TP53-Mutated Preneoplastic Colonies, Adaptive Immune Response, and Risk for Lung Cancer and COPD

James C. Willey[1], Andrew R. Boring[1], Erin L. Crawford[1], Kevin N. Lei[1], Thomas M. Blomquist[1], Christopher I. Amos[2], Stephen A. Deppen[3], Eric L. Grogan[3], Heidi Chen[3], Mohamed Omballi[4], and Daniel J. Craig[1]

ABSTRACT

The purpose of this study was to better characterize the prevalence and size distributions of *TP53*-mutated preneoplastic colonies in relation to the risk for lung cancer and chronic obstructive pulmonary disease (COPD). Ultrasensitive next-generation sequencing was used to assess colonies comprising as few as 100 cells in airway epithelium. Specimens from 59 subjects with a heavy smoking history were assessed in four cohorts: lung cancer with ($n = 14$) or without COPD ($n = 16$) and non–lung cancer with ($n = 7$) or without COPD ($n = 22$). *TP53*-mutated preneoplastic colony prevalence was higher in lung cancer subjects with or without COPD and in COPD subjects without lung cancer ($P < 0.01$). A higher prevalence of colonies driven by the subset of *TP53* mutations known to be neoantigenic in cancer tissue was nominally significant ($P < 0.05$). *TP53*-mutated colony distributions skewed to larger sizes among lung cancer subjects and skewed greater among subjects with both COPD and active cigarette smoking ($P < 0.01$), each of which is associated with adaptive immune inflammation. Thus, TP53-mutated colony prevalence and size distribution measured in nonmalignant airway epithelium improved characterization of lung cancer and COPD risk. Further, these colonies may stimulate the adaptive immune response and inflammation that is mechanistically associated with COPD. A skew to larger colonies is consistent with selection for immune surveillance avoidance characteristics and may provide a metric for both lung cancer risk and immune surveillance function. Among healthier non-COPD smokers who are ineligible for lung cancer screening based on demographic criteria, a positive *TP53* biomarker test may identify lung cancer risk sufficiently high to warrant screening eligibility.

Prevention Relevance: *TP53*-mutated colony prevalence and size distribution measured in nonmalignant airway epithelium improved characterization of lung cancer and COPD risk. The observed skew to larger colonies in lung cancer subjects is consistent with selection for immune surveillance avoidance and may enable measurement of immune surveillance function to guide the development of prevention therapeutics.

Introduction

Annual low-dose CT (LDCT) screening reduces mortality from lung cancer among high-risk patients based on smoking history, age, and chronic obstructive pulmonary disease (COPD; refs. 1–3). However, a substantial fraction of lung cancer occurs in individuals who do not meet the eligibility criteria for LDCT screening (2, 4). Moreover, age, smoking, and COPD are associated with adverse effects on screening benefits due to a corresponding increase in inoperability or competing causes of death (5–7). Thus, there is a need to develop biomarkers to determine the lung cancer risk among healthy individuals.

Cigarette smoke inhalation induces genetic changes throughout the lungs, often referred to as a "field effect" (8–10). Quantification of somatic genetic changes in lung tissue is a successful strategy for identifying new biomarkers of lung cancer risk. Recently, we used ultrasensitive next-generation sequencing (US-NGS) to test for very low variant allele fraction (VAF) mutations in the TP53 gene, targeting the DNA-binding region spanned by exons 5, 6, and 7 (11). This analysis revealed a high prevalence of microscopic *TP53*-mutated airway epithelial cell (AEC) colonies in normal-appearing airways of heavy smokers, an observation that was corroborated using orthogonal methods (12, 13). Moreover, the prevalence of *TP53*-mutated colonies is significantly higher in the lungs of smokers with lung cancer and represents a promising biomarker for the detection of lung cancer risk and/or early diagnosis of lung cancer (11, 13). Further, the lung cancer risk associated with TP53-mutated colony prevalence is independent of the risk associated with smoking pack-years or age (13). Notably, the prevalence of *TP53*-mutated colonies was 10-fold higher than the prevalence of

[1]The University of Toledo College of Medicine, Toledo, Ohio. [2]University of New Mexico, Albuquerque, New Mexico. [3]Vanderbilt University Medical Center, Nashville, Tennessee. [4]Promedica, Toledo, Ohio.

Corresponding Author: James C. Willey, The University of Toledo Health Sciences Campus, Block Health Sciences Building, Room 408, 3000 Arlington Avenue, Toledo, OH 43614. E-mail: james.willey2@utoledo.edu

Cancer Prev Res 2026;19:631–43

doi: 10.1158/1940-6207.CAPR-26-0112

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2026 The Authors; Published by the American Association for Cancer Research

colonies driven by mutated *PIK3CA* or *BRAF* oncogenes. No colonies driven by oncogenic *KRAS* or *EGFR* were observed.

These findings highlight the need to better understand inter-individual variation in interactions between microscopic preneoplastic *TP53*-mutated colonies and the lung tissue microenvironment, including the immune response. The immune system plays an important role in both preventing and genetically modifying tumors, and the risk of lung cancer is increased in immunosuppressed individuals, as well as in mice with defects in genes responsible for an adaptive immune T-cell response (14). Recent studies have advanced our knowledge in this area by characterizing the immune response to macroscopic precancerous lesions in the airways (15). There remains a need to investigate the interactions between the microenvironment and even earlier stages of cancer development (16).

The application of recently available US-NGS methods is a promising approach for addressing this need (11, 13, 17). With these methods it possible to assess the prevalence and size distributions of microscopic *TP53*-mutated preneoplastic colonies in the airways of subjects with conditions known to alter the level of adaptive immune inflammation, including COPD and active cigarette smoking.

The early stages of COPD are characterized by inflammation associated with adaptive immune responses (18–20). However, the stimulus for this response is not yet well understood. One possible stimulus is neoantigens. Neoantigens may be produced by somatic mutations that express aberrant proteins not accounted for in development when negative selection limits autoimmunity by quenching T-cell reactivity against the germline proteome. We hypothesized that a key driver of the adaptive immune inflammation in COPD lung tissue is neoantigens associated with the many microscopic TP53-mutated preneoplastic colonies induced in lung tissue by inhaled cigarette smoke (13). According to the hypothesis, once these *TP53*-mutated colonies reach a threshold size, neoantigenic properties will stimulate an adaptive immune response, and they will be eliminated (14). However, in some individuals the adaptive immune response will damage adjacent normal lung structures, resulting in COPD. Alternatively, the intense immune response to the continuously generated large number of *TP53* colonies may result in the selection of colonies with immune surveillance avoidance characteristics, enabling further colony expansion and additional somatic mutations resulting in lung cancer. Notably, approximately 50% of non–small cell lung cancer (NSCLC) tissues have a *TP53* driver mutation (21, 22). These cancers also have an oncogene comutation, most commonly *KRAS* or *EGFR*. The most common NSCLC *TP53* driver mutations express TP53 proteins that are documented to be neoantigenic in cancer tissues (23) and are associated with both activation of a $CD8^+$ T-cell response and PD-L1 upregulation (24) within lung cancer tissue. Thus, based on the higher prevalence of *TP53*-mutated colonies in the airway epithelium of patients with lung cancer compared with those driven by activated oncogenes, it is reasonable to hypothesize that each *TP53* mutation–driven lung cancer originates from a particular *TP53*-mutated cell in the lung epithelium, consistent with the Boveri hypothesis (25). According to this hypothesis, this progenitor gives rise to a colony of cells; certain cells within the colony acquire the capacity to escape immune surveillance, in some cases through upregulation of PD-L1 expression (26), and additional somatic mutations occur in cells within these colonies, including *KRAS* or *EGFR* oncogene activation, resulting in lung cancer.

To better understand the interactions between early preneoplastic lesions and the immune microenvironment, we tested whether the prevalence of microscopic *TP53*-mutated colonies in the airway epithelium is associated with lung cancer risk after controlling for the diagnosis of COPD and/or with COPD risk after controlling for the diagnosis of lung cancer. In addition, because COPD and cigarette smoking status are independently associated with a T cell–driven immune inflammatory response in lung tissue (18–20), we assessed differences in *TP53*-mutated colony size distributions based on the presence or absence of lung cancer and/or COPD and current or former cigarette smoking status.

Materials and Methods

Subject enrollment and specimen collection

The cohort of 59 subjects used in this study has been described previously, and only subjects from the previous publication were included in the present study (13). Briefly, subjects undergoing bronchoscopy for clinical purposes provided informed consent and were enrolled according to one of the two Institutional Review Board (IRB)–approved research protocols at the Vanderbilt University Medical Center (VUMC), 0398 (Molecular Predictors of Lung Cancer Behavior) or 1078 (Nashville Lung Screening Trial). The studies were conducted in accordance with the Belmont Report and the US Common Rule. Reanalysis of deidentified data was conducted under the University of Toledo IRB exemption # 301971. Determination of lung cancer risk by demographic characteristics for each subject was based on the calculation of the $PLCO_{m2012}$ lung cancer risk score (3). COPD status was determined by results of the forced expiratory volume in one second (FEV1)/forced vital capacity (FVC) pulmonary function test results, specifically FEV1/FVC <0.7 = COPD. Demographic characteristics, including age, sex, smoking history, and disease status of the subjects in this cohort, are provided in Supplementary Table S1.

