is driven by histopathologic assessment that identifies morphologic and structural changes including bronchial dysplastic

changes in the airway associated with LUSC and adenomatous hyperplastic changes to alveolar tissue associated with LUAD. The definition of what premalignant means is likely to continue to evolve to focus on further defining term(s) that differentiate those lesions that are most likely to progress to aggressive invasive cancers, with associated increased morbidity and mortality.

Over the past decade, the significant advances in the understanding and treatment of advanced NSCLC, with the expansion of targeted and immune therapies[5], catalyzed by large-scale molecular studies including The Cancer Genome Atlas (TCGA), which characterized the most predominant subtypes of NSCLC; LUAD[6] and LUSC[7]. In addition to improvements in treating advanced stage NSCLC, low-dose computed tomography (LDCT) [G] screening for early detection among high-risk smokers has been demonstrated to reduce lung cancer mortality up to 20% in the US National Lung Screening (NLST) and Dutch-Belgian Randomized Lung Cancer Screening Trial (NELSON) Trials[8,9]. With the adoption of LDCT screening in the US, more lung cancer cases are being detected at earlier stages of disease with 50% or greater of screen detected lung cancers diagnosed as stage IA/IB[10]. Sustained cures ($\geq$ 5years) can be achieved with surgery for most Stage IA/IB, with increasing survival with neoadjuvant/adjuvant therapy, which has contributed to an overall reduction of 24-33% in lung cancer mortality[8,11,12]. However, numerous emerging unmet needs persist, particularly related to the management of premalignant lesions identified with LDCT screening as single or multiple hazy areas described as ground glass opacities (GGOs) [G] in the distal lung[13]. Distinguishing non-neoplastic GGOs (infection, etc.,) from those that represent a progressive or indolent adenocarcinoma premalignancy or early-stage malignancy is a major challenge. The confirmation of malignancy often requires either following a GGO serially and/or being subject to biopsy[14]. Recent advances in robotic bronchoscopy have improved the ability to biopsy potential premalignant parenchymal lesions in the lung periphery aimed to reduced transthoracic needle biopsy through the chest wall[15], however there are a lack of clinical tools or molecular test to identify which premalignant lesions are high-risk and destined to progress to invasive and/or aggressive lung cancer. In addition to peripheral precancerous lesions, auto-fluorescent and narrow-band imaging bronchoscopy [G] can be used to identify premalignant squamous lesions in the central airway, but these techniques are not routinely available at all medical centers, nor are they recommended as part of routine lung cancer screening due to the limited benefit-risk ratio[16] as the natural history of premalignant lesions are variable with only a subset of lesions ever progressing to invasive disease[17], however the exact proportion and biologic distinction are still under investigation.

Smoking cessation has been proven effective in reducing lung cancer incidence[18], although a significant proportion of lung cancer diagnoses in the US are now among former smokers[19], with an increasing incidence in individuals who have never smoked[20]. The impact of smoking cessation in patients harboring premalignant lesions has not be clearly established however, persistence and progression of LUSC premalignant lesions has been demonstrated to be higher active smokers[21]. The efforts to develop lung prevention or intervention [G] strategies have fallen short due to several critical challenges[17]. First and foremost is the inability to accurate to identify at-risk individuals who are harboring progressive premalignant lesion either in their central airway or deep in the lung periphery and then distinguishing progressive lesions that are likely to become invasive. Identifying progressive lesions is a key path to test therapeutic agents like those prescribed in advanced disease, as intervention agents that may have toxicity not generally tolerable in the primary and secondary prevention [G] setting in healthy individuals with low risk for progression. The second critical challenge is translating the molecular findings from large scale profiling to mechanistic studies to refine the identification well-validated therapeutic targets for lung cancer interception to prevent premalignant and or early invasive transitions to aggressive tumors, where no intervention options currently are available for those with premalignant disease. The translation of intervening in premalignant progression has historically been challenging in part due to the lack of pre-clinical models that effectively reflect the cellular changes identified in human lesions and allow for refining or deriving novel intervention strategies, however discovery efforts in murine and other models is rapidly evolving with molecular techniques that employ CRISPR/Cas9 technology to alter many molecular profiles simultaneously[22–25]. Lastly, clinical endpoints for prevention trials have been previously difficult to achieve as trials need to enroll large number of individuals with premalignant lesions to identify positive outcomes[17]. Historically prevention trials have enrolled patients with squamous premalignant lesion due to the accessibility with bronchoscopy[17], where bronchial premalignant lesions

have the capacity regress naturally making it nearly impossible to know which resolved due to intervention. However surrogate endpoint biomarkers based on biological pathways, enabled by mechanistic studies could facilitate better assessment of and intervention approach. There is a clear need refine interventions to pathways driving early lung carcinogenesis from the mounting data from large-scale molecular characterization of early lesions to test in refined pre-clinical models for the validation and testing of intervention strategies.

In this Review, we discuss the advances in the characterization of the molecular and cellular phenotypes associated with lung or airway injury and the evolution and progression of precancerous cells that evolved from normal-appearing airway and alveoli to preinvasive lesions and ultimately progress to invasive NSCLC highlighting possible gaps that on-going atlas efforts aim to resolve. We go on to propose central biologic processes that span the precancerous lesions of the NSCLC subtypes of LUSC and LUAD, which we suggest that the translational research field can and will build on the momentum established in advanced NSCLC to translate the growing advances into early detection and interception strategies. We detail possible paths to intervention, which require the development of biomarkers that can predict premalignant lesions at greatest risk of progression and identify candidates for interception of the most promising targetable pathways.

The lung "field of injury" & "field cancerization"

Understanding of the biology of early lung carcinogenesis begins with the initial molecular changes in the cells of the bronchial or alveolar compartments or "field" often due to exposures of known (i.e., tobacco) or unknown carcinogen, genomic events of unknown origin (i.e. replication error[26]) that can lead to the establishment of histologic changes that appear as precancerous lesions and/or carcinomas. The term "field of cancerization" [G], was first described in 1953, where "non-lesional" tissues neighboring oropharyngeal carcinoma were found to comprise cellular states (e.g., pathologic atypia) and phenotypic traits shared with malignancy[27]. With increasing resolutions of field of cancerization, it is now further defined by pathologic and molecular[27] (*i.e.* genomic[28] epigenomic and transcriptional[29]) alterations that are shared between histologically non-tumor areas in proximity to or prior to the establishment of an adjacent cancer, for example *TP53* genomic alterations in dysplastic lesions of resected tissues[28] or alterations in genes that promote cell growth in normal areas of resected tissue[29]. A related concept that followed field cancerization is the "field of injury [G]" or "etiologic field effect" that precedes the shared with tumor constraints of field cancerization to alterations often incurred by exposure to carcinogens of known or unknown origin reflecting damage directly caused by carcinogens and the respective host mechanisms of defense throughout the life of an individual. The field of injury reflects alterations that likely preceded carcinogenesis that may create a niche to mediate or combat carcinogenesis in normal appearing epithelium. The understanding of the biologic consequences of the field of injury that progress toward cancerization and premalignant transformation, should be leveraged to identify and intervene lesions early[30] (**Figure 1**).

Active smoking has been shown to alter genes in the field of injury reversibly and irreversible which may give rise to the cancerized field with potential to progress to carcinoma. Genes associated with xenobiotic detoxification[31] are rapidly reversable with smoking cessation[32]. In contrast, genes associated with cell adhesion, polarity and differentiation are often irreversibly altered[32]. Bronchial airway cell proportions and states also reflect the consequence of smoking, both histologically and resolved with single cell gene expression, show shifts towards increased secretory cells and reduced multiciliated cells, that may contribute to niches where field cancerization may emerge[33]. The establishment of the smoking field, drove the identification of field of injury changes in cancer-associated pathways in the normal-appearing nasal and bronchial airway epithelial cells. Gene changes included increased cell cycle and inflammation as well as down regulation of antioxidant defense and DNA repair suggesting the transition to field cancerized state in the absence of histologic changes[32,34]. These cancer-associated gene expression changes in the bronchial airway field in airway brushings collected by bronchoscopy have been used to develop a clinical biomarker to aid in the clinical diagnosis of NSCLC[35,36], which has recently been extended to the nasal epithelium in order to serve as a non-invasive biomarker for lung cancer diagnosis in the pulmonary nodule setting[37]. More recent work has bridged the transcriptomic field changes in active smokers and those harboring lung

cancer to the gene expression changes in the bronchial airway field of smokers with premalignant lesions that suggest increased oxidate phosphorylation and increased cell cycle in the normal airway which may be leveraged to stratify smokers at high risk for progression of premalignancy to invasive disease with further refinement to narrow down those most at risk[38,39]. Studies are on-going to establish whether field level changes can be identified in the proximal and distal airway in association with preinvasive LUAD as the detection of these lesions increases with LDCT screening[37].

