



Smoking status and self-reported race affect the frequency of clinically relevant oncogenic alterations in non-small-cell lung cancers at a United States-based academic medical practice

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Yamaguchi, Norihiro, Paul A. VanderLaan, Erik Folch, David H. Boucher, Hannah M. Canepa, Michael S. Kent, Sidharta P. Gangadharan, et al. 2013. "Smoking Status and Self-Reported Race Affect the Frequency of Clinically Relevant Oncogenic Alterations in Non-Small-Cell Lung Cancers at a United States-Based Academic Medical Practice." <i>Lung Cancer</i> 82 (1) (October): 31–37. doi:10.1016/j.lungcan.2013.07.013.
Published Version	10.1016/j.lungcan.2013.07.013
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:36304198
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Published in final edited form as:

Lung Cancer. 2013 October ; 82(1): 31–37. doi:10.1016/j.lungcan.2013.07.013.

Smoking status and self-reported race affect the frequency of clinically-relevant oncogenic alterations in non-small-cell lung cancers at a United States-based academic medical practice

Norihiro Yamaguchi, MD, MPH¹, Paul A. VanderLaan, MD, PhD², Erik Folch, MD, MSc^{1,3}, David H. Boucher, BS¹, Hannah M. Canepa, BA¹, Michael S. Kent, MD³, Sidharta P. Gangadharan, MD³, Adnan Majid, MD^{1,3}, Olivier N. Kocher, MD, PhD², Michael A. Goldstein, MD¹, Mark S. Huberman, MD¹, and Daniel B. Costa, MD, PhD^{1,*}

¹Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

²Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

³Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

Abstract

Introduction—The identification of somatic genomic aberrations in non-small-cell lung cancer (NSCLC) is part of evidence-based practice guidelines for care of patients with NSCLC. We sought to establish the frequency and correlates of these changes in routine patient-tumor sample pairs.

Methods—Clinicopathologic data and tumor genotype were retrospectively compiled and analyzed from an overall cohort of 381 patient-tumor samples.

Results—Of these patients, 75.9% self-reported White race, 13.1% Asian, 6.5% Black, 27.8% were never-smokers, 54.9% former-smokers and 17.3% current-smokers. The frequency of *EGFR* mutations was 23.9% (86/359), *KRAS* mutations 34.2% (71/207) and *ALK* FISH positivity 9.1% (23/252) in tumor samples, and almost all had mutually exclusive results for these oncogenes. In tumors from White, Black and Asian patients, the frequencies of *EGFR* mutations were 18.4%, 18.2% and 62%, respectively; of *ALK* FISH positivity 7.81%, 0% and 14.8%, respectively; and of *KRAS* mutations 41.6%, 20% and 0%. These patterns changed significant

© 2013 Elsevier Ireland Ltd. All rights reserved.

*Correspondence to: Daniel B. Costa, MD, PhD - Division of Hematology/Oncology, Beth Israel Deaconess Medical Center, 330 Brookline Av., Boston, MA 02215 Phone: 617-667-9236, Fax: 617-975-5665, dbcosta@bidmc.harvard.edu.

Conflict of interest: DBC has received consulting fees from Pfizer, Roche and AstraZeneca. No other conflict of interest is stated.

CONFLICT OF INTEREST STATEMENT

Daniel B. Costa has received consulting fees from Pfizer, Roche and AstraZeneca. Paul A. VanderLaan, Erik Folch, David H. Boucher, Hannah M. Canepa, Michael S. Kent, Sidharta P. Gangadharan, Adnan Majid, Olivier N. Kocher, Michael A. Goldstein, Mark S. Huberman have no conflicts to disclose. No other conflict of interest is stated.

Contributors: NY and DBC were involved in the conception of this study; PAV, EF, DHB, HMC, MSK, SG, AM, ONK, MAG, MSH and DBC were involved in data acquisition, analysis and interpretation; DBC provided administrative and funding support; NY, PAV, EF and DBC were involved in writing the report; all authors approved the final version.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

with increasing pack-year history of smoking. In White patients, the frequencies of *EGFR* mutations and *ALK* FISH positivity decreased with increasing pack-year cohorts; while the frequencies of *KRAS* mutations increased. Interestingly, in Asian patients the frequencies of *EGFR* mutations were similar in never smokers and in the cohorts with less than 45 pack-year histories of smoking and only decreased in the 45 pack-year plus cohort.