Sample processing

Methods for subject enrollment, specimen collection, and *TP53* mutation prevalence analyses were performed as described previously (13). AEC specimens were collected by bronchoscopic brush biopsy of macroscopically normal airways (i.e., not overtly metaplastic or abnormal appearing to the trained pulmonologist) from 30 lung cancer cases, all of whom had a smoking history, and 29 noncancer controls. Brushings were taken from the opposite lung or a region

distant from a known or suspected disease, immediately frozen on dry ice, and stored at −80°C until DNA extraction. Genomic DNA (gDNA) was extracted from each AEC specimen using a DNeasy Blood & Tissue Kit (QIAGEN). Aliquots of gDNA extracted from AEC specimens were deidentified and blinded by VUMC and provided to the University of Toledo.

Targeted US-NGS method

Targeted US-NGS was used to measure very low VAF mutations in gDNA extracted from each AEC specimen, as described previously (13, 17). Sequencing data for the cohort are available through the National Library of Medicine National Center for Biotechnology Information under Sequence Read Archive (SRA) number PRJNA1001394, https://www.ncbi.nlm.nih.gov/sra/PRJNA1001394. Three *TP53* exons spanning 193 bp were targeted (11, 13, 17). Target-specific synthetic DNA internal standards (IS) were spiked into each specimen prior to NGS library preparation. Each reported *TP53* mutation VAF was validated as significant relative to a nucleotide-specific limit of detection (LOD) based on the IS VAF value, as described previously (17). This step controlled for nucleotide substitution errors associated with library preparation and sequencing, thereby enabling the measurement of variants with a VAF as low as 0.01%.

Measurement of TP53-mutated colony size relative to colony prevalence

TP53-mutated colony size calculation

The colony size was calculated based on the VAF and the number of cells sampled using a bronchoscopy brush (average of 1 million cells/subject). As is standard practice, we assumed that each variant measured was attributable to a single heterozygous cell for the variant (27, 28). Thus, a variant with a VAF of 0.01% represented 1 cell in 10,000 cells. Because an average of 10^6 total cells was sampled in each brush specimen, a VAF of 0.01% represented 100 cells/1,000,000 cells sampled.

TP53-mutated colony prevalence calculation

Cohort colony prevalence was calculated as the mean number of colonies of any size (i.e., cell number) per 10^6 cells sampled from the subjects within a cohort.

TP53-mutated colony cumulative frequency distributions

To assess the relationship between *TP53*-mutated colony size and cumulative colony prevalence, each of the 192 variants identified among the 59 subjects was binned according to a colony size VAF interval of 0.000001 (i.e., a colony size interval of one cell/10^6 cells) and presented as the independent variable. Cohort-specific cumulative colony prevalence was the dependent variable and was calculated as the number of colonies/10^6 cells greater than the colony size threshold (T) normalized by dividing by the number of patients in each cohort, resulting in cumulative colonies/10^6 cells/subject.

$$y(T) = \frac{[\#Colonies\ with\ VAF \geq T/10E6\ cells]/Cohort}{Subjects/Cohort}$$

Statistical analysis

Prediction of lung cancer by cancer and COPD status

The four cohorts included in the analysis were lung cancer–non-COPD (LC-NCO), lung cancer–COPD (LC-CO), non-lung cancer–COPD (NLC-CO), and non-lung cancer–non-COPD (NLC-NCO). The receiver operator characteristic (ROC) area under the curve (AUC) and Kruskal–Wallis test were used to test the performance of the *TP53* mutation prevalence biomarker to classify lung cancer or COPD among the four cohorts.

Test for association with lung cancer after controlling for COPD: The LC-NCO cohort ($n = 16$) was compared with NLC-NCO the cohort ($n = 22$), and the LC-CO cohort ($n = 14$) was compared with the NLC-CO cohort ($n = 7$).

Test for association with COPD after controlling for lung cancer: All COPD subjects (LC-CO and NLC-CO; $n = 21$) were compared with all NCO subjects (LC-NCO and NLC-NCO; $n = 38$); NLC-CO subjects ($n = 7$) were compared with NLC-NCO subjects ($n = 21$); and LC-CO subjects ($n = 14$) were compared with LC-NCO subjects ($n = 16$).

Analysis of TP53-mutated colony size distributions

For each of the four cohorts (LC-NCO, LC-CO, NLC-CO, and NLC-NCO), the cumulative number of *TP53*-mutated colonies for each colony size was plotted. Cox proportional hazards analysis was used to test for inter-cohort differences in cumulative colony survival relative to colony size, starting with a colony size LOD of 100 cells. We first $\log_2$-transformed colony size (cells/colony) values. At the LOD [min ($\log_2$cells/colony], the colony size was $\log_2(100\ \text{cells/colony}) = 2^{6.645}$ cells/colony. This value was then subtracted from all higher $\log_2$ cells/colony values so that the starting independent value was zero. In each plot, the dependent values were the cumulative colonies/10^6 cells/subject for each colony size value.

Results

Summary statistics are provided in Supplementary Table S1. In this table, the *TP53*-mutated colony prevalence measurement units are presented as colonies/193 bp/10^6 cells/subject. Because the same 193 bp region was assessed in each measurement, the units were converted to colonies/10^6 cells/subject for presentation in the figures. This unit conversion did not alter the results of statistical analysis. The NLC-NCO, NLC-CO, LC-NCO, and LC-CO cohorts were generally well matched with respect to age and smoking. However, as indicated in Supplementary Table S1.4, there were eight never-smokers among the NLC-NCO group. The remaining 14 LC-NCO subjects were heavy smokers (i.e., >20 pack-years of smoking history) with a mean of 42 pack-years of smoking

history, and this was not significantly different from a mean of 59 pack-years for the 16 NLC-CO subjects. Importantly, (i) the *TP53*-mutated colony prevalence was very low in all NLC-NCO subjects, (ii) there was not a significant difference in colony prevalence between never-smokers (0.5 colonies/10^6 cells/subject) and smokers (0.8 colonies/10^6 cells/subject) in the NLC-NCO cohort, and (iii) removing the never-smokers from the NLC-NCO cohort did not alter the results of comparisons between NLC-NCO and NLC-CO or LC-NCO.

Performance of the *TP53* biomarker to identify lung cancer after controlling for COPD

NCO subjects: comparison of NLC-NCO with LC-NCO

TP53-mutated colony prevalence (colonies/10^6 cells/subject) classified LC and NLC among 38 NCO subjects ($P < 0.001$, Kruskal–Wallis). When fit to an ROC curve, the AUC for identification of lung cancer by the *TP53* biomarker in this NCO group was 0.84 (**Fig. 1A**). At the optimal cutoff point for accuracy based on the Youden index (lung cancer > 4 *TP53* mutations/bp), the *TP53* biomarker displayed 85% specificity, 69% sensitivity, 79% positive predictive value, and 85% negative predictive value. We also assessed the performance of the demographic risk index $PLCO_{m2012}$ in the NCO group alone (**Fig. 1B**) and in combination (**Fig. 1C**) with the *TP53* biomarker. $PLCO_{m2012}$ alone had an AUC of 0.87 (**Fig. 1B**), and a biomarker that combined $PLCO_{m2012}$ with the *TP53* biomarker had an AUC of 0.94 (**Fig. 1C**).

COPD Subjects (NLC-CO vs. LC-CO)

The *TP53*-mutated colony prevalence classified lung cancer and NLC in 21 COPD subjects ($P = 0.036$, Kruskal–Wallis), and the ROC AUC was 0.86 (**Fig. 1D**), with an optimal cutoff point of COPD >5 *TP53* mutations/bp. Among the COPD subjects, the $PLCO_{m2012}$ index alone had an AUC of 0.81 (**Fig. 1E**), and when combined with the TP53 biomarker, the AUC increased to 0.95 (**Fig. 1F**).

Performance of the TP53 mutation biomarker to identify subjects with COPD after controlling for lung cancer

NLC subjects (NLC-NCO vs. NLC-CO)

Among the 29 NLC subjects, classification of COPD versus NCO according to the *TP53*-mutated colony prevalence biomarker was significant ($P < 0.001$, Kruskal–Wallis).