Fig.1 comments

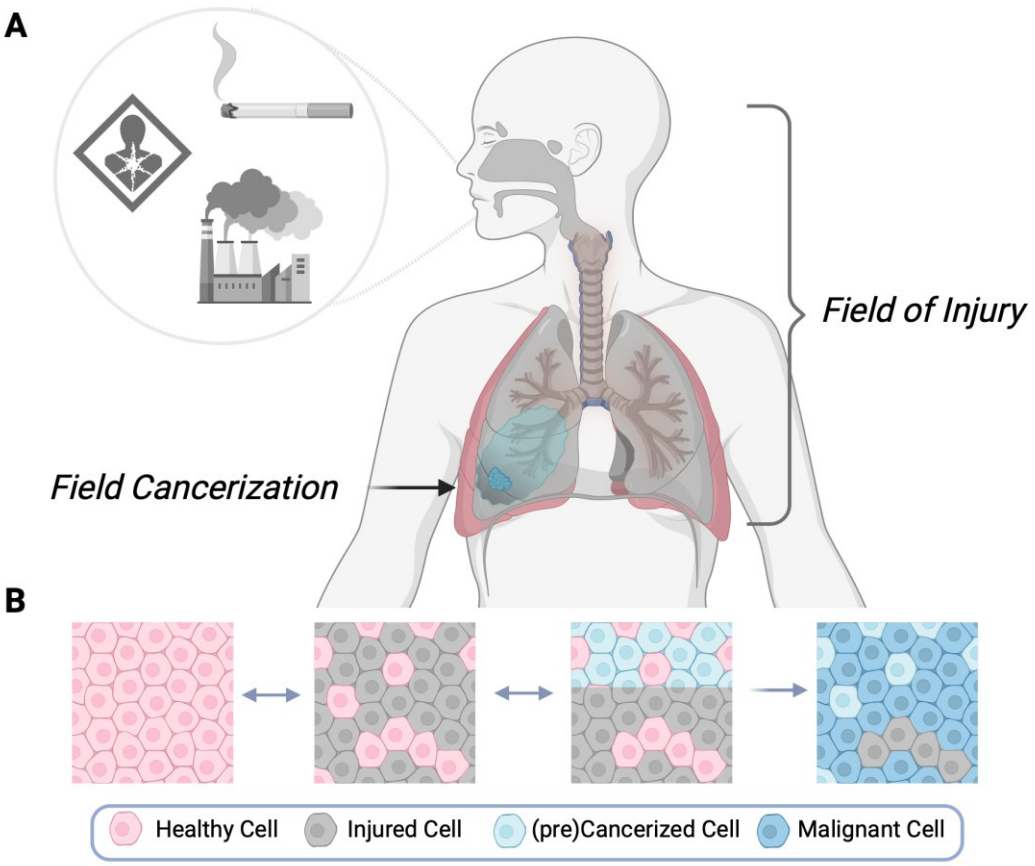


Figure 1. Field level alterations in the lung. A| Inhaled exposures span the entire airway due to create and etiologic field of injury (grey) that can lead to cancer risk over the course of ones life. By contrast focal molecular changes that create field of cancerized cells that can precede or are adjacent to the development of lung cancer at a specific pulmonary site which can be removed when a lesion/tumor is resected (teal). **B|** The cellular landscape of changes that accumulate throughout field include cells undergoing injury response to exposure (injured cell) such increased antioxidant defense and DNA repair. Continued exposures further increase cellular injury and DNA damage that result in both morphologic and genomic changes driving precancerous lesions to develop ((pre)cancerized cell) with additional alterations cells can transition to become cancerized (malignant cell) and become invasive cancer or in many cases areas of damage undergo repair and regress back toward a health state.

In addition to altered gene expression in the field, there is growing evidence that extends the mutation burden of field cancerization to the mutagenized field of injury in the lung that precede the accumulation of known tumor mutations or histologic changes found in the cancerized field. This mutagenized field of injury can yield an array of somatic mutations that tilt the odds toward a localized field cancerization with the potential progress to a precancerous lesion or tumor. However, this may is not always the case as some expanding positively selected mutant clones may outcompete micro-tumors formed within mutagenized fields as shown in the context of

esophageal[40] and colon[41] carcinogenesis in mice and in the esophagus[42] of humans. However this remains to be demonstrated formally in the lung, although somatic point mutations or chromosomal aberrations (allelic imbalance, copy number variation) in key oncogene drivers (e.g., *KRAS*, *PIK3CA*) and tumor suppressors (*TP53*, *KEAP1*, *CDKN2A*) are found in normal-appearing tissues and shared with NSCLCs, in tumor-adjacent relative areas opposed to more distant normal-appearing regions[43]. Further insights to mutational field of the airway have come from work studying colonies of primary airway basal cells derived from single-cell sorted bronchial epithelial cells of cancer-free patients. Basal cells harbored considerable heterogeneity in mutational burdens across cells based on smoking status; with a subset of cells from current and former smokers with mutational burdens in the range expected of age-matched never smokers, while others cells have a four-fold higher relative frequency of mutations than former smokers[44]. Mutational signatures in cancer-free smokers were positively correlated with the predominant signature of smoking-induced lung cancer, albeit to a considerably lesser extent in former smokers when compared to current[45,46], suggesting that these mutagenized fields could portend a possible malignant transformation which might be halted by stopping exposure to the insult (either by smoking cessation or by mitigating inflammation). A strong inverse relationship is also found between telomere length and mutational burden in current and ex-smokers, where cells with a near-normal mutational burden in the latter group had longer telomeres, suggesting that these cells would have undergone fewer cell divisions[44,47]. This raises the question of whether these cells exist as quiescent stem cells, perhaps in a physically protected niche.

Despite major advances in our understanding of the field of injury and cancerization fueled novel approaches, like single-cell technology, the field still requires continued refinement needed to translate field alterations detected in the proximal airway (the lateral aspect of the inferior nasal turbinate or endoscopic brushings of the mainstem bronchus) to the development of prognostic biomarkers defined by transcriptional and/or mutational landscape of neighboring "normal- appearing" or "non-lesional" tissues. Furthermore, field changes may also provide an opportunity to identify and develop targeted treatments that can increase the fitness of "favorable" clones or alterations to halt the initiation or progression of carcinogenesis. It is hypothesized that favorable clones could explain how the lung tissue covered by field cancerization (i.e., in the local niche of the lung cancer or lesion) is more fertile for tumorigenesis to the distinct NSCLC subtypes of LUAD and LUSC compared with other regions in the lung field of injury.

Premalignant lesions and early lung carcinogenesis

Major advances in the characterization of premalignant LUSC. LUSC is the second most common type of lung cancer accounting for ~ 20-25% of cases[48]. LUSC is thought to develop through a stepwise process of progressive epithelial changes predominantly in the proximal bronchial airway[17]. The normal bronchial epithelium is comprised of pseudostratified layer of primarily basal, goblet and club secretory, and multiciliated cells and more rare cell types[49,50] like neuroendocrine, ionocytes[51,52], hillock[53] and tuft cells. The airway epithelium functions to perform mucociliary clearance and respond to irritants such as cigarette smoke[33,54]. As the normal-appearing bronchial airway sustains injury from chronic exposures or infection, histologic changes result. These changes can range from mild and largely reversible histologies, that change from the normal pseudostratified airway to hyperplasia characterized by the hyper-proliferation and expansion of cell populations (*i.e.* basal (reserve) or goblet cells) or the development of squamous metaplasia that acts to protect the airway epithelium through the flattening of the apical cell layer. More severe histologic changes, that can be preceded by hyperplasia or metaplasia, include bronchial or squamous dysplasia and squamous carcinoma in situ (CIS), which have varying reversibility[17] (**Figure 2**)**.** Dysplasia histology can be further classified as mild, moderate, or severe based on increasing levels of cellular abnormalities and polymorphisms within the basal or basal-like cells of the epithelium, from alterations in the lower third of the epithelium (mild) to the complete epithelium (severe)[55] with CIS has extensive cytologic changes that resemble carcinoma, however have and intact basement membrane and lack stromal invasion[56]. The cell of origin for LUSC is thought to have properties of basal cells, although various basal cell states have been elucidated[57], continued linage studies in human and murine squamous progression are needed.

A major challenge in classifying dysplasia as a preinvasive lesion has been driven by lack of uniformly accepted in pathologic assessments[58]. The World Health Organization (WHO) classification of tumors of the lung in 1999 and further in 2004 began to provided granular definition and were inclusive, but details of dysplasia were limited in the 2015 update and lacked inclusion of specific international classification of diseases for oncology (ICD-O) codes to uniformly name lesions[59]. The 2021 update now specifies and provides ICD-O codes by which to capture these pathologic findings[58,60–62]. However, variability across pathologists in what is considered a true preinvasive lesion continues to challenge assessment and treatment of these lesions. The identification of additional molecular markers identifying lesions will enable assessment of disease pathogenesis. Digital pathology combined with rapid development of machine learning approaches, combined with low or mid-plex protein or RNA spatial localization of key cell states may provide opportunity to further discern features indictive of high progression risk.

Early work to establish the molecular indicators of a stepwise trajectory of progression from dysplastic lesions to invasive LUSC highlighted molecular changes including the contribution of p53 mutations[63,64] and the loss of heterozygosity (LOH) on chromosome arms 3p, 9p, 8p, 13q and 17p in bronchial dysplasia and CIS lesions[65,66], as well as the association of specific chromosomal aberrations within more severe lesions[67,68]. Genomic analysis uncovered observed field cancerization, reflected by cell migration of clonal CIS lesions[69] and somatic chromosomal alterations at multiple sites[70]. These studies, although often challenged by limited sample sizes, proximity to resected tumors and their largely cross-sectional nature, provide evidence that genomic changes associated with carcinogenesis can occur in some lesions that lack invasive pathology. Additionally, evidence from observational longitudinal studies following patients with bronchial dysplastic lesions for many years suggest that up to 34% of patients progress to LUSC within 10 years[71] and that persistent high-grade dysplasia is significantly associated (adjusted HR, 7.84; 95% confidence interval, 1.56-39.39) with the development of invasive LUSC in 10 years[21]; however, clinical tests for identifying which of these patients would progress have not been established.

In the more recent past with improving technology, significant advances have been made in our understanding the biology of preinvasive lesions. This has been facilitated by the application of genomic, transcriptomic, and proteomic tools to longitudinally followed cohorts of high-risk (ever smokers, history of NSCLC, COPD, etc.) patients that have specific airway locations biopsied via bronchoscopy every 6-12 months (**Figure 2, Supplemental Table 1**). This approach has offered greater insight into the natural history of carcinogenic airway changes. Multiomic profiling has been performed on endobronchial forceps biopsies of normal (including hyperplasia and squamous metaplasia), dysplastic (mild to severe) and/or CIS lesions. Building on historic genomic studies that characterized LOH in cross sectional studies of LUSC lesions of resected tumors, whole-genome sequencing of CIS biopsy samples absent of invasive tumors revealed that lesions which regress to normal epithelium or low-grade dysplasia histology on subsequent biopsy have markedly lower mutational load and fewer copy number alterations compared to other CIS samples[72]. *TP53, CDKN2A,* and *SOX2* were mutated in CIS lesions with identified mutations, paralleling those found in late-stage disease according to TCGA data[7]. Ongoing studies aim to bridge the gap between the mutational field findings associated with tobacco and other exposures found in the normal-appearing airway and the various grades of dysplasia over time as they regress or progress to identify the mutational signatures that define early carcinogenesis to translate these historic and novel findings to biomarkers to identify progressive lesions, for intervention approaches. There have been limited large scale DNA methylation profiling of dysplastic lesions, however early work by Belinsky et al identified aberrant promoter methylation of *CDKN2A (p16)*, *MGMT* and *RASSF1A1* in atypic cells from sputum[73,74]. However, in CIS lesions, significant differences in methylation between progressive and regressive CIS lesions have been observed, where NKX2-1 (TTF-1) protein loss is a diagnostic feature of LUSC and was found to be a putative driver gene of interest with reduced expression and hypermethylation across both gene expression and methylation analyses in progressive compared to regressive CIS lesions[72].