Conclusions—The frequencies of somatic *EGFR*, *KRAS*, and *ALK* gene abnormalities using routine lung cancer tissue samples from our United States-based academic medical practice reflect the diverse ethnicity (with a higher frequency of *EGFR* mutations in Asian patients) and smoking patterns (with an inverse correlation between *EGFR* mutation and *ALK* rearrangement) of our tested population. These results may help other medical practices appreciate the expected results from introduction of routine tumor genotyping techniques into their day-to-day care of NSCLC.

Keywords

lung cancer; non-small-cell lung cancer; never smokers; epidermal growth factor receptor; *EGFR*; anaplastic lymphoma kinase; *ALK*; *KRAS*; tumor genotype; ethnicity; Asian; White; Black

INTRODUCTION

The most common form of lung cancer - non-small-cell lung cancer (NSCLC) – is characterized by subgroups enriched for driver oncogene aberrations. The most prevalent mutated or rearranged oncogenes identified in NSCLCs are v-ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*), epidermal growth factor receptor (*EGFR*), anaplastic lymphoma kinase (*ALK*), among others (1;2). The latter two changes, mutations in *EGFR* and rearrangements (either inversions or translocations) involving *ALK*, are robust predictive biomarkers for improved outcomes with tyrosine kinase inhibitors (TKIs), such as erlotinib and crizotinib, respectively (3–5). Interestingly, *EGFR* mutations and *ALK* rearrangements are more frequently identified in adenocarcinomas and in never and/or light smokers; while *KRAS* mutations are more common in former or current smokers (6–9). Significant variations in the frequency of these somatic tumor molecular changes are evident in different racial and ethnic groups with a diagnosis of NSCLC; with *EGFR* mutations being more prevalent in tumors from patients with self-reported Asian race and *KRAS* mutations predominating in self-reported White race (6;9).

In most United States-based academic medical practices that treat NSCLC the population of patients seen is mixed with a variety of heterogeneous racial groups and smoking habits; and the frequency of common driver oncogenes in tumors from these patients is not well defined using commercially-available genotyping techniques. In metropolitan areas of the North-East New England region of the United States, most NSCLCs are diagnosed in former or current smokers of White race; however never smokers and patients with Asian or Black/African-American race can comprise close to 20% of the diagnosed NSCLC cases (10). We sought to better elucidate how self-reported race and smoking status affected the frequency of somatic tumor genotype aberrations of NSCLCs in an academic medical practice based in Boston (Massachusetts in the United States) that sees a heterogeneous group of patients afflicted with NSCLC.

MATERIALS AND METHODS

Patient selection

Patients with a diagnosis of lung cancer, who were seen by our providers and whose tumors were genotyped for at least *EGFR* mutations were identified through an ongoing Institutional Review Board (IRB) approved protocol at Beth Israel Deaconess Medical

Center (BIDMC). Patients and tumor pairs were excluded if genotyping was not performed. There were 381 patient-tumor specimens that were submitted to a commercial vendor for tumor genotype techniques between 2007 and 2012. The data cut off for analyses was December 19th, 2012. The tumor specimens selected during routine clinical care that were retrospectively reviewed in this dataset reflect evolving tumor genotyping recommendations from 2007 to 2012 (1;2).

Tumor genotype

After a routine diagnosis of lung cancer was established on histologic and/or immunohistologic slides from surgical specimens, core needle biopsies or cell aspirates/concentrates, the residual material in the formalin-fixed paraffin-embedded (FFPE) tissue blocks were submitted for molecular analysis. *EGFR* mutation analysis was performed using standard DNA sequencing techniques and exons 18 to 21 were sequenced (11;12). For *KRAS* mutation analysis, DNA from exon 2 was amplified and subjected to single nucleotide primer extension to detect mutations at codons 12 and 13. *ALK* translocation status was analyzed using the Vysis ALK Break-Apart fluorescence in situ hybridization (FISH) probe (Abbott Molecular, Inc., Des Plaines, IL), as previously described (8).