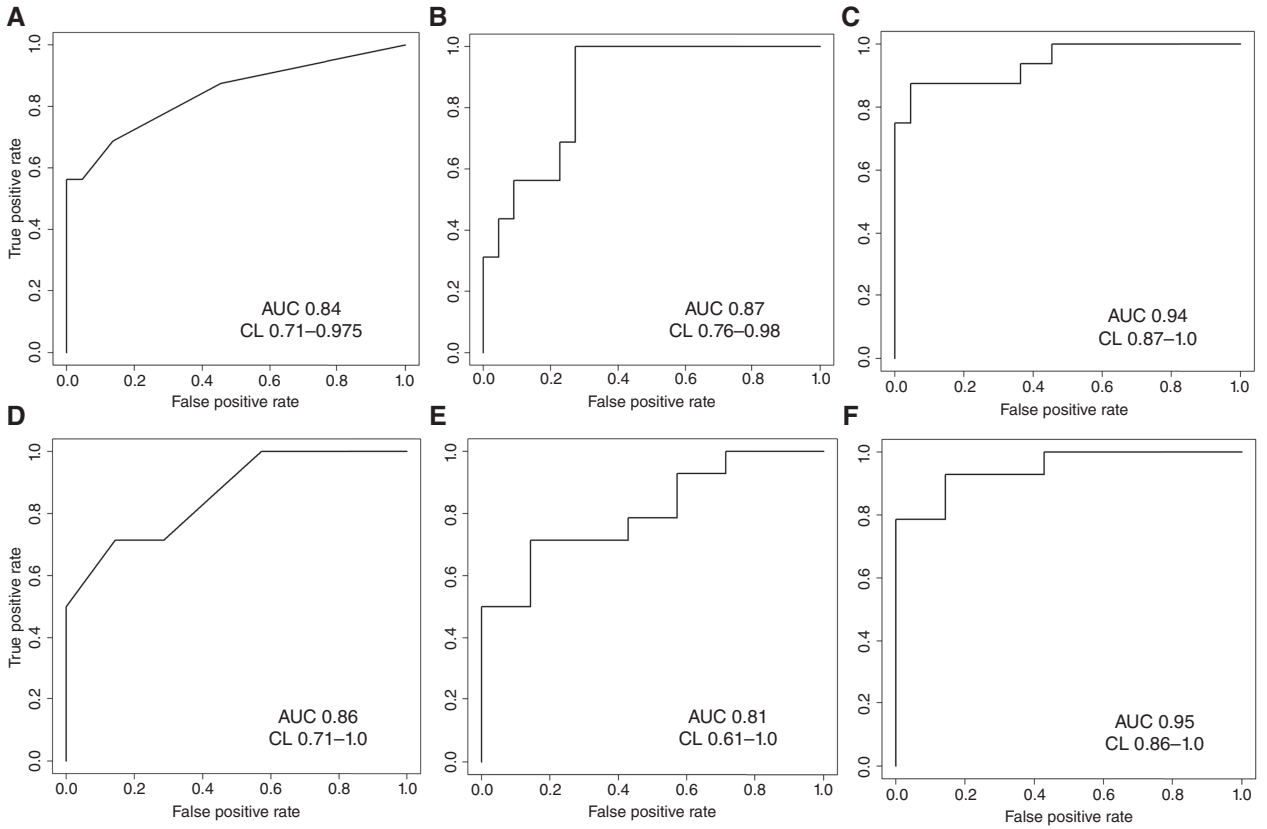


Figure 1.

TP53 biomarker classification for lung cancer status after controlling for COPD. Classification of subjects according to lung cancer status based on *TP53*-mutated colony prevalence (mutations/10^6 cells/subject). ROC AUC and 95% confidence limits (CL) are presented on each graph. **A–C,** NCO subjects ($n = 38$). **A,** TP53 biomarker alone; (**B**) $PLCO_{m2012}$ alone; and (**C**) combined biomarker. **D–F,** COPD subjects ($n = 21$). **D,** *TP53* biomarker alone; (**E**) $PLCO_{m2012}$ alone; and (**F**) combined biomarker.

The ROC AUC was 0.84, with an optimal cutoff point of >4 *TP53* mutations/bp (**Fig. 2**).

Lung cancer subjects (LC-NCO vs. LC-CO)

Among lung cancer subjects, this biomarker did not classify according to COPD status ($P > 0.05$), with an AUC of 0.52.

Effect of lung cancer, COPD, and smoking status on colony prevalence and size distributions

The *TP53*-mutated colony size distribution plots for each of the four cohorts: LC-NCO, LC-CO, NLC-CO, and NLC-NCO, are presented in **Fig. 3A**. For each cohort, the total cumulative *TP53*-mutated colony prevalence value (colonies/10^6 cells/subject) was represented by the Y-intercept value at the LOD value of 100 cells/colony. The values were LC-NCO: 5.4 (87 colonies/10^6 cells/16 subjects), LC-CO: 5.14 (72 colonies/10^6 cells/14 subjects), NLC-CO: 2.71 (19 colonies/10^6 cells/7 subjects), and NLC-NCO: 0.64 (14 colonies/10^6 cells/22 subjects).

Based on direct inspection of these plots, it is evident that for each cohort, the number of colonies was inversely proportional to the colony size. Moreover, the distributions for the NLC cohorts (NLC-CO and NLC-NCO) were almost entirely restricted to colony sizes of <900 cells. In contrast, for the lung cancer cohorts (LC-CO and LC-NCO), a large fraction of colonies comprised >900 cells, with a maximum of 60,000 cells. Through retrospective analysis, we determined that these findings reproduce observations from a previously reported discovery cohort (11).

Among the lung cancer cohorts, we also analyzed the *TP53*-mutated colony size distributions after controlling for both COPD and smoking, each of which has a known effect on the adaptive immune response (**Fig. 3B**; ref. 29). Thus, a high adaptive immune lung cancer group comprising currently smoking COPD subjects (24 colonies/six subjects) was compared with a low adaptive immune lung cancer group comprising NCO subjects (87 colonies/10^6 cells/16 subjects). As is evident by inspection of **Fig. 3B**, the fraction of large colonies (>900 cells/colony) was higher among actively smoking COPD subjects (purple line) compared with NCO subjects (black line). This difference was nominally significant; the ratio of large to small colonies in the COPD active smoker group (2) was higher ($P < 0.05$) than the ratio of 0.72 for the NCO group.

Analysis of colonies driven by *TP53* mutations documented to be neoantigenic

Of the 192 *TP53*-mutated colonies measured, a subgroup of 29 colonies were driven by seven *TP53* mutations with the potential to be neoantigenic based on analysis of cancer tissues, including G245S, M237I, R175H, R248Q, R248W, R273C, and R273H (23). The intercohort differences in colony prevalence for colonies driven by this subset closely paralleled the differences for all *TP53* mutations (Supplementary Table S2) although significance testing

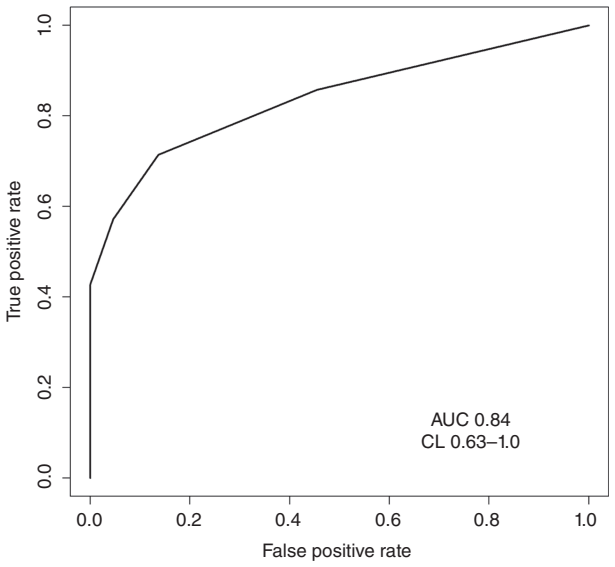


Figure 2.
Diagnostic accuracy of the *TP53* biomarker for COPD in NLC subjects ($N = 29$). Classification of subjects according to COPD status based on *TP53*-mutated colony prevalence (mutations/10^6 cells/subject) with ROC AUC and 95% confidence limits (CL) presented on the graph.

was limited by the small number. Specifically, there was a nominally significant 4.6-fold higher prevalence of this colony subset ($P = 0.017$, FDR $= 0.107$) among all lung cancer compared with all NLC and 4.6-fold higher among LC-NCO compared with NLC-NCO ($P = 0.029$, FDR $= 0.107$). There was also a 3.6-fold higher prevalence among LC-CO compared with NLC-CO and a 2.2-fold higher prevalence in COPD compared with NCO, but the colony numbers were too low for meaningful statistical comparison.

Effect of *TP53* mutation type on colony prevalence and size distribution

The relative prevalence of different *TP53* mutation types in preneoplastic colonies in the airway epithelium was compared with that previously reported by Barta and McMahon (30) for three lung cancer subtypes, including adeno, squamous, and small cell. Because the lung cancer tissue analysis reported mutations at the codon level, we compared preneoplastic mutations at the codon level as well. The representation of G245 mutations in preneoplastic colonies was 10.7%, which was >2 standard deviations higher than the mean reported for the three lung cancer subtypes (Supplementary Table S3). In a retrospective analysis, this finding is consistent with observations from our discovery cohort, in which G245 *TP53* mutations constituted the highest portion at 16% (11). Notably, the G245 *TP53* mutation fraction also varied according to the characteristics associated with adaptive immune inflammation, including COPD, cigarette smoking status, and colony size. Specifically, the G245 representation was

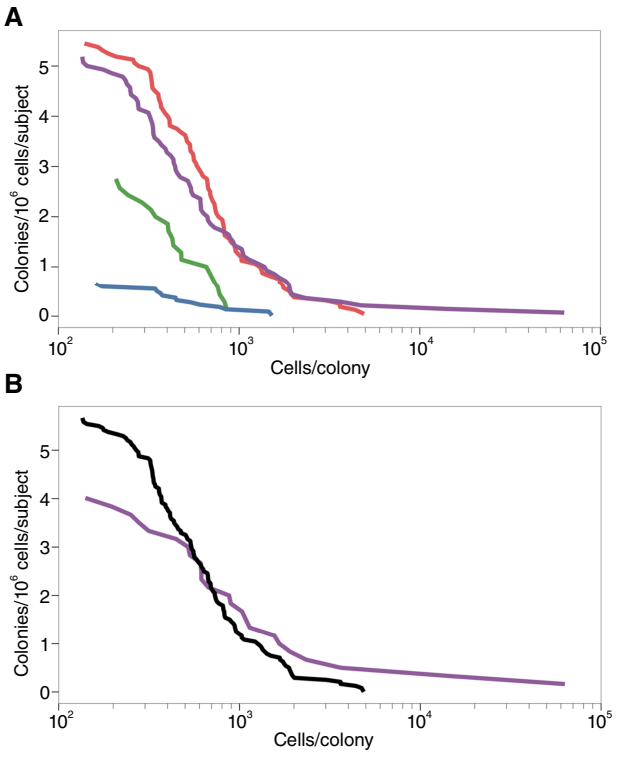


Figure 3.