The longitudinal collection of these biospecimens over two or more time points and often at multiple locations in each patient prior to the progression to cancer, have also highlighted transcriptional alterations associated with temporal changes in histology and progression. Dysplastic and CIS lesions harbored gene expression changes with the upregulation cell cycle and DNA replication genes including MKI67[39] with increased

expression of PLK1 in lesions that progress[75] or persist at high grades of dysplasia. In addition, genes associated with altered DNA damage response, metabolism (OXSPHOS) and cell adhesion pathways[39,75]. In CIS lesions, epithelium profiled alone also has shown increased expression of genes associated with chromosomal instability,

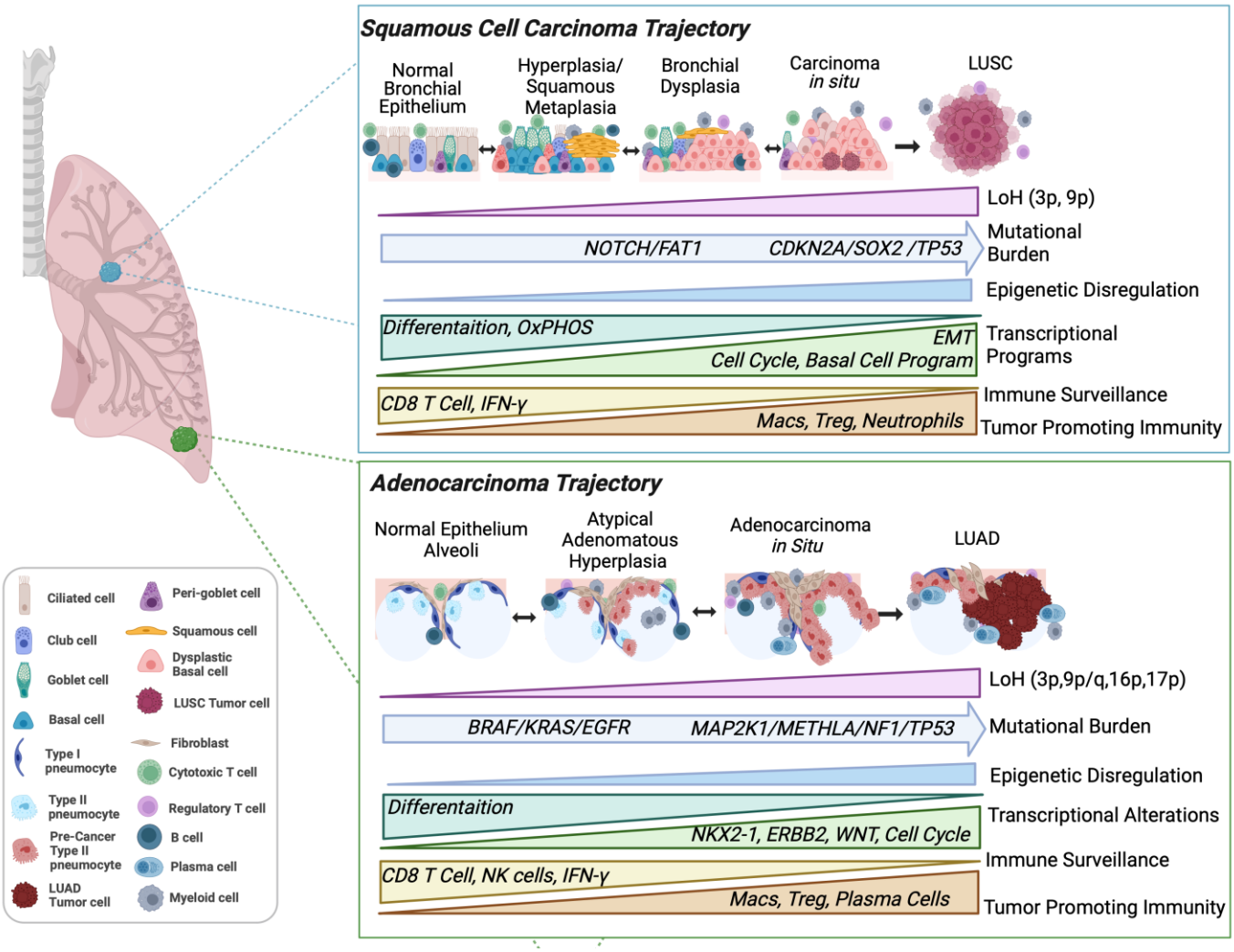


Figure 2. Major advances in the biology of pre-invasive and early stage LUSC and LUAD. (Top) The premalignant progression to LUSC lesions is histological associated with increasing cellularity of the basal like (pink) population. Characterization of bronchial lesions longitudinally as they progress and regress away from LUSC have identified the accumulation of genomic, transcriptional and epigenetic alterations that support a pro-tumor microenvironment in lesions that progress with the accumulation of genomic alterations in *TP53, CDKN2A* and *SOX2* in CIS. Transcriptional signatures of progressive LUSC lesions demonstrate and evolution from increased proliferation (*i.e.* increase in OXPHOS metabolism and cell cycle) in early dysplasia to altered cell programs (*i.e.* increased differentiation and EMT) and immune evasion (reduced cytotoxic T cells and suppressed IFN-γ signaling) as lesions progress to CIS. (Bottom) The natural history of premalignant progression to LUAD is beginning to emerge from cross-sectional approaches to identify molecular alterations in the preinvasive lesions adjacent to resected tumors with increasing alveolar cell intermediates (pink) as lesions develop into LUAD. Genomic and transcriptional alterations bear similarities to LUSC lesions that include altered differentiation, increased proliferation and pro-tumor immune profiles with increasing T regulatory and plasma cell phenotypes. *(LUSC- lung squamous cell carcinoma; LUAD- lung adenocarcinoma; LoH -loss of heterozygosity; EMT – epithelial to mesenchymal transition; IFN-γ – interferon gamma; Macs- macrophages.)*

most frequently in lesions areas that progress at a subsequent biopsy[72]. Follow on atlas studies aim to provide single resolution of the cell type and state changes, cellular plasticity and transcriptional programs associated with lesions progression such as regulation of the hippo pathway[76] are poised to aid in the refinement of targets for interception.

In addition to the epithelial changes, the evolution of the microenvironment surrounding the lesion has been increasingly shown to play a vital role in early carcinogenesis[77]. Dysplastic lesions that persist or progress towards invasive cancer have decreased expression genes regulating anti-tumor immunity and a reduction of immune infiltrate of proinflammatory macrophages recruiting CD 8+ T lymphocytes which are abundant in regressive and more low grade (hyperplasia, metaplasia mild dysplasia) and normal lesions[75]. Beane *et al.* transcriptionally define molecular subtypes of normal, dysplastic and CIS lesions that classified lesions by gene programs. This included the identification of a subset of samples that had upregulated genes associated with cell cycle named the Proliferative subtype, which was found to be enriched for lesions that were dysplastic (mild - severe) and from active smokers. Among the lesions classified as Proliferative that progressed over time to higher grades of dysplasia or remained dysplastic, they had reduced expression of genes associated with interferon gamma signaling and T cell mediated immunity, specifically antigen processing and presentation. This correlated with decreased numbers of innate and adaptive immune cells by immunofluorescence[39]. These immune-suppressed phenotypes have been further corroborated and refined with the inclusion of trajectories that included additional CIS and early LUSC lesions[78,79]; transcriptomic analysis identified 'immune cold' clusters with associated immune escape through immune checkpoints in lesions that progressed to or were LUSC. In contrast 'immune hot' clusters were evident in lesions that regressed [78]. Pennycuick *et al.* also identified enrichment of *CCL27* (*epithelium*):*CCR10* (*T cell*) interaction in progressive samples, as a possible key immune escape mechanism via the PIK/Akt pathway[78]. Together these studies depict the suppressed immune environment of progressing lesions, however more work is needed to fully understand the mechanisms and signaling cascades that drive these phenotypes, and a refined understanding of the innate immune surveillance contributions from populations such as dendritic, natural killer and neutrophil populations, that may provide continued resolution to areas that could be targeted for intervention approaches. Additionally, more effort to resolve the stromal and fibroblast contributions to the lesion microenvironment, both in the contribution to the stiffness and permeability of basement membrane and the signaling cascades that may contribute to cellular plasticity as suggested in advances LUSC[80] and with esophageal precancer transitions[81].

Collectively the work to profile LUSC premalignant lesions highlights progress in the field in identifying epithelial and microenvironmental cues that likely enable the development and progression of premalignant lesions. However, the discovery efforts described above have yet to be translated although with continued integration of these genomic, transcriptional and epigenetic landscapes with ongoing single cell and spatial studies the field is poised to provide greater insight into cellular trajectory, subcellular signaling networks, and spatial neighborhoods that can be leveraged for refining the prognostic identification of premalignant lesions destined to progress to LUSC. Continued work to improve and test they hypothesis put forth by discovery efforts is essential to drive identification of key targets for both biomarker and interception strategy development bolstering the epithelial-immune axis to facilitate the elimination of premalignant or early cancers possibly before any surgical intervention may be needed.