Colony size distributions. Cohort-specific cumulative colony prevalence (colonies/10^6 cells/subject) was plotted relative to the colony size (cells/colony). **A,** Colony distributions for LC-NCO (red), LC-CO (purple), NLC-CO (green), and NLC-NCO (blue). **B,** Colony distribution for lung cancer subjects with low adaptive immune inflammation (LC-NCO; black) or high adaptive immune inflammation (active smoker LC-CO; purple).

2.7-fold higher among the actively smoking COPD subjects (25% vs. 9.2%). Moreover, there was a difference in the degree of reduction in the fraction of G245 mutated colonies among larger colonies (>900 cells). The G245 portion in large colonies decreased to 20% in actively smoking COPD and 4% in NCO. As such, the representation of G245 mutated increased from 2.7-fold higher among smaller colonies to 5-fold higher (20%/4%) among larger colonies. None of the other *TP53* mutation types were enriched to a different degree in preneoplastic colonies compared with lung cancer tissue or varied according to colony size, COPD, and cigarette smoking status.

Effect of colony size, lung cancer status, and COPD status on the probability of colony survival

We used Cox proportional hazards analysis to test the relationship between *TP53*-mutated colony size and the probability of colony survival in lung cancer and NLC subjects. Larger colonies were more likely to survive among patients with lung cancer ($P = 0.0026$; **Fig. 4A**). When the analysis was restricted to the NCO or COPD cohorts, this difference remained significant, and there were qualitative differences (**Fig. 4B** and **C**). Among NCO subjects, the colony survival difference between lung cancer and NLC was distributed across the range of colony sizes (**Fig. 4B**), whereas among COPD subjects, the difference was restricted to colonies with sizes greater than approximately 900 cells (**Fig. 4C**).

To explore the relationships between these Cox proportional hazards plots and the performance of the biomarker to classify lung cancer or COPD subjects, we tested the effect of a *TP53*-mutated colony size lower limit on ROC AUC for classification of lung cancer or COPD subjects (Supplementary Fig. S1). For each group, the AUC was close to the maximum when all colonies were included, regardless of their size. However, the effect of increasing the colony size lower limit was different for each cohort. Specifically, for the classification of COPD in NLC subjects, the AUC decreased to a minimum as the lower limit increased from 100 to 900 cells because there were no colonies larger than this (**Fig. 3**). In contrast, for the diagnosis of lung cancer among COPD subjects, as the lower limit increased over this same span (100–900 cells), the AUC was relatively unchanged from the maximum. However, as the lower limit increased from 900 to 60,000 cells, the AUC decreased steadily to a minimum. Thus, among COPD subjects, the classification of lung cancer versus NLC was primarily due to the large difference in the prevalence of colonies comprising 900 or more cells, consistent with the data presented in **Fig. 4C**.

In contrast to COPD subjects, the AUC for the diagnosis of lung cancer among NCO subjects decreased steadily across the range of colony sizes as the lower limit increased from 100 cells to the maximum observed colony size of 10,000 cells (Supplementary Fig. S1), consistent with the data presented in **Fig. 4B**. This pattern is due to the very low prevalence of *TP53*-mutated colonies of any size in subjects with neither lung cancer nor COPD (NLC-NCO; **Fig. 3**).

Discussion

We used US-NGS to analyze the prevalence and size distributions of microscopic preneoplastic *TP53*-mutated colonies in the airways of subjects with conditions known to be associated with suboptimal DNA protection (31–33) and/or increased adaptive immune inflammation in the lungs (18–20, 29). These conditions included lung cancer (present or absent), COPD (present or absent), and cigarette smoking (current or former).

In **Fig. 5A**, we schematically integrated qualitative and quantitative differences among the four cohorts of heavy smokers based on published knowledge and data presented here. As indicated by cohort circle sizes, (i) approximately 35% of screening-eligible heavy smokers have COPD (represented by NLC-CO + LC-CO; ref. 5), whereas 65% do not (represented by NLC-NCO + LC-NCO), and (ii) COPD diagnosis is enriched among heavy smokers with lung cancer; about half have COPD (LC-CO) and half do not (LC-NCO). Each cohort was positioned along the *x*-axis according to the prevalence of *TP53*-mutated colonies measured in this study (**Fig. 3**), and this value corresponded to the known relative

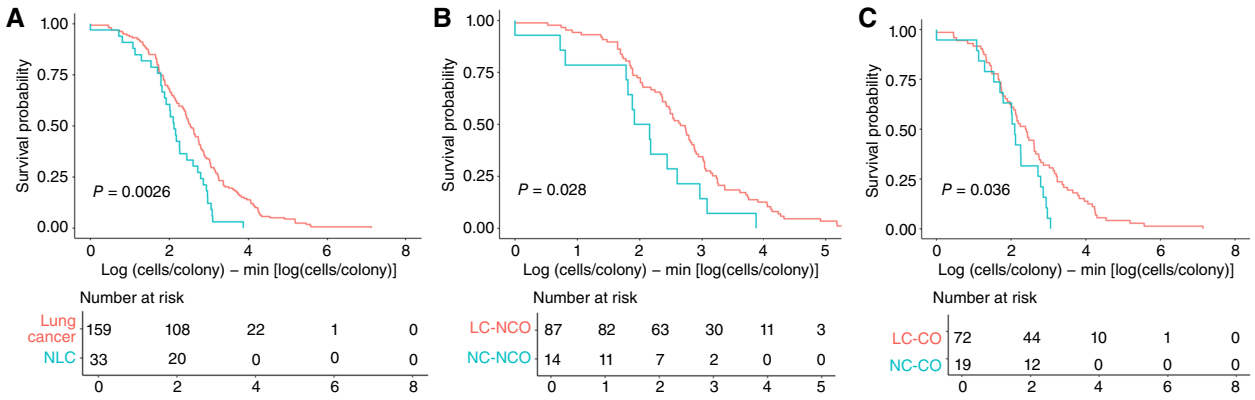


Figure 4.

Effect of colony size on colony survival in lung cancer vs. NLC subjects based on Cox proportional hazards analysis. **A,** All lung cancer vs. all NLC. **B,** Lung cancer vs. NLC among NCO subjects. **C,** Lung cancer vs. NLC among CO subjects. The colony size values (cells/10^6 cells) were $\log_2$-transformed. The colony size LOD [min ($\log_2$(cells/colony)] was 100 cells ($2^{6.554}$ cells), and this was given the value 0 on the *x*-axis. All other values were calculated as follows: Log_2 (cells/colony) − min [$\log_2$(cells/colony)]. The tables provide the number of colonies surviving at each colony size interval.

DNA repair capacity (33, 34). The position along the *y*-axis was based on immune surveillance capability, which is closely associated with COPD diagnosis (18, 19). Significant intercohort classification by the *TP53* mutation biomarker measured in this study is indicated with *P* values (Kruskal–Wallis).

TP53 biomarker classification of lung cancer subjects after controlling for COPD

Currently, LDCT screening eligibility is based on demographic criteria designed to identify patients with a lung cancer risk of at least 0.03%/year (35, 36). However, about 50% of individuals who develop lung cancer do not meet this threshold due to their younger age and lower smoking history (37). The *TP53* biomarker improves accuracy in diagnosing lung cancer in subjects without COPD (**Figs. 1A–C** and **5A**), who also tend to be younger and have a lower smoking history. We expect that the measurement of the *TP53* biomarker in such individuals, in combination with demographic criteria, will identify a large fraction with a lung cancer risk >0.03%/year and enable them to qualify for screening.

TP53 biomarker classification of COPD subjects

The *TP53* biomarker performed well in classifying COPD subjects among heavy smokers without lung cancer (**Figs. 2** and **5A**). If validated in future clinical trials, this biomarker may provide additional avenues for investigating the mechanistic basis of COPD causation by cigarette smoke inhalation. Furthermore, it may facilitate the identification of

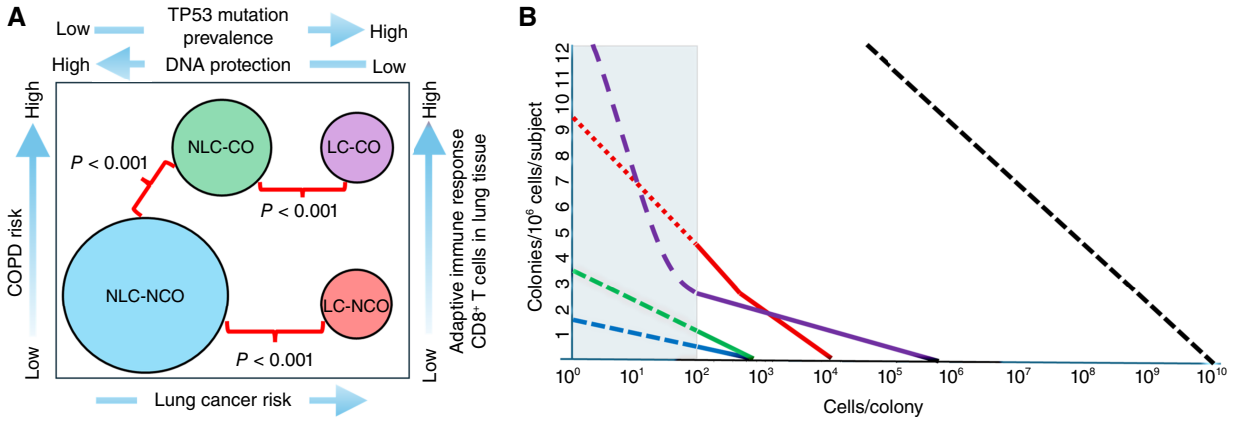


Figure 5.

Schematic models for the effect of DNA repair capacity and adaptive immune activity on *TP53*-mutated colonies in the airway epithelium of heavy cigarette smokers. **A,** Population distribution model. The four cohorts of screening-eligible heavy smokers are represented according to approximate relative size in the population (circle size) and location with respect to DNA repair capacity (*x*-axis) and immune surveillance capability (*y*-axis). *P* values (Kruskal–Wallis) associated with the classification of representative cohorts by *TP53*-mutated colony prevalence are indicated. **B,** Colony distribution model. Solid lines correspond to measured results displayed in **Fig. 3**. Dashed lines are hypothetical. The blue shaded area displays hypothesized relationships below the LOD. Black line indicates an imaginary cohort of heavy smokers with a high mutagenesis rate and no adaptive immune control. Purple line indicates LC-CO; lung cancer subjects with COPD. Red line indicates LC-NCO subjects. Green line indicates NLC-CO subjects. Blue line indicates NLC-NCO subjects.

individuals most likely to benefit from early intervention toward the goal of COPD prevention. A positive biomarker test may identify patients for whom the risk is sufficiently high to justify LDCT screening for COPD although it is not currently recommended based on demographic criteria alone (38). Coincidentally, asymptomatic emphysema evident on LDCT chest imaging is a significant risk factor for lung cancer, independent of obstruction (39). This group tends to be younger, has a lower smoking history, and is more likely to benefit from lung cancer screening.

A recent review compiled evidence that genetic selection pressure caused by chronic inhalation of wood smoke used for cooking and warmth resulted in the evolution of robust DNA repair capacity in human lung tissues (40). This is consistent with the observation that most people who smoke heavily (i) do not develop lung cancer or COPD (5, 41) and (ii) have a low to nonmeasurable prevalence of *TP53*-mutated colonies in their airways, as observed for cohort NLC-NCO (**Figs. 3** and **5A**). However, there is wide interindividual variation for this trait in the population, and a significant fraction of heavy smokers have poor DNA repair capacity, which increases their risk of lung cancer, COPD, or both. According to the model presented in **Fig. 5A**, heavy smokers with suboptimal DNA repair were represented in this study by three cohorts, each of which displayed a high *TP53*-mutated colony prevalence (**Fig. 3A**). These cohorts are represented schematically in **Fig. 5A** as follows: (i) *Lung cancer subjects without COPD* (*LC-NCO*): These subjects had a high prevalence of TP53-mutated colonies but did not have a corresponding adaptive immune response to *TP53*-mutated colonies. Because the majority of screening-eligible subjects (i.e., >20 pack-years smoking, 50 years or more of age) have not been diagnosed with COPD (5), it is reasonable to hypothesize that the adaptive immune response in the lungs, although generally robust, has evolved in many individuals to be relatively tolerant of the acquired neoantigens represented by *TP53*-mutated colonies. For example, it is possible that certain *TP53* mutations only serve as neoantigens in individuals with certain HLA types. (ii) *NLC subjects with COPD* (*NLC-CO*): In these subjects, suboptimal DNA repair results in an increased prevalence of *TP53*-mutated colonies, but they are protected from lung cancer due to a robust adaptive immune response. However, chronic adaptive immune inflammation damages the lung tissue, resulting in COPD (18, 19). (iii) *Lung cancer subjects with COPD* (*LC-CO*): There is a robust adaptive immune response in cigarette smoke–induced *TP53*-mutated colonies. However, according to this model, the DNA repair capacity is so poor that the adaptive immune response cannot compensate for the high rate at which mutated colonies are generated. Moreover, the accompanying inflammation causes COPD damage to the lung tissue, similar to NLC-CO.

Analysis of TP53-mutated colony size distributions

The proposed schematic model in **Fig. 5B** is an effort to integrate the *TP53*-mutated colony size distributions measured in this study with established evidence about intercohort differences in DNA repair and adaptive immune capacities associated with lung cancer, COPD, and active cigarette smoking.

Arbitrary reference cohort assigned colony distribution

As a point of reference, the dashed black line in **Fig. 5B** represents the *TP53*-mutated colony distribution in a hypothetical cohort of long-term heavy smokers under assumed conditions in which there was (i) a continuous rate of cigarette smoke–induced *TP53* mutagenesis, (ii) a continuous rate of population doubling within a colony, (iii) new *TP53*-mutated colonies generated at a rate of one colony/population doubling, (iv) no adaptive immune surveillance, and (v) no selection for resistance to adaptive immune surveillance. In normal nonsmokers, the replication rate for basal AECs is 30 to 50 days, with only 1% of cells replicating at any one time; however, in heavy smokers, up to 30% of basal cells are in replication (42, 43). In this arbitrary cohort, it was conservatively assumed that *TP53*-mutated cells had a slow replication rate associated with normal basal cells. Even at this relatively slow rate, under the assumed conditions, it would take less than 36 months for a *TP53*-mutated colony induced by inhaled cigarette smoke to reach the size of lung cancer detectable as a nodule on CT ($2^{33} = 10^{10}$ cells). In this model, a continuous rate of cigarette smoke–induced *TP53* mutagenesis over many years was assumed, with different stem cells becoming mutated at different times. Under these conditions, the number of colonies in each colony size bin would be approximately the same, resulting in a straight line with a negative slope in the cumulative distribution plot, as schematically presented.

Observed colony distributions

The observed *TP53*-mutated colony distributions (**Fig. 3**) were in marked contrast to those of the cohort with the arbitrarily assigned conditions. We propose that interindividual variation in the adaptive immune response to TP53 neoantigens explains several features of these differences.

Lung cancer subjects: The maximum observed colony size of 60,000 cells in lung cancer subjects (LC-CO and LC-NCO; **Fig. 3**) was much smaller than the 10^{10} cells expected under very conservative assumptions in the arbitrary model. Furthermore, in a model that assumes no inhibition due to immune surveillance or other factors, the cumulative colony distribution associated with a maximum colony size of 60,000 cells (2^{16} population doublings) would be a straight line with a maximum colony prevalence of 18 colonies/10^6 cells/subject at an LOD of 100 cells/colony (Y-intercept). In contrast, the observed maximum colony prevalence for lung cancer subjects was only five colonies/10^6 cells/subject, and the distribution plot was coved, not straight.

Effect of COPD and smoking: For lung cancer subjects with COPD (represented by cohort LC-CO), cigarette smoke–induced mutagenesis is high due to relatively ineffective DNA

repair (34), but this is partially offset by a more robust adaptive immune surveillance compared with those without COPD (12, 18, 19). According to the LC-CO model [**Fig. 5B** (purple line)], abundant microscopic neoantigenic *TP53*-mutated colonies are subjected to an intense adaptive immune response, and their prevalence begins to decrease rapidly at colony sizes below the LOD for our test (100 cells/colony), as indicated by the dashed line. This would explain the observation that at the LOD, the cumulative colony prevalence for LC-CO subjects was lower than that for LC-NCO subjects, as observed in **Fig. 5A**. Furthermore, according to the model, the intense immune response in LC-CO selects for colonies with greater immune avoidance characteristics. This model explains the skew to larger colonies among LC-CO compared with LC-NCO subjects. Moreover, it is well-documented that the cigarette smoke–induced mutagenesis rate and adaptive immune response are particularly high among COPD subjects who are actively smoking (12, 18, 19), and our observations presented in **Fig. 5B** are consistent with this. Specifically, among actively smoking COPD subjects, we observed a further reduction in colony prevalence at the LOD and a further increase in the ratio of large to small colonies.

Lung cancer subjects without COPD: According to the LC-NCO model (**Fig. 5B**), the mutagenesis rate was lower, and the starting number of mutated colonies was lower than that in LC-CO, as modeled by the red dashed line below the LOD. However, in the less robust immune environment in NCO subjects, the pool of preneoplastic colonies is large enough to result in the selection of a measurable number of resistant colonies, with an observed skew to larger colonies (red solid line) compared with NLC subjects. This model is consistent with the observation that NCO individuals who develop lung cancer are younger and smoke less (37).