Major advances in premalignant LUAD. Lung adenocarcinoma (LUAD) represents the most common subtype of lung cancer (57% in the US)[48], primarily affecting ever smokers with a growing proportion of nearly 20% of lung cancer cases being in individuals that never smoked[19,20,82]. LUAD forms in the lung parenchyma and alveolar tissue and precursor lesions can arise in a cancerized field as solitary foci in the normal-appearing lung tissues or in continuity with LUAD of varying histologic subtypes[83,84,85]. The major epithelial cell types line the alveoli include type I pneumocytes (AT1 cells), thin, squamous cells responsible for gas exchange, and type II pneumocytes (AT2 cells) which secrete pulmonary surfactants necessary for structural integrity and proliferate to regenerate the lung after injury[86]. LUAD is thought to arise from the AT2 like cells lining the alveoli of the lung parenchyma[87]. Premalignant histological progression that precedes invasive LUAD is thought to include atypical adenomatous hyperplasia (AAH) and adenocarcinoma in situ (AIS) (**Figure 2**) which is supported by the progressive increase in genomic alterations , however natural history studies have been partly hindered by the peripheral location of the tumor, limiting our ability to longitudinally access and evaluate presumptive precursor lesions as they progress to invasive tumors, so the stepwise process remains unconfirmed. AAH develops as a small proliferative lesion

consisting of monolayered atypical cells lining the alveolar septa[60]. Adenocarcinoma in situ (AIS), previously referred to as bronchioloalveolar carcinoma, is now accepted as a preinvasive lesion and share histologic characteristics of AAH and are often differentiated by size with AIS being greater that 0.5 cm[60]. AIS is distinct from minimally invasive adenocarcinoma (MIA) based on size (> 3 cm) and absence of invasive features (≤5 mm), which make MIA considered invasive carcinoma and not premalignant. However, pathologically distinguishing between AAH, AIS, MIA, and regions of small lepidic[88] (absent of stromal or vascular invasion) LUAD tumors regions is challenging although the WHO now specifies and provides ICD-O codes by which to capture these pathologic findings[58,60–62], the review of the histology largely depend on the availability of tissue from a resection and the reviewing pathologist, highlighting the need for better assessment tools to distinguish lesions that may give rise to invasive and aggressive subtypes of LUAD like solid predominant patterns which are associated with poor disease free survival[89]

Although the natural history of LUAD remains to be fully resolved, a large body of work has examined precancerous histology adjacent or within the resected margin of formalin fixed paraffin embedded lung tissue from individuals that underwent surgical resection for treatment of LUAD. (**Figure 2, Supplemental Table 2**). Early work depicted increasing LOH in a subset of AAH lesions on chromosome arms 3p, 9p, 9q, 16p, and 17p[90,91]. More recent studies have highlighted HLA LOH, disrupting antigen presentation machinery, as well as increased frequency of 6p arm-level copy number alterations in precursor lesions that progress in histologic grade to malignancy[92,93]. Of note though, some AAH lesions were found to exhibit shared allelic imbalance changes with LUAD including losses harboring tumor suppressor on 17p, 9p, 9q and gains encompassing oncogenes on 8q, 12p, and 1q[94], however without longitudinal studies it's unclear if these alterations would suggest progression to invasion if identified in and AAH alone. More recently, targeted DNA sequencing of macro-dissected preinvasive lesions and early LUAD found that *EGFR, KRAS* and *BRAF* mutations were the most abundant alterations among all AAH lesions, were individual AAH were divergent on *BRAF* or *KRAS* alterations AAH[93,95]. The roles *KRAS* have been heavily described as a driver of LUAD in human and mice[96], however the characteristics of these *BRAF*-mutant clones in the lung are yet to be elucidated, given their exclusiveness to AAHs and their absence from LUADs. It could be postulated that carrying *BRAF* mutations in AAHs limits the propagation of these clones into malignancy or confers better recognition by the immune system leading to their elimination. Interestingly, Dankort and colleagues found that mice genetically engineered to express *BrafV600E* develop lung lesions that rarely progress to LUAD and that *BrafV600E* expression initially induced proliferation followed by senescence[97] suggesting the requirement for additional driver mutations or other hallmarks. Other significantly mutated genes reported in AIS and MIA lesions include *RBM10*, *ERBB2*, *MAP2K1*, and *MET* although limited is known about the sequence of events and how they contribute to progression to invasion[93]. Additionally, frequencies of *TP53* and *NF1* mutations were found to increase with lesion histological grade from AIS and MIA to LUAD, with *TP53* being a key mediator for the invasiveness of LUAD[93,98]. A higher proportion of clonal mutations were identified in invasive lesions compared to precursor AAHs, supporting a selective sweep of unfit subclones during the development of this malignancy[99]. In addition to mutational studied, targeted epigenetic modifications including promoter hypermethylation of *CDKN2A (p16)* and *PTPRN2* in AAH lesions potentially inferring growth advantages and *HOXA1, HOXA11*, *TMEFF* in AIS suggestive of plasticity in differentiation, and increasing genes methylated in invasive LUAD confirming invasive potential (*SFRP1, TWIST1*) [100,101].

In parallel to establishing the genomic landscape of developing lesions, invaluable insights into the early development of LUAD from transcriptional profiling have identified increases in the *NKX2-1, ERBB2, WNT* and cell cycle pathways[94]. Although insights from bulk profiling inadvertently obscuring some of the individual roles of epithelial cells subsets and their interactions with the microenvironment and redirected the focus to the transcriptome to delineate the evolutionary trajectory of epithelial programs, identifying AT2-like cells and loss of alveolar differentiation in development of locally invasive LUAD[102–104]. These findings are corroborated by a study that also found an increased frequency of AT2-like cells in nodules that are solid appearing on CT (usually invasive LUAD) as opposed to part solid (potentially for preinvasive & MIA) further supporting early hypotheses that AT2 cells might be the cell of origin of LUAD lesions[105]. A more recent study that performed single-cell sequencing of ~250,000 sorted epithelial cells from LUAD patients identified "normal" alveolar intermediate cells (Krt8+ alveolar

intermediate cells; KACs) that were enriched in the LUAD ecosystem, harbored driver *KRAS* mutations, and while they were associated with AT2 to AT1 transdifferentiation they also functioned as progenitors for LUAD in mouse carcinogenesis models with lineage-specific reporters[106–108]. In addition to the alterations in the epithelial compartment, single cell studies also highlight the role of the non-epithelial cells in contributing to pro-tumorigenic environments. Cancer-associated fibroblasts, which increase with histological grade of LUAD are thought to contribute to the deregulation of the extracellular matrix, potentially affecting immune infiltration in subsolid nodules through ligand-receptor interactions[105]. Endothelial cell changes have also been identified between normal and early LUAD samples, highlighting MYC target genes, possibly associated with angiogenic growth and IFN-γ pathway genes associated with inflammation.

In addition to alterations regulating the molecular changes to the epithelium, suppression of genes associated with IFN-γ and inflammatory responses have been identified with increasing premalignant histological grade[109]. These initial immune insights have also been correlated with progression-associated mutations and infiltration of CD8+ and CD4+ T cells, suggesting the presence of an adaptive immune response in a subset of early lesions[92,110]. Evidence of immune suppressive phenotypes and evasion of immune surveillance includes a reduction in NK and NKT cells as well as elevation of T regulatory cells associated with increasing histology in single cell analysis[102,105,111,112], with T regulatory cell phenotypes increasing with proximity to the invasive LUAD[102]. suggesting that in addition to mutations in epithelial cells, immunological changes may also underlie field cancerization specific to normal-appearing areas surrounding lung tumors or premalignant lesion that would support progression toward invasion. Although the diminished presence of CD8+ T cells has been one of the most striking findings as lesion increase in histologic severity, recent studies highlighted an increase in tumor-infiltrating memory B (TIB) cells and plasma cells, which are in more differentiated states and exhibit increased frequencies of somatic hypermutation. The TIB and plasma cells were highly prominent in early LUAD tumors compared to adjacent normal tissues, particularly in localization of CXCL13+ lymphoid aggregates, prominent in tumors. Using pre-existing RNA bulk profiling the findings were proposed to be enriched with increasing histology in AAH and AIS[113].

Collectively, the continued profiling of adjacent normal (normal in the margin to tumor or within the resected lung), premalignant lesions and early LUAD tumors has provided several avenues for continued refinement to translate toward improving prognostication and actualization of intervention strategies that focus on opportunities to target or stimulate the immune phenotypes in addition to novel mechanism to halt the progression by tipping the scales of transdifferentiation away from malignancy. On-going efforts to extend these findings particularly through longitudinal studies mapping the natural history of LUAD preinvasive lesions driven in part by the development of robotic bronchoscopy can now extend highly accurate trans-bronchial biopsies in the distal lung parenchyma. These modalities, combined with increased collaborations between surgeons and pathologists, have begun to enable the detection and sampling of early-stage lesions, and enhanced our capacity to decode the earliest events driving LUAD that can be translated into targets for its interception for LUAD. Current and on-going studies are now moving to use frozen tissue post-resection, with a growing move toward robotic biopsy methods that parallel LUSC premalignant lesion collections

Hallmarks of premalignant lung lesion progression. Advances in our understanding of preinvasive lesions has highlighted specific alterations in proximal airway and distal parenchymal lesions (**Figure 2**) includes mutations and transcriptional alterations in molecular pathways that evoke stem cell-like changes and portend growth advantage changes to the different epithelial cell subsets that populate the airway and alveolar compartments[33,107,108]. Although not all lesions identified histologically as premalignant harbor genomic and epigenetic changes that support cancer-associated phenotypes, a subset of patients will have lesions that exhibit aberrant cell cycle control, altered epithelial differentiation, increased metabolism, suppressed DNA damage response and altered cell adhesion and cell matrix, however which phenotype initiates or indicates progression remain to be fully resolved. The epithelial and stromal change and the regulation of tumor promoting inflammation and immune suppression in the lesional microenvironments, can elicit immune evasion pathways, increasing the potential to persist as a high grade

preinvasive dysplasia or CIS/AIS and ultimately progress to invasive carcinoma. Among the updated 14 defined hallmarks of cancer[114], many have been identified prior to the presence of invasive disease and the early molecular events altered in lung precancerous lesions and the airway field of injury/cancerization that may be leveraged for translational investigation. The accumulation of tumor promoting genomic, epigenetic and transcriptional alterations highlight the genome instability (LOH), mutations, and non-mutational epigenetic reprogramming (*CDKN2A, MGMT*) which enable sustaining proliferative signaling (*CDKN2A/ KRAS*) evasions of growth suppressors (*TP53*), deregulation of changes in cellular energetics, cellular plasticity and altered cell differentiation. A major parallel between progressive LUSC and LUAD lesions is the accumulation tumor promoting inflammation, immune suppression & evasion highlighted by a reduction in recruitment of effective cytotoxic T cells NK and NKT cells and increase in T regulatory cells and increased tumor associated macrophages. Further mechanistic insights in on-going studies are needed to translate the discovery of the markers of progressive lesions to interception.

Translating discoveries to accelerate lung cancer interception.

Discovery efforts have uncovered numerous molecular features that define premalignant lesions, however there is a need to translate these findings to address major unmet clinical needs for lung cancer interception. Major unmet needs the field must address to successfully translate the discovery efforts into clinical care are to: 1) develop biomarkers that can identify individuals harboring premalignant lesions utilizing non-invasive tissues such as airway brushings, blood or sputum to enhance lung cancer screening; 2) develop biomarkers that can predict premalignant lesions at greatest risk of progression to carcinoma to identify candidates for early intervention or increased screening; and 3) identify the most promising targetable pathways and the establishment of mechanistic studies in optimized representative pre-clinical models (Figure 3).

Non-invasive biomarkers to expand lung cancer screening. A major hurdle in developing successful tools for enhancing screening eligible populations is identifying individuals harboring premalignant lesions. Current approaches to assess risk for presence of established lung cancer can include assessment of demographics and clinical risk factors including age, family history of cancer, COPD status, smoking status[115]. Although, only a subset factors suggesting risk are used to for lung cancer screening, including those age 50-80 years with 20 pack-years in active smokers and previously only included those that have quit within the past 15 years[115,116]. Despite recent updates to the screening guidelines, they are still not adequate or inclusive of all those who could benefit from screening, including underrepresented populations who smoke less, but have higher rates of mortality due both to unknown biology and social determinates of health[117] as well as the growing population of lung cancer patients that have never smoked[118,119]. In individuals screened with LDCT, standard radiologic assessments of pulmonary nodules consider the size (6mm or greater) and density (non- solid or solid, with solid suggestive of invasive LUAD) to evaluate lung cancer risk and need for subsequent biopsy[116]. A growing number of computational, machine learning and artificial intelligence approaches are being developed to refine radiologic assessments and to integrate radiologic findings with molecular biomarkers measuring genomic or proteomic signatures to enhance the evaluation of risk of invasive disease[120–123]. This approach is predicted to not only improve the diagnosis of established cancer but could also be applied to the identify preinvasive disease.

Non-invasive molecular biomarkers are also being established to detect the presence of circulating malignant/premalignant cells and analytes (DNA [known primary tumor mutations, fragment size, methylation status] tumor/lesion RNA/miRNA, altered immune subsets, exosomes, autoantibodies or microbes) and many of which have demonstrated to have efficacy in detecting established cancers[124,125]. However, due to the limited size and focal nature of premalignant lesions, the threshold for detection remains challenging in this context. As new targets are identified, and assays sensitivity is improved, detection may become feasible. Sputum has also been suggested as a potential source for assessing biomarkers for premalignancy as this is an effective approach for identifying patients with bronchial dysplasia and has been long studied as a tool for detecting LUSC, but the opportunity to extend to LUAD remains to be resolved[21,126]. Airway epithelial transcriptomic alterations within the "field of injury" may also serve as a biomarker in airway epithelia brushed for detection of premalignant squamous

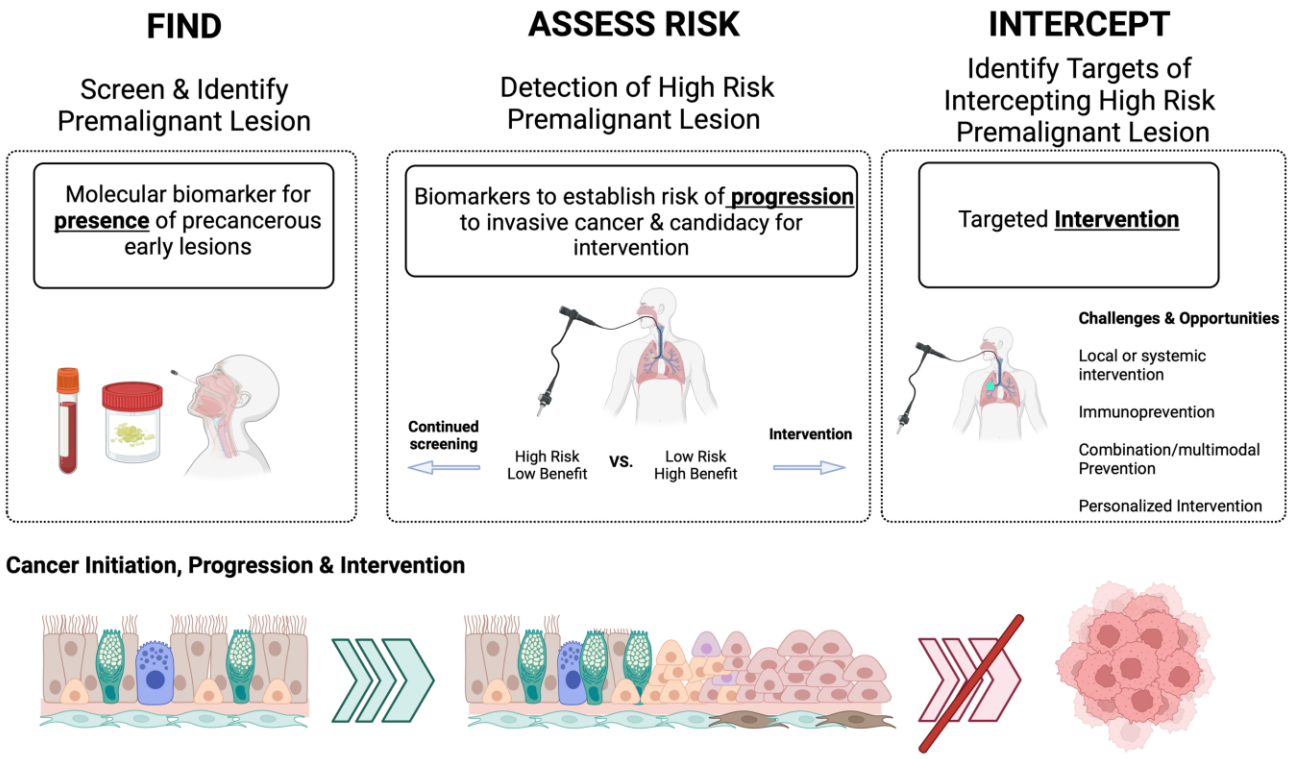
lesions[34,39,127]. Additional studies are needed to evaluate whether airway epithelial brushings can be used to detect premalignant adenocarcinoma lesions that are more distally located.

Non-invasive biomarkers of progression. Biomarkers of progressive phenotypes of premalignant disease are likely to be multi-modal in nature that bring together clinical evaluation and diagnostics using image-incorporating LDCT and/or digital histology images (e.g. H&E and multiplexed target protein stained) with genomic signatures that may include mutation (genes), transcriptional signatures (e.g. gene pathways) and/or epigenetic modifications (e.g. methylation profiles) as has been demonstrated to aid in the adjudicating Nodules with the Percept biomarker for presence of lung cancer[35]. Furthermore, the identification of biomarkers of progression will serve to identify molecular phenotype of a lesion to stratify who might be amenable and/or respond to an intervention that may not be tolerated by a healthy individual with and unknown likelihood of progression. This would enable the adoption of employing targeted intervention strategies that are likely to be developed to parallel targeted approaches for invasive lung cancer including mutation (EGFR, ALK, ROS1 etc.) and immune (PD-1/PDL-1, CTLA-4) alterations.

Identifying druggable targets. The rapid advancement and success of neoadjuvant immunotherapy (nivolumab – anti-PD1) in combination with chemotherapy (platinum doublet)[11,128], along with numerous ongoing trials for additional neoadjuvant strategies highlights the ability to eliminate disease earlier and improve outcomes following surgery[129]. However, to transition from neoadjuvant treatment to interception strategies, the initial challenge will be leveraging the discoveries discussed above to refine targets for intervention to prevent premalignant progression to invasive lung cancer. A first step in translating these discoveries to the identification of intervention targets that have low to moderate toxicity, is the narrowing of key drivers and mechanisms that are targetable. Current work points to the suppression of key anti-tumor immune pathways such as cytotoxic T cell recruitment opposed to T cell exhaustion found in late-stage disease[130] in addition to novel insights cell plasticity that results from genetic and epigenetic modifications[107]. However, to further narrow targets to move through the drug discovery pipeline, more mechanistic studies are needed as a second key step in this process. These can be accomplished via establishment of well-characterized preclinical models that accurately reflect premalignant human lesions that parallel the patient population of future intervention trials, to test targets and develop measurable end points. As highlighted a decade ago by Keith *et al*, historic lung cancer prevention studies relied on data from less aggressive pre-clinical LUAD models for interventions aimed at preventing the progression of bronchial dysplasia in humans[131], which although we propose lesions may develop through shared hallmarks differ in the molecular drivers that may explain the lack of success of agents trialed. Significant advances have been made in establishing and characterizing existing genetic, and carcinogen-induced murine models across LUSC and LUAD[25,132,133] presenting growing opportunities for preclinical studies. The field has also driven forward new approaches for introducing drive mutations to expand the number of genetic models to represent human disease in addition to the explosion of in vitro 3-D models of lung carcinogenesis that constitute differentiated and molecularly tunable systems[25]. However, careful characterization of precancerous lesions in many models remains limited[25,133]. With the increasing knowledge of alterations identified in early lesions, more work is essential to understand the interplay between genomic, epithelial, immune, and stromal phenotypes within lung compartments and respective models and their concordance with human disease fully elucidated to be best positioned for future pre-clinical intervention studies that will test agents that support anti-tumor and pro-wound healing and repair from known or unknow cellular injury.