NLC subjects

It is reasonable to conclude that among cigarette smokers at low risk for COPD or lung cancer, there is, in various combinations, (i) robust DNA repair that prevents cigarette smoke–induced mutagenesis, resulting in a low prevalence of *TP53*-mutated colonies, as is evident in NLC-NCO subjects [**Fig. 5B** (blue line)], and (ii) a robust adaptive immune response that eliminates *TP53*-mutated colonies, as is evident in NLC-CO subjects (green line). Notably, these lines are not concave, consistent with the data presented (**Fig. 3A**). This is what we would expect to observe in the absence of colonies with immune surveillance avoidance properties. According to the model, we would not expect such colonies to arise due to the much lower number of colonies generated compared with lung cancer subjects; thus, a lower pool from which to select resistant colonies.

Effect of TP53 mutation type

A subfraction of *TP53* mutations measured in this study (Supplementary Tables S2 and S3) has been experimentally confirmed to be neoantigenic and to drive a T-cell adaptive immune response in cancer tissues (23). The association of this subfraction with lung cancer and COPD parallels that observed for all *TP53* mutations although the small numbers limit significance testing. We hypothesize that most, if not all, *TP53* mutations measured in this study are neoantigenic and that colonies driven by these mutations stimulate an adaptive immune response when the colonies are still relatively small, comprising fewer than 100,000 cells. We believe that the shapes of the colony distribution plots (**Fig. 3A** and **B**) modeled in **Fig. 5B** are consistent with this hypothesis. Proof of this hypothesis will require focused experimentation.

Proposed role of TP53-mutated preneoplastic colonies in lung cancer development

The observed *TP53*-mutated colony distributions associated with COPD and active smoking among lung cancer subjects also support the role of the adaptive immune response in determining the genetic characteristics of the cells that form lung cancers, as depicted in the hypothetical model displayed in **Fig. 6**. In **Fig. 6** (top), cigarette smoke induces *TP53* mutations that drive colony formation. As previously reported, microscopic *TP53*-mutated colonies are about 10-fold more prevalent than colonies driven by any other driver gene in the airways of heavy smokers with lung cancer (11). In **Fig. 6** (middle), neoantigenic *TP53* mutation-driven colonies induce an adaptive immune response that causes cytotoxic damage (represented by the gray cell). As presented in **Fig. 6** (bottom left) the chronic intense adaptive immune inflammation damages normal lung tissue, resulting in COPD. cytotoxic adaptive immune response selects cells that are immune-resistant, commonly through PD-L1 upregulation, and are able to expand into larger colonies with a selective growth advantage over nonresistant colonies that continue to be formed due to persistent cigarette smoke exposure. As presented in **Fig. 6** (bottom right), cytoxic adaptive immune response selects for cells that are immune-resistant, commonly through PD-L1 upregulation, and are able to expand into larger colonies with a selective growth advantage over nonresistant colonies that continue to be formed due to persistent cigarette smoke exposure. These colonies of immune-resistant cells with defective DNA repair due to mutated TP53 present more numerous targets for additional mutations, including *EGFR* or *KRAS* oncogene mutations, that result in lung cancer. This model is consistent with data presented here and with established conceptual models (14) and is supported by the analysis of lung cancer tissues. For example, a *TP53* mutation drives about 50% of lung cancers, and the presence of a *TP53* driver mutation in lung cancer tissue is strongly correlated with both the infiltration of $CD8^+$ T cells and the upregulation of the PD-L1 inhibitor of adaptive immune surveillance (21, 22, 24, 44). The most prevalent mutations observed in the *TP53*-mutated microscopic colonies reported here (Supplementary Tables S2 and S3) are also (i) the most prevalent driver mutations in lung cancers (21, 22) and (ii) documented to be neoantigenic in cancer tissues (23). Our analyses thus far have not included the study of *TP53*-mutated preneoplastic

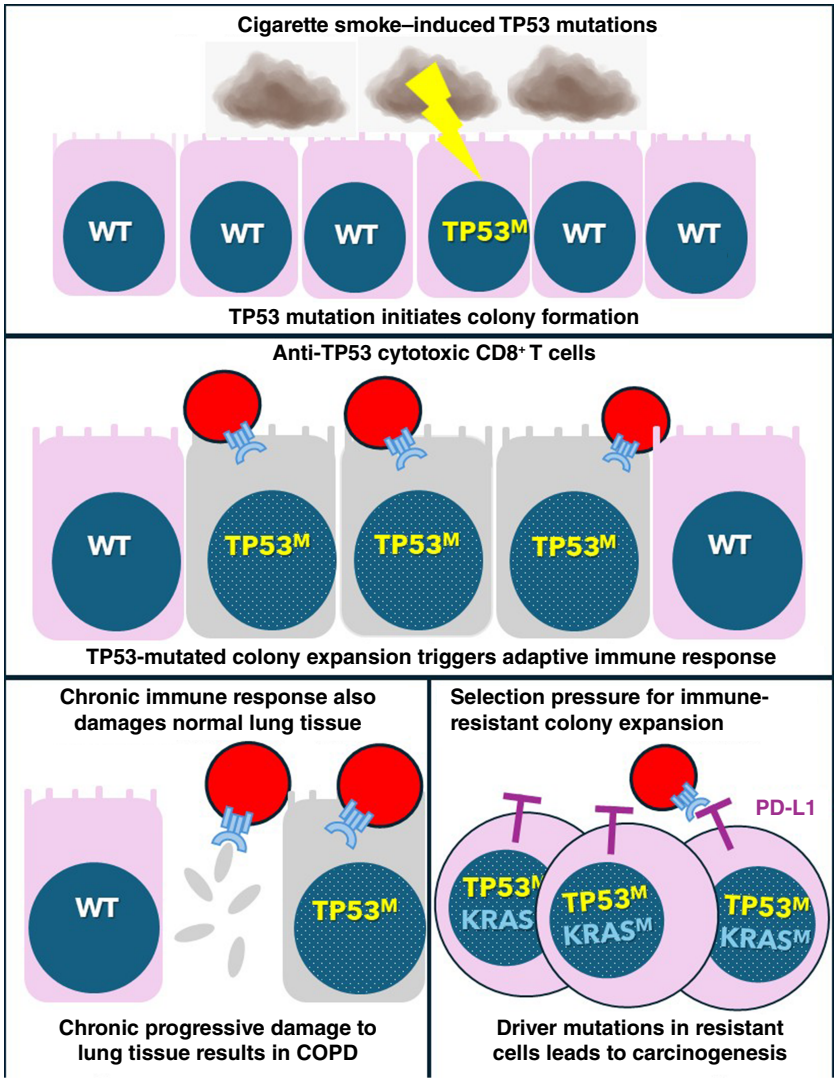


Figure 6.

A proposed model of lung cancer development beginning with *TP53*-mutated preneoplastic colonies. Top, Inhalation exposure to mutagens in cigarette smoke induces mutations in the airway epithelium, some of which are *TP53* mutations that induce proliferation and colony formation. Middle, Neoantigens produced by the expanding colony draw an adaptive immune response that is cytotoxic and selects for immune resistance. Bottom left, Intense chronic adaptive immune inflammation damages normal lung tissue, resulting in COPD, Bottom right, An expanding colony of *TP53*- mutated cells that is immune-resistant due to PD-L1 upregulation is an enlarging target for additional cigarette smoke–induced mutations. A *TP53*-mutated, immune-resistant cell acquires a *KRAS* activating mutation and expands to become a cancer with upregulated PD-L1 and abundant infiltrating $CD8^+$ T cells.

colony prevalence related to lung cancers in nonsmokers. However, recent data support the role of noncigarette smoke–induced *TP53* mutation as the first step in a significant fraction of EGFR-driven lung cancers in nonsmokers. For example, as with cigarette smoke–associated lung cancer, about half of EGFR-driven lung cancers in nonsmokers have a *TP53* comutation (45). Furthermore, it was recently reported that *TP53*-mutated preneoplastic colony prevalence is higher in the tumor-adjacent normal colon epithelium in colon cancer subjects compared with matched controls (46). Although we have not yet studied lung cancer in nonsmokers, it is reasonable to hypothesize that the prevalence of *TP53*-mutated colonies in tumor-adjacent lung tissue will be higher in nonsmoker subjects with *EGFR* and *TP53* co-mutated lung cancers compared with matched controls, as was observed for *KRAS* and *TP53* co-mutated colon cancer. If so, the basis for this higher prevalence in colon or nonsmoker lung cancer subjects could be in part due to inherited suboptimal DNA repair and adaptive immune response. Additional exogenous risk factors, including radon exposure, could further increase the prevalence of *TP53*-mutated colonies and the risk for subsequent driver mutations within these colonies, whereas germline inherited predisposition to *EGFR* mutations may predispose to *EGFR* driver mutations within these TP53-mutated preneoplastic colonies.

Evidence that an adaptive immune response to *TP53*-mutated preneoplastic colonies induced by cigarette smoke exposure is mechanistically associated with COPD causation

The early stages of COPD are associated with chronic activation of the adaptive immune system, manifested as infiltration of $CD8^+$ T cells into the airway epithelium and lung parenchyma (18, 19). $CD8^+$ T-cell infiltration is associated with additional inflammatory signals, and the effects of T cells and associated inflammatory cells are major contributors to lung tissue damage observed in COPD (18–20).