Historic lung cancer prevention trials focused primarily on natural agents and vitamins thought to have limited toxicity, most had little to no efficacy whereas some were found to be toxic highlighting the need for new approaches[131]. More recent studies utilized prostacyclin analogues (e.g. iloprost) and nutrient modifying compounds (myoinositol and metformin) show some promise in early studies, although none have progressed to late stage clinical trials, responses in these early trials in a subset of patients provide an opportunity improve future trials via examining intermediate endpoints in non-invasive and lesional tissue to define the molecular changes that reflect regression or progression [134–136]. With a deeper understanding of immunology and biology of early lesions, there is a growing opportunity to reinvigorate the field including through the identification of possible immunoprevention approaches that target the immune suppressed premalignant microenvironment.

Further, insights into the targeted approaches will also support the approach to tuning dosing regimens and local delivery, which are moving to early phase studies. Early insights from on-going phase II studies include PD1 immune checkpoint inhibition (nivolumab) for the reversal of squamous dysplasia in high-risk current and former smokers with or without a history of lung cancer (NCT03347838)[137] and pembrolizumab in preventing lung cancer in patients with high-risk pulmonary nodules, the IMPRINT-Lung Study (NCT03634241)[138] suggest a path for immunoprevention and opportunities to refine dosing, however the both studies are small, and recruitment remains a challenge due to the absence of tools to better enroll patients. Studying intervention approaches in patients with high-risk pulmonary nodules that may be part solid and may histologically be premalignant is an innovative approach, yet the natural history of AAH and AIS lesions, which precede LUAD is not well understood so identifying those likely to progress to invasive disease no feasible. Thus, establishing the genomic, transcriptomic,



Figures 3. The Future: Finding and eradicating lung cancer before it becomes invasive. Biologic insights of precancerous transitions are positioned to drive the field toward the establishment of tools to stratifying lung cancer risk in all patients as current guidelines only screen a subset of the eligible population. The establishment of non-invasive molecular biomarkers will provide an opportunity to manage screening and identify those that would benefit from more additional screening or invasive diagnostics to identify those with progressive lesions that would benefit from intervention due to presence of progressive premalignant lesions. Intervention approaches will need to overcome challenges of toxicities but opportunities of to target early lesions locally and with approaches viable or not viable in the advanced disease setting may benefit from approaches like locally delivered therapeutics.

and proteomic signatures at the cellular and spatial levels will be crucial for assessing the success of such intervention approaches.

The growing adoption of lung cancer screening and the increased identification of GGOs and partially solid lesions that may represent progressive AAH and AIS or very early-stage LUADs underscores one of the most urgent unmet clinical needs in the field. The advent of electromagnetic navigation and robot-assisted bronchoscopy increases the ability to access and treat these early lesions less invasively than conventional surgical approaches and is poised to improve diagnosis and treatment. These new endoscopic approaches for tracking the natural history of GGOs via longitudinal sampling will enable identification of cellular and molecular markers associated with lesions destined

to progress to invasive adenocarcinoma and allow for the identification of prognostic biomarkers to stratify high-risk premalignant lesions into interception trials that have been limited due to ability to study the natural history of LUAD before invasion as the risk of sampling these nodules outweighed the benefit.. Beyond prognostic markers, multimodal profiling of these lesions will elucidate mechanisms by which epithelial and stromal cells interact with immune cells, providing novel approaches to harness the immune microenvironment for disease interception. Further, electromagnetic and robot-assisted bronchoscopies sampling GGOs in the lung periphery will establish foundational approaches for potential intra-lesional delivery of immune-activating therapeutics with an acceptable benefit-risk profile for otherwise healthy patients with lung pre-cancerous lesions. Local delivery of immunotherapies into premalignant lesions may reduce systemic toxicity and increase local drug concentration and efficacy, transforming the paradigm for lung cancer interception (**figure 3**).

Conclusions & Perspectives

The increased understanding of the molecular and cellular biology that enable normal bronchial or alveolar tissue to transform and evade cellular cues and immune detection, thereby colonizing the lung as a lesion that may become cancerous, will pave the way for novel lung cancer interception approaches. Further work in this space is needed to more fully characterize the spatial and temporal changes that are associated with early lung carcinogenesis for both squamous and adenocarcinoma. This will refine the vulnerabilities of the neoplastic changes that initiate carcinomas is being modeled in preclinical testing for targeted intervention strategies. Several key barriers remain to fully establish an approach for identifying patients harboring high-risk premalignant lesions that can be halted through effective intervention strategies. However, these barriers can be overcome with collaborative efforts from all stakeholders across academic, foundations, federal agencies and industry partners. These collaborative approaches include 1) longitudinal collection of lung tissue along the continuum of premalignancy to malignancy including in lung adenocarcinoma like done in on-going LUSC studies; 2) employment of multi-modal tissue profiling and innovative data integration strategies to identify novel therapeutic targets; 3) mechanistic testing and validation studies in relevant pre-clinical models, 4) development of therapeutic strategies with the appropriate benefit :risk ratio in an otherwise healthy population including procedural workflow and drug formulation for intralesional treatment, 5) establishing novel clinical trial designs suited for intervention where progression to invasive disease is variable, surrogate and intermediate molecular endpoints may improve efficacy and further reduce toxicity that will be less tolerable in the absence of invasive disease. Looking ahead, we anticipate that the use of data generated and widely disseminated coming from the large-scale number pre-cancer atlas efforts in lung and other organs within the scientific community by current and forthcoming will serve as the first step to usher in a new era for lung cancer prevention.

Supplemental Table 1. Key work in LUSC premalignant lesions

Supplemental Table 2. Key work in LUAD premalignant lesions

References:

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).
2. Cancer Facts & Figures 2023. (1930).
3. Zhang, Y. *et al.* Global variations in lung cancer incidence by histological subtype in 2020: a population-based study. *Lancet Oncol* **24**, 1206–1218 (2023).
4. Gazdar, A. F. & Brambilla, E. Preneoplasia of lung cancer. *Cancer Biomark* **9**, 385–396 (2010).
5. Howlader, N. *et al.* The Effect of Advances in Lung-Cancer Treatment on Population Mortality. *N Engl J Med* **383**, 640–649 (2020).
6. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).
7. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
8. de Koning, H. J. *et al.* Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N Engl J Med* **382**, 503–513 (2020).
9. **National Lung Screening Trial Research Team *et al.* Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 365, 395–409 (2011).**
10. Silvestri, G. A. *et al.* Outcomes From More Than 1 Million People Screened for Lung Cancer With Low-Dose CT Imaging. *CHEST* **164**, 241–251 (2023).
11. Forde, P. M. *et al.* Neoadjuvant Nivolumab plus Chemotherapy in Resectable Lung Cancer. *N Engl J Med* **386**, 1973–1985 (2022).
12. Altorki, N. *et al.* Lobar or Sublobar Resection for Peripheral Stage IA Non-Small-Cell Lung Cancer. *N Engl J Med* **388**, 489–498 (2023).
13. Kobayashi, Y., Ambrogio, C. & Mitsudomi, T. Ground-glass nodules of the lung in never-smokers and smokers: clinical and genetic insights. *Transl Lung Cancer Res* **7**, 487–497 (2018).
14. Migliore, M. *et al.* Ground glass opacities management in the lung cancer screening era. *Ann Transl Med* **6**, 90 (2018).

15. Jiang, J., Chang, S. H., Kent, A. J., Geraci, T. C. & Cerfolio, R. J. Current Novel Advances in Bronchoscopy. *Front Surg* **7**, 596925 (2020).

16. Tremblay, A. *et al.* Low Prevalence of High-Grade Lesions Detected With Autofluorescence Bronchoscopy in the Setting of Lung Cancer Screening in the Pan-Canadian Lung Cancer Screening Study. *Chest* **150**, 1015–1022 (2016).

17. Keith, R. L. *et al.* Lung cancer: Premalignant biology and medical prevention. *Semin Oncol* S0093-7754(22)00013–6 (2022) doi:10.1053/j.seminoncol.2022.02.001.

18. Shiels, M. S., Graubard, B. I., McNeel, T. S., Kahle, L. & Freedman, N. D. Trends in smoking-attributable and smoking-unrelated lung cancer death rates in the United States, 1991-2018. *J Natl Cancer Inst* **116**, 711–716 (2024).

19. Siegel, D. A., Fedewa, S. A., Henley, S. J., Pollack, L. A. & Jemal, A. Proportion of Never Smokers Among Men and Women With Lung Cancer in 7 US States. *JAMA Oncol* **7**, 302–304 (2021).

20. Corrales, L. *et al.* Lung cancer in never smokers: The role of different risk factors other than tobacco smoking. *Crit Rev Oncol Hematol* **148**, 102895 (2020).

21. **Merrick, D. T. *et al.* Persistence of Bronchial Dysplasia Is Associated with Development of Invasive Squamous Cell Carcinoma. *Cancer Prev Res (Phila)* 9, 96–104 (2016).**

22. Yang, D. *et al.* Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell* **185**, 1905-1923.e25 (2022).

23. Thege, F. I. *et al.* A Programmable In Vivo CRISPR Activation Model Elucidates the Oncogenic and Immunosuppressive Functions of MYC in Lung Adenocarcinoma. *Cancer Res* **82**, 2761–2776 (2022).

24. Blair, L. M. *et al.* Oncogenic context shapes the fitness landscape of tumor suppression. *Nat Commun* **14**, 6422 (2023).

25. Hynds, R. E. *et al.* Progress towards non-small-cell lung cancer models that represent clinical evolutionary trajectories. *Open Biol* **11**, 200247 (2021).

26. Curtius, K., Wright, N. A. & Graham, T. A. An evolutionary perspective on field cancerization. *Nat Rev Cancer* **18**, 19–32 (2018).

27. Slaughter, D. P., Southwick, H. W. & Smejkal, W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* **6**, 963–968 (1953).

28. Chang, Y.-L. *et al.* Clonality and prognostic implications of p53 and epidermal growth factor receptor somatic aberrations in multiple primary lung cancers. *Clin Cancer Res* **13**, 52–58 (2007).

29. **Kadara, H., Scheet, P., Wistuba, I. I. & Spira, A. E. Early Events in the Molecular Pathogenesis of Lung Cancer. *Cancer Prev Res (Phila)* 9, 518–527 (2016).**

30. Lochhead, P. *et al.* Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Mod Pathol* **28**, 14–29 (2015).

31. Gower, A. C., Steiling, K., Brothers, J. F., Lenburg, M. E. & Spira, A. Transcriptomic studies of the airway field of injury associated with smoking-related lung disease. *Proc Am Thorac Soc* 8, 173–179 (2011).

32. **Beane, J. *et al.* Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biol* 8, R201 (2007).**

33. Duclos, G. E. *et al.* Characterizing smoking-induced transcriptional heterogeneity in the human bronchial epithelium at single-cell resolution. *Sci Adv* **5**, eaaw3413 (2019).

34. AEGIS Study Team. Shared Gene Expression Alterations in Nasal and Bronchial Epithelium for Lung Cancer Detection. *J Natl Cancer Inst* **109**, djw327 (2017).

35. **Silvestri, G. A. *et al.* A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer. *N Engl J Med* 373, 243–251 (2015).**

36. Lee, H. J. *et al.* Impact of the Percepta Genomic Classifier on Clinical Management Decisions in a Multicenter Prospective Study. *Chest* **159**, 401–412 (2021).

37. Lamb, C. R. *et al.* A Nasal Swab Classifier to Evaluate the Probability of Lung Cancer in Patients with Pulmonary Nodules. *CHEST* S0012369223058282 (2023) doi:10.1016/j.chest.2023.11.036.

38. Beane, J. *et al.* Detecting the Presence and Progression of Premalignant Lung Lesions via Airway Gene Expression. *Clin Cancer Res* **23**, 5091–5100 (2017).

39. **Beane, J. E. *et al.* Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. *Nat Commun* 10, 1856 (2019).**

40. Colom, B. *et al.* Mutant clones in normal epithelium outcompete and eliminate emerging tumours. *Nature* **598**, 510–514 (2021).

41. Karlsson, K. *et al.* Deterministic evolution and stringent selection during preneoplasia. *Nature* **618**, 383–393 (2023).

42. Colom, B. *et al.* Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium. *Nat Genet* **52**, 604–614 (2020).

43. Kadara, H. *et al.* Driver Mutations in Normal Airway Epithelium Elucidate Spatiotemporal Resolution of Lung Cancer. *Am J Respir Crit Care Med* 200, 742–750 (2019).

44. Yoshida, K. *et al.* Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* 578, 266–272 (2020).

45. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150, 1107–1120 (2012).

46. Alexandrov, L. B. *et al.* Mutational signatures associated with tobacco smoking in human cancer. *Science* **354**, 618–622 (2016).

47. Kim, E. S. *et al.* Telomere length and recurrence risk after curative resection in patients with early-stage non-small-cell lung cancer: a prospective cohort study. *J Thorac Oncol* **10**, 302–308 (2015).

48. SEER*Stat Databases: November 2022 Submission. *SEER* https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/index.html.

49. Davis, J. D. & Wypych, T. P. Cellular and functional heterogeneity of the airway epithelium. *Mucosal Immunol* **14**, 978–990 (2021).

50. Vieira Braga, F. A. *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* **25**, 1153–1163 (2019).

51. Plasschaert, L. W. *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).

52. Montoro, D. T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 560, 319–324 (2018).

53. Lin, B. *et al.* Airway hillocks are injury-resistant reservoirs of unique plastic stem cells. *Nature* 629, 869–877 (2024).

54. Crystal, R. G., Randell, S. H., Engelhardt, J. F., Voynow, J. & Sunday, M. E. Airway epithelial cells: current concepts and challenges. *Proc Am Thorac Soc* **5**, 772–777 (2008).

55. WD, T., E, B., HK, M.-H. & CC, H. *Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart*.

56. Wistuba, I. I. & Gazdar, A. F. Lung cancer preneoplasia. *Annu Rev Pathol* 1, 331–348 (2006).

57. Ruiz García, S. *et al.* Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. *Development* **146**, dev177428 (2019).

58. Nicholson, A. G. *et al.* The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. *J Thorac Oncol* **17**, 362–387 (2022).

59. WD, T., E, B., AP, B., A, M. & AG, N. *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart*.

60. Board, W. C. of T. E. *Thoracic Tumours*.

61. Brambilla, E., Travis, W. D., Colby, T. V., Corrin, B. & Shimosato, Y. The new World Health Organization classification of lung tumours. *Eur Respir J* **18**, 1059–1068 (2001).

62. Beasley, M. B., Brambilla, E. & Travis, W. D. The 2004 World Health Organization classification of lung tumors. *Semin Roentgenol* **40**, 90–97 (2005).

63. Franklin, W. A. *et al.* Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis. *J Clin Invest* **100**, 2133–2137 (1997).

64. Wistuba, I. I. *et al.* Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene* **18**, 643–650 (1999).

65. Wistuba, I. I. *et al.* Allelic losses at chromosome 8p21-23 are early and frequent events in the pathogenesis of lung cancer. *Cancer Res* **59**, 1973–1979 (1999).

66. Wistuba, I. I. *et al.* High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. *Cancer Res* **60**, 1949–1960 (2000).

67. Massion, P. P. *et al.* Recurrent genomic gains in preinvasive lesions as a biomarker of risk for lung cancer. *PLoS One* **4**, e5611 (2009).

68. McCaughan, F. *et al.* Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. *Am J Respir Crit Care Med* **182**, 83–91 (2010).

69. Pipinikas, C. P. *et al.* Cell migration leads to spatially distinct but clonally related airway cancer precursors. *Thorax* **69**, 548–557 (2014).

70. Nakachi, I. *et al.* Application of SNP microarrays to the genome-wide analysis of chromosomal instability in premalignant airway lesions. *Cancer Prev Res (Phila)* **7**, 255–265 (2014).

71. van Boerdonk, R. A. A. *et al.* Close Surveillance with Long-Term Follow-up of Subjects with Preinvasive Endobronchial Lesions. *Am J Respir Crit Care Med* 192, 1483–1489 (2015).

72. Teixeira, V. H. *et al.* Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat Med* 25, 517–525 (2019).

73. Belinsky, S. A. *et al.* Gene promoter methylation in plasma and sputum increases with lung cancer risk. *Clin Cancer Res* **11**, 6505–6511 (2005).

74. Belinsky, S. A. *et al.* Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Res* **62**, 2370–2377 (2002).

75. Merrick, D. T. *et al.* Altered Cell-Cycle Control, Inflammation, and Adhesion in High-Risk Persistent Bronchial Dysplasia. *Cancer Res* 78, 4971–4983 (2018).

76. Ning, B. *et al.* Convergence of YAP/TAZ, TEAD and TP63 activity is associated with bronchial premalignant severity and progression. *J Exp Clin Cancer Res* **42**, 116 (2023).

77. Krishna, C. *et al.* An immunogenetic basis for lung cancer risk. *Science* **383**, eadi3808 (2024).

78. Pennycuick, A. *et al.* Immune Surveillance in Clinical Regression of Preinvasive Squamous Cell Lung Cancer. *Cancer Discov* 10, 1489–1499 (2020).

79. Mascaux, C. *et al.* Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature* 571, 570–575 (2019).

80. Bons, J. *et al.* Data-independent acquisition and quantification of extracellular matrix from human lung in chronic inflammation-associated carcinomas. *Proteomics* **23**, 2200021 (2023).

81. Strasser, M. K. *et al.* Concerted epithelial and stromal changes during progression of Barrett's Esophagus to invasive adenocarcinoma exposed by multi-scale, multi-omics analysis. Preprint at https://doi.org/10.1101/2023.06.08.544265 (2023).

82. Cho, J. *et al.* Proportion and clinical features of never-smokers with non-small cell lung cancer. *Chin J Cancer* **36**, 20 (2017).

83. Travis, W. D. Pathology of lung cancer. *Clin Chest Med* **32**, 669–692 (2011).

84. Tian, S. Classification and survival prediction for early-stage lung adenocarcinoma and squamous cell carcinoma patients. *Oncol Lett* **14**, 5464–5470 (2017).

85. Mori, M., Rao, S. K., Popper, H. H., Cagle, P. T. & Fraire, A. E. Atypical adenomatous hyperplasia of the lung: a probable forerunner in the development of adenocarcinoma of the lung. *Mod Pathol* **14**, 72–84 (2001).

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

87. Sainz de Aja, J., Dost, A. F. M. & Kim, C. F. Alveolar progenitor cells and the origin of lung cancer. *J Intern Med* **289**, 629–635 (2021).

88. Iwata, H. Adenocarcinoma containing lepidic growth. *J Thorac Dis* **8**, E1050–E1052 (2016).

89. Dong, Z.-Y. *et al.* Genetic and Immune Profiles of Solid Predominant Lung Adenocarcinoma Reveal Potential Immunotherapeutic Strategies. *J Thorac Oncol* **13**, 85–96 (2018).

90. Takamochi, K. *et al.* Loss of heterozygosity on chromosomes 9q and 16p in atypical adenomatous hyperplasia concomitant with adenocarcinoma of the lung. *Am J Pathol* **159**, 1941–1948 (2001).

91. Kitaguchi, S., Takeshima, Y., Nishisaka, T. & Inai, K. Proliferative activity, p53 expression and loss of heterozygosity on 3p, 9p and 17p in atypical adenomatous hyperplasia of the lung. *Hiroshima J Med Sci* **47**, 17–25 (1998).