The colony distribution plots presented here (**Fig. 3**) support the hypothesis that inhaled cigarette smoke–induced neoantigenic *TP53* mutations signal the immune surveillance system that preneoplastic colonies are present and need to be cleared. In this context, it is notable that the subfraction of colonies driven by *TP53* mutations, documented to be neoantigenic in cancer tissues, also has a higher prevalence among subjects with COPD although this does not reach significance in this study (Supplementary Table S2). Further, in mouse models of COPD, cigarette smoke inhalation induced a strong adaptive immune response after 24 weeks, with increased infiltration of $CD8^+$ cells into lung tissue (47), and knockout of the CD8 gene prevented the development of COPD in this mouse model. Moreover, in a lung cancer mouse model in which one allele of *TP53* was knocked out, inhalation of cigarette smoke over the same time period of 24 weeks was associated with a marked increase in lung tumors over the subsequent 4 months, consistent with the generation of *TP53*-mutated colonies during the exposure period (48). Thus, based on data from the mouse lung cancer model (48), it is reasonable to assume that abundant *TP53*-mutated colonies formed in the mouse COPD model over the same cigarette smoke exposure time in which there was documented $CD8^+$ T-cell infiltration (47). Importantly, these data are correlative and will require focused experimentation to confirm a mechanistic association.

A proposed explanation for the association of COPD with response to anti–PD-1 treatment

Patients with advanced-stage NSCLC and COPD treated with nivolumab, an anti–PD-1 antibody, have longer progression-free survival than patients without COPD (29). Further, there is an association between the severity of smoking (based on pack-years) and response to nivolumab, and this association is preferentially observed in patients with COPD (29). It was proposed that heavier smoking in COPD subjects is associated with more mutations in the tumors, which is why they respond to immune modulator therapy. The data presented here raise the possibility of an additional, not mutually exclusive, explanation. Specifically, we report that *TP53*-mutated colony prevalence and colony size distribution in airway epithelium are strongly correlated with COPD and active smoking (**Fig. 3**), conditions documented to be associated with the upregulation of $CD8^+$ T-cell–driven adaptive immune response (18, 20, 29). We hypothesize that the increased lung cancer tumor response to anti–PD-1 antibody in subjects with COPD who are also active smokers may be in part due to activated $CD8^+$ T cells infiltrating the highly prevalent and widely dispersed preneoplastic *TP53*-mutated colonies in these patients.

Strategy to identify biomarkers that predict anti–PD-L1 treatment response

Based on the data presented here, it would be reasonable to explore *TP53*-mutated colony size distribution, particularly for G245 *TP53*-mutated colonies, as an indirect measure of adaptive immune capacity, as a predictor of response to immune modulation (PD-L1 modulation or other) therapy, and possibly to assess response to immune modulation in relatinship to prevention. For example, a reduction in the survival of the large *TP53*-mutated colony fraction among lung cancer subjects following anti–PD-1 treatment compared with that observed in untreated lung cancer subjects would support the hypothesis that large colonies survived due to immune surveillance avoidance characteristics. In support of this proposal, it was recently reported that patients with metastatic lung cancer have a reduced risk for a second primary lung cancer after treatment with immune modulation versus chemotherapy (49) and that active cigarette smoking is a significant second factor. As such, a reduction in the survival of large *TP53*-mutated colonies may be tested for the prevention of both primary lung cancer and second primary lung cancer.

Data Availability

Sequencing data for the cohort are available through the National Library of Medicine National Center for Biotechnology Information under SRA number PRJNA1001394, https://www.ncbi.nlm.nih.gov/sra/PRJNA1001394. Readers can request the deidentified demographic data from the corresponding author.

Authors' Disclosures

J.C. Willey reports grants from the NCI during the conduct of the study, as well as other support from Accugenomics, Inc. and GRAIL, Inc. outside the submitted work, and a patent for US patent 12460247 issued, licensed, and with royalties paid from Accugenomics, Inc.; a patent for US patent application 20250369055 pending; and a patent for US patent 9944973 issued, licensed, and with royalties paid from Accugenomics, Inc. E.L. Crawford reports grants from NIH/NCI, NIH/NHGRI, and the George Isaac Cancer Research Fund during the conduct of the study, as well as other support from Accugenomics, Inc. outside the submitted work, and a patent for US Patent 12,460,247 B2 issued, licensed, and with royalties paid from Accugenomics, Inc.; a patent for US Patent 11,312,987 issued, licensed, and with royalties paid from Accugenomics, Inc.; a patent for US Patent 9,944,973 B2 issued, licensed, and with royalties paid from Accugenomics, Inc.; a patent for US Patent 62/897,343 pending; and a patent for US Patent 62/935,705 pending. T.M. Blomquist reports a patent for WO2014082032 A1 issued, licensed, and with royalties paid from Accugenomics LLC. S.A. Deppen reports grants from NCI and nonfinancial support from NIH during the conduct of the study; and GO2 For Lung Cancer Travel Funds, unpaid consultant. D.-J. Craig reports grants from NIH/NCI, NIH/NHGRI, and the George Isaac Cancer Research Fund during the conduct of the study, as well as a patent for PCT/US2020/049629 pending. No disclosures were reported by the other authors.

Authors' Contributions

J.C. Willey: Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **A.R. Boring:** Conceptualization, data curation, software, formal analysis, investigation, writing–review and editing. **E.L. Crawford:** Data curation, investigation, methodology, project administration, writing–review and editing. **K.N. Lei:** Resources, software, methodology, writing–review and editing. **T.M. Blomquist:**

Software, methodology, writing–review and editing. **C.I. Amos:** Writing–review and editing. **S.A. Deppen:** Resources, writing–review and editing. **E.L. Grogan:** Resources, writing–review and editing. **H. Chen:** Data curation, formal analysis, writing–review and editing. **M. Omballi:** Resources. **D.J. Craig:** Software, methodology, writing–review and editing.

Acknowledgments

This study was funded in part by the US NCI, 5U24CA086368 (J.C. Willey, E.L. Grogan, and S.A. Deppen), U01CA152662 (E.L. Grogan, S.A. Deppen, H. Chen, and J.C. Willey), R01CA243483-05 (C.I. Amos and J.C. Willey), and the George Isaac Family Endowed Chair for Cancer Research (J.C. Willey).

Note

Supplementary data for this article are available at Cancer Prevention Research Online (http://cancerprevres.aacrjournals.org/).

Received April 10, 2026; revised June 23, 2026; accepted July 22, 2026; posted first July 23, 2026.

References

1. Humphrey LL, Deffebach M, Pappas M, Baumann C, Artis K, Mitchell JP, et al. Screening for lung cancer with low-dose computed tomography: a systematic review to update the US Preventive Services Task Force recommendation. Ann Intern Med 2013;159:411–20. https://doi.org/10.7326/0003-4819-159-6-201309170-00690.
2. Wang Y, Midthun DE, Wampfler JA, Deng B, Stoddard SM, Zhang S, et al. Trends in the proportion of patients with lung cancer meeting screening criteria. JAMA 2015;313:853–5. https://doi.org/10.1001/jama.2015.413.
3. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection criteria for lung-cancer screening. N Engl J Med 2013;368:728–36. https://doi.org/10.1056/NEJMoa1211776.
4. US Preventive Services Task Force; Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, Caughey AB, et al. Screening for lung cancer: US Preventive Services Task Force recommendation statement. JAMA 2021;325:962–70. https://doi.org/10.1001/jama.2021.1117.
5. Young RP, Hopkins RJ. Chronic obstructive pulmonary disease (COPD) and lung cancer screening. Transl Lung Cancer Res 2018;7:347–60. https://doi.org/10.21037/tlcr.2018.05.04.
6. Arenberg D, Silvestri GA. Understanding when cancer screening is likely to help and when it may not. Chest 2026;169:24–5. https://doi.org/10.1016/j.chest.2025.09.023.
7. Rivera MP, Tanner NT, Silvestri GA, Detterbeck FC, Tammemägi MC, Young RP, et al. Incorporating coexisting chronic illness into decisions about patient selection for lung cancer screening. An official American Thoracic Society research statement. Am J Respir Crit Care Med 2018;198:e3–13. https://doi.org/10.1164/rccm.201805-0986ST.
8. Bennett WP, Colby TV, Travis WD, Borkowski A, Jones RT, Lane DP, et al. p53 protein accumulates frequently in early bronchial neoplasia. Cancer Res 1993;53:4817–22.
9. Breuer RH, Pasic A, Smit EF, van Vliet E, Vonk Noordegraaf AV, Risse EJ, et al. The natural course of preneoplastic lesions in bronchial epithelium. Clin Cancer Res 2005;11:537–43. https://doi.org/10.1158/1078-0432.537.11.2.
10. Wistuba II, Mao L, Gazdar AF. Smoking molecular damage in bronchial epithelium. Oncogene 2002;21:7298–306. https://doi.org/10.1038/sj.onc.1205806.
11. Craig DJ, Morrison T, Khuder SA, Crawford EL, Wu L, Xu J, et al. Technical advance in targeted NGS analysis enables identification of lung cancer risk-associated low frequency TP53, PIK3CA, and BRAF mutations in airway epithelial cells. BMC cancer 2019;19:1081. https://doi.org/10.1186/s12885-019-6313-x.
12. Yoshida K, Gowers KHC, Lee-Six H, Chandrasekharan DP, Coorens T, Maughan EF, et al. Tobacco smoking and somatic mutations in human bronchial epithelium. Nature 2020;578:266–72. https://doi.org/10.1038/s41586-020-1961-1.
13. Craig DJ, Crawford EL, Chen H, Grogan EL, Deppen SA, Morrison T, et al. TP53 mutation prevalence in normal airway epithelium as a biomarker for lung cancer risk. BMC Cancer 2023;23:783. https://doi.org/10.1186/s12885-023-11266-7.
14. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 2002;3:991–8. https://doi.org/10.1038/ni1102-991.
15. Mascaux C, Angelova M, Vasaturo A, Beane J, Hijazi K, Anthoine G, et al. Immune evasion before tumour invasion in early lung squamous carcinogenesis. Nature 2019;571:570–5. https://doi.org/10.1038/s41586-019-1330-0.
16. Moghaddam SJ, Savai R, Salehi-Rad R, Sengupta S, Kammer MN, Massion P, et al. Premalignant progression in the lung: knowledge gaps and novel opportunities for interception of non-small cell lung cancer. An official American Thoracic Society research statement. Am J Respir Crit Care Med 2024;210:548–71. https://doi.org/10.1164/rccm.202406-1168ST.
17. Willey JC, Morrison TB, Austermiller B, Crawford EL, Craig DJ, Blomquist TM, et al. Advancing NGS quality control to enable measurement of actionable mutations in circulating tumor DNA. Cell Rep Methods 2021;1:100106. https://doi.org/10.1016/j.crmeth.2021.100106.
18. Cosio MG, Saetta M, Agusti A. Immunologic aspects of chronic obstructive pulmonary disease. N Engl J Med 2009;360:2445–54. https://doi.org/10.1056/NEJMra0804752.
19. Soriano JB, Polverino F, Cosio BG. What is early COPD and why is it important? Eur Respir J 2018;52:1801448. https://doi.org/10.1183/13993003.01448-2018.
20. Maeno T, Houghton AM, Quintero PA, Grumelli S, Owen CA, Shapiro SD. $CD8^+$ T cells are required for inflammation and destruction in cigarette smoke-induced emphysema in mice. J Immunol 2007;178:8090–6. https://doi.org/10.4049/jimmunol.178.12.8090.
21. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50. https://doi.org/10.1038/nature13385.
22. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519–25. https://doi.org/10.1038/nature11404.
23. Kim SP, Vale NR, Zacharakis N, Krishna S, Yu ZY, Gasmi B, et al. Adoptive cellular therapy with autologous tumor-infiltrating lymphocytes and T-cell receptor- engineered T cells targeting common p53 neoantigens in human solid tumors. Cancer Immunol Res 2022;10:932–46. https://doi.org/10.1158/2326-6066.CIR-22-0040.
24. Cha YJ, Kim HR, Lee CY, Cho BC, Shim HS. Clinicopathological and prognostic significance of programmed cell death ligand-1 expression in lung adenocarcinoma and its relationship with p53 status. Lung Cancer 2016;97:73–80. https://doi.org/10.1016/j.lungcan.2016.05.001.
25. Boveri T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. J Cell Sci 2008;121 Suppl 1:1–84. https://doi.org/10.1242/jcs.025742.

26. Huang PW, Chang JWC. Immune checkpoint inhibitors win the 2018 Nobel Prize. Biomed J 2019;42:299–306. https://doi.org/10.1016/j.bj.2019.09.002.
27. Kandoth C, McLellan MD, Vandin F, Ye K, Niu BF, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333–9. https://doi.org/10.1038/nature12634.
28. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods 2014;11:396–8. https://doi.org/10.1038/nmeth.2883.
29. Biton J, Ouakrim H, Dechartres A, Alifano M, Mansuet-Lupo A, Si H, et al. Impaired tumor-infiltrating T cells in patients with chronic obstructive pulmonary disease impact lung cancer response to PD-1 blockade. Am J Respir Crit Care Med 2018;198:928–40. https://doi.org/10.1164/rccm.201706-1110OC.
30. Barta JA, McMahon SB. Lung-enriched mutations in the p53 tumor suppressor: a paradigm for tissue-specific gain of oncogenic function. Mol Cancer Res 2019;17:3–9. https://doi.org/10.1158/1541-7786.MCR-18-0357.
31. Blomquist T, Crawford EL, Mullins D, Yoon Y, Hernandez DA, Khuder S, et al. Pattern of antioxidant and DNA repair gene expression in normal airway epithelium associated with lung cancer diagnosis. Cancer Res 2009;69:8629–35. https://doi.org/10.1158/0008-5472.CAN-09-1568.
32. Yeo J, Morales DA, Chen T, Crawford EL, Zhang X, Blomquist TM, et al. RNAseq analysis of bronchial epithelial cells to identify COPD-associated genes and SNPs. BMC Pulm Med 2018;18:42. https://doi.org/10.1186/s12890-018-0603-y.
33. Wu HC, Kehm R, Santella RM, Brenner DJ, Terry MB. DNA repair phenotype and cancer risk: a systematic review and meta-analysis of 55 case-control studies. Sci Rep 2022;12:3405. https://doi.org/10.1038/s41598-022-07256-7.
34. Sauler M, Lamontagne M, Finnemore E, Herazo-Maya JD, Tedrow J, Zhang X, et al. The DNA repair transcriptome in severe COPD. Eur Respir J 2018;52:1701994. https://doi.org/10.1183/13993003.01994-2017.
35. Kim EE, Cenzer I, Graham FJ, Kang J, Lee SJ, Rustagi AS. Time to benefit for lung cancer screening: a systematic review and survival meta-analysis. Am J Prev Med 2025;69:107736. https://doi.org/10.1016/j.amepre.2025.107736.
36. Kovalchik SA, Tammemagi M, Berg CD, Caporaso NE, Riley TL, Korch M, et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. N Engl J Med 2013;369:245–54. https://doi.org/10.1056/NEJMoa1301851.
37. Zhai R, Yu X, Shafer A, Wain JC, Christiani DC. The impact of coexisting copd on survival of patients with early-stage non-small cell lung cancer undergoing surgical resection. Chest 2014;145:346–53. https://doi.org/10.1378/chest.13-1176.
38. Mulshine JL, Pyenson B, Healton C, Aldige C, Avila RS, Blum T, et al. Paradigm shift in early detection: lung cancer screening to comprehensive CT screening. Eur J Cancer 2025;218:115264. https://doi.org/10.1016/j.ejca.2025.115264.
39. de Torres JP, Bastarrika G, Wisnivesky JP, Alcaide AB, Campo A, Seijo LM, et al. Assessing the relationship between lung cancer risk and emphysema detected on low-dose CT of the chest. Chest 2007;132:1932–8. https://doi.org/10.1378/chest.07-1490.
40. Soriano JB, Polverino F. The resilience of the human lungs: evolution, smoke, and disease. Am J Respir Crit Care Med 2025;211:1343–5. https://doi.org/10.1164/rccm.202504-0793VP.
41. Bruder C, Bulliard JL, Germann S, Konzelmann I, Bochud M, Leyvraz M, et al. Estimating lifetime and 10-year risk of lung cancer. Prev Med Rep 2018;11:125–30. https://doi.org/10.1016/j.pmedr.2018.06.010.
42. Demoly P, Simonylafontaine J, Chanez P, Pujol JL, Lequeux N, Michel FB, et al. Cell-proliferation in the bronchial-mucosa of asthmatics and chronic bronchitics. Am J Resp Crit Care 1994;150:214–17. https://doi.org/10.1164/ajrccm.150.1.7912988.
43. Boers JE, Ambergen AW, Thunnissen FB. Number and proliferation of basal and parabasal cells in normal human airway epithelium. Am J Resp Crit Care 1998;157:2000–6. https://doi.org/10.1164/ajrccm.157.6.9707011.
44. Xu C, Hua HJ, Chen TY, Zhang WM, Song GX, Zhang ZH. PD-L1 is correlated with p53 expression in patients with lung adenocarcinoma. Int J Clin Exp Patho 2017;10:2701–11.
45. Roeper J, Christopoulos P, Falk M, Heukamp LC, Tiemann M, Stenzinger A, et al. TP53 co-mutations as an independent prognostic factor in 2nd and further line therapy-EGFR mutated non-small cell lung cancer IV patients treated with osimertinib. Transl Lung Cancer Res 2022;11:4–13. https://doi.org/10.21037/tlcr-21-754.
46. Matas J, Kohrn B, Fredrickson J, Carter K, Yu M, Wang T, et al. Colorectal cancer is associated with the presence of cancer driver mutations in normal colon. Cancer Res 2022;82:1492–502. https://doi.org/10.1158/0008-5472.CAN-21-3607.
47. Motz GT, Eppert BL, Sun G, Wesselkamper SC, Linke MJ, Deka R, et al. Persistence of lung CD8 T cell oligoclonal expansions upon smoking cessation in a mouse model of cigarette smoke-induced emphysema. J Immunol 2008;181:8036–43. https://doi.org/10.4049/jimmunol.181.11.8036.
48. Yan Y, Tan Q, Wang Y, Wang DL, Jin M, Gordon T, et al. Enhanced lung tumor development in tobacco smoke-exposed transgenic and heterozygous deficient mice. Inhal Toxicol 2007;19:183–7. https://doi.org/10.1080/08958370701496160.
49. Provencio M, Cobo M, Rodriguez-Abreu D, Carcereny E, Castro RL, Fernández Bruno M, et al. Incidence of second primary cancers in lung cancer survivors by oncological treatment: a nationwide prospective cohort study in Spain. Lancet Reg Health Eur 2025;56:101402. https://doi.org/10.1016/j.lanepe.2025.101402.