92. Krysan, K. *et al.* The Immune Contexture Associates with the Genomic Landscape in Lung Adenomatous Premalignancy. *Cancer Res* 79, 5022–5033 (2019).

93. Chen, H. *et al.* Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nat Commun* 10, 5472 (2019).

94. Sivakumar, S. *et al.* Genomic landscape of allelic imbalance in premalignant atypical adenomatous hyperplasias of the lung. *EBioMedicine* 42, 296–303 (2019).

95. Sivakumar, S. *et al.* Genomic Landscape of Atypical Adenomatous Hyperplasia Reveals Divergent Modes to Lung Adenocarcinoma. *Cancer Res* 77, 6119–6130 (2017).

96. Cascetta, P. *et al.* KRAS in NSCLC: State of the Art and Future Perspectives. *Cancers (Basel)* **14**, 5430 (2022).

97. Dankort, D. *et al.* A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. *Genes Dev* **21**, 379–384 (2007).

98. Qian, J. *et al.* Genomic Underpinnings of Tumor Behavior in In Situ and Early Lung Adenocarcinoma. *Am J Respir Crit Care Med* **201**, 697–706 (2020).

99. Hu, X. *et al.* Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. *Nat Commun* **10**, 2978 (2019).

100. Licchesi, J. D. F., Westra, W. H., Hooker, C. M. & Herman, J. G. Promoter hypermethylation of hallmark cancer genes in atypical adenomatous hyperplasia of the lung. *Clin Cancer Res* **14**, 2570–2578 (2008).

101. Selamat, S. A. *et al.* DNA methylation changes in atypical adenomatous hyperplasia, adenocarcinoma in situ, and lung adenocarcinoma. *PLoS One* **6**, e21443 (2011).

102. Sinjab, A. *et al.* Resolving the Spatial and Cellular Architecture of Lung Adenocarcinoma by Multiregion Single-Cell Sequencing. *Cancer Discov* 11, 2506–2523 (2021).

103. Han, G. *et al.* An atlas of epithelial cell states and plasticity in lung adenocarcinoma. 2022.05.13.491635 Preprint at https://doi.org/10.1101/2022.05.13.491635 (2022).

104. Wang, Z. *et al.* Deciphering cell lineage specification of human lung adenocarcinoma with single-cell RNA sequencing. *Nat Commun* **12**, 6500 (2021).

105. Yanagawa, J. *et al.* Single-Cell Characterization of Pulmonary Nodules Implicates Suppression of Immunosurveillance across Early Stages of Lung Adenocarcinoma. *Cancer Res* **83**, 3305–3319 (2023).

106. Yang, D. *et al.* Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell* **185**, 1905-1923.e25 (2022).

107. Han, G. *et al.* An atlas of epithelial cell states and plasticity in lung adenocarcinoma. *Nature* 627, 656–663 (2024).

108. Li, Z. *et al.* Alveolar Differentiation Drives Resistance to KRAS Inhibition in Lung Adenocarcinoma. *Cancer Discov* **14**, 308–325 (2024).

109. X, H. *et al.* Evolution of DNA methylome from precancerous lesions to invasive lung adenocarcinomas. *Nature communications* **12**, (2021).

110. Dejima, H. *et al.* Immune evolution from preneoplasia to invasive lung adenocarcinomas and underlying molecular features. *Nat Commun* **12**, 2722 (2021).

111. Lavin, Y. *et al.* Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell* **169**, 750-765.e17 (2017).

112. Altorki, N. K. *et al.* Global evolution of the tumor microenvironment associated with progression from preinvasive invasive to invasive human lung adenocarcinoma. *Cell Rep* **39**, 110639 (2022).

113. Hao, D. *et al.* The Single-Cell Immunogenomic Landscape of B and Plasma Cells in Early-Stage Lung Adenocarcinoma. *Cancer Discov* 12, 2626–2645 (2022).

114. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31–46 (2022).

115. Tammemägi, M. C. *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* **368**, 728–736 (2013).

116. Wood, D. E. *et al.* NCCN Guidelines® Insights: Lung Cancer Screening, Version 1.2022. *J Natl Compr Canc Netw* **20**, 754–764 (2022).

117. Brock, B. A. *et al.* Social and Biological Determinants in Lung Cancer Disparity. *Cancers (Basel)* **16**, 612 (2024).

118. Tammemägi, M. C. *et al.* Evaluation of the lung cancer risks at which to screen ever- and never-smokers: screening rules applied to the PLCO and NLST cohorts. *PLoS Med* **11**, e1001764 (2014).

119. Aldrich, M. C. *et al.* Evaluation of USPSTF Lung Cancer Screening Guidelines Among African American Adult Smokers. *JAMA Oncol* **5**, 1318–1324 (2019).

120. Tammemagi, M. *et al.* Predicting Malignancy Risk of Screen-Detected Lung Nodules-Mean Diameter or Volume. *J Thorac Oncol* **14**, 203–211 (2019).

121. Mikhael, P. G. *et al.* Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk From a Single Low-Dose Chest Computed Tomography. *J Clin Oncol* **41**, 2191–2200 (2023).

122. Kammer, M. N. *et al.* Integrated Biomarkers for the Management of Indeterminate Pulmonary Nodules. *Am J Respir Crit Care Med* **204**, 1306–1316 (2021).

123. Kammer, M. N. & Massion, P. P. Noninvasive biomarkers for lung cancer diagnosis, where do we stand? *J Thorac Dis* **12**, 3317–3330 (2020).

124. Kalinke, L., Thakrar, R. & Janes, S. M. The promises and challenges of early non-small cell lung cancer detection: patient perceptions, low-dose CT screening, bronchoscopy and biomarkers. *Mol Oncol* **15**, 2544–2564 (2021).

125. Mathios, D. *et al.* Detection and characterization of lung cancer using cell-free DNA fragmentomes. *Nat Commun* **12**, 5060 (2021).

126. Hubers, A. J., Prinsen, C. F. M., Sozzi, G., Witte, B. I. & Thunnissen, E. Molecular sputum analysis for the diagnosis of lung cancer. *Br J Cancer* **109**, 530–537 (2013).

127. Spira, A. *et al.* Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* **13**, 361–366 (2007).

128. Leal, T. A. & Ramalingam, S. S. Neoadjuvant therapy gains FDA approval in non-small cell lung cancer. *Cell Rep Med* **3**, 100691 (2022).

129. Heiden, B. T. *et al.* Association Between Surgical Quality Metric Adherence and Overall Survival Among US Veterans With Early-Stage Non-Small Cell Lung Cancer. *JAMA Surg* **158**, 293–301 (2023).

130. Wu, Y. *et al.* The identification of genes associated T-cell exhaustion and construction of prognostic signature to predict immunotherapy response in lung adenocarcinoma. *Sci Rep* **13**, 13415 (2023).

131. Keith, R. L. & Miller, Y. E. Lung cancer chemoprevention: current status and future prospects. *Nat Rev Clin Oncol* **10**, 334–343 (2013).

132. Marjanovic, N. D. *et al.* Emergence of a High-Plasticity Cell State during Lung Cancer Evolution. *Cancer Cell* **38**, 229-246.e13 (2020).

133. Gómez-López, S., Whiteman, Z. E. & Janes, S. M. Mapping lung squamous cell carcinoma pathogenesis through in vitro and in vivo models. *Commun Biol* **4**, 937 (2021).

134. Keith, R. L. *et al.* Oral iloprost improves endobronchial dysplasia in former smokers. *Cancer Prev Res (Phila)* **4**, 793–802 (2011).

135. Lam, S. *et al.* A Randomized Phase IIb Trial of myo-Inositol in Smokers with Bronchial Dysplasia. *Cancer Prev Res (Phila)* **9**, 906–914 (2016).

136. Roos, J. F. *et al.* Metformin for lung cancer prevention and improved survival: a novel approach. *Eur J Cancer Prev* **28**, 311–315 (2019).

137. Study Details | Nivolumab for the Reversal of Squamous Dysplasia in High Risk Current and Former Smokers | ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT03347838.

138. Study Details | Pembrolizumab in Preventing Lung Cancer in Patients With Stage I-II Non-Small Cell Lung Cancer or High-Risk Pulmonary Nodules, the IMPRINT-Lung Study | ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT03634241.

139. Auerbach, O., Stout, A. P., Hammond, E. C. & Garfinkel, L. Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. *N Engl J Med* **265**, 253–267 (1961).

Funding
Investigators supported by in part by: the National Cancer Institute (NCI) Human Tumor Atlas Network (HTAN) (U2CCA233238-01) (AS, SD, SM, HK, SJ); Stand Up To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team Translational Cancer Research Grant (grant number: SU2C-AACR-DT23-17 to Steven M. Dubinett and Avrum E. Spira). Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C.

Competing Interests
AS is and employee of Johnson and Johnson Innovative Medicine. SM, SD and HK have sponsor research agreements with Johnson and Johnson supporting research in the area.

Author Contribution
SM, RZ, MR AS made substantial contributions the research and writing of the article with SJ, HK, SD providing guidance, discussion of the content, reviewed and edits to the manuscript before submission.

Glossary [G]:

Precancerous, premalignant, preneoplastic and preinvasive lesions: histopathologic changes to the normal tissue architecture driven by genomic, transcriptional, epigenetic changes that suggest progression toward carcinoma, but have the capacity to naturally resolved.

Low dose computed tomography (LDCT): medical imaging technique that uses a low dose x-ray to image the chest employed for lung cancer screening.

Ground glass opacities (GGOs): light hazy areas of increased opacity that appear on LDCT

Bronchoscopy: a clinical endoscopic procedure that visualizes the airways.

Intervention: a treatment that acts to prevent or treat disease

Primary and secondary prevention: In cancer, primary prevention is the prevention of cancer from every developing & secondary prevention is the prevention of the progression of early disease to aggressive lethal disease.

Field of cancerization: The occurrence of molecular alterations in normal appearing tissue in proximity to a tumor that harbors phenotypic cancer alterations.

Field of injury: The presence of molecular alterations throughout normal appearing tissue as a consequence of exposure.