Data collection

Clinical, pathologic, radiographic and tumor genotyping (for *EGFR* and *KRAS* mutation status and *ALK* translocation status) was collected from chart extraction. The histology of the tumor was retrospectively obtained from diagnostic pathology reports; and the incorporation of histological and immunohistochemical markers was not prospectively standardized. BIDMC's medical chart system provides details on patient's self-reported race following United States census (White, Black or African American, Asian, etc). Pack-years history of smoking was extracted from patient's self-reported smoking status. Study data were collected and managed using REDCap electronic data capture tools hosted at BIDMC.

Statistical methods

Fisher's exact test was used to compare categorical variables and Wilcoxon rank for numerical variables. All p-values reported were two-sided. Statistical analyses were performed with STATA version 12 (STATA Corp, College Station, TX).

RESULTS

Patient and tumor characteristics

The complete cohort comprised 381 patient-tumor pairs. The mean age at time of diagnosis was 65 and 61.2% were women. Self-reported racial groups in this cohort were 75.9% White, 13.1% Asian, 6.5% Black and 4.4% others or mixed racial groups. 17.3% of patients were current smokers, 54.9% were former smokers and 27.8% never smokers. Out of these lung cancers, 6.3% were staged by the 7th TNM staging system as stage I, 5.7% as stage 2, 12.1% as stage III, and 73.8% as stage IV. The latter predominance of advanced stages reflects current evidence-based guidelines to genotype advanced tumors (1;5). Most genotyped tumors had adenocarcinoma histology (86.1%), followed by NSCLC not otherwise specified (10.2%), squamous cell carcinoma (2.89%) and others (0.7%). *EGFR* mutation analysis was successful in 94.2% (359/381) of tumors, *KRAS* mutation in 91.6% (207/226) of tested tumors, and *ALK* FISH in 91.6% (252/275) of tested samples.

The overall frequency of *EGFR* mutations was 23.9% (86/359), of *KRAS* mutations 34.2% (71/207) and of *ALK* FISH positivity 9.1% (23/252) in tumor samples with successful genotype. Out of the 182 samples that had successful assays for all three tests, only 3 (1.6%) had concurrent positive results (2 cases with *ALK* FISH and *KRAS* mutations, and 1 case

with an atypical *EGFR* G724fs mutation and a *KRAS* mutation). These were excluded from further analysis. All other cases had mutually exclusive findings when one of the three aberrations was identified.

Table 1 summarizes the patient and tumor characteristics of the cases with positive findings for *EGFR* mutation, Table 2 for *ALK* FISH, and Table 3 for *KRAS* mutation. *EGFR* mutated NSCLCs were enriched in patients with a never smoking history, with non-White ethnicity and women (Table 1). *ALK* rearranged tumors were more prevalent in patients with a never smoking history and younger age (Table 2). *KRAS* mutated NSCLCs were enriched in former/current smokers and in patients with non-Asian ethnicities (Table 3).

Due to the significant interactions of *EGFR*, *ALK* and *KRAS* aberrations with baseline self-reported race and smoking history, we decided to study these factors in more detail.

Frequency of *EGFR* mutation, *KRAS* mutation and *ALK* FISH by smoking status

We subgrouped the patient-tumor samples in four distinct smoking categories: never smokers (<100 cigarettes over lifetime), more than 0 but less than or equal to 15 pack-years (1–15 pack-years), more than 15 but less than or equal to 45 pack-years (16–45 pack-years), and more than 45 pack-years (>45 pack-years). The frequency of tumor somatic mutations in these different cohorts for *EGFR* (Figure 1A), *ALK* (Figure 1B) and *KRAS* (Figure 1C) changed with increasing smoking. To analyze the effect of smoking in the frequency of these three genomic changes, we restricted our analysis to the 179 patient-tumor samples that had all three tumor genotype techniques performed, had successful assays and had no overlap in test results. Out of the 50 never smokers, *EGFR* mutations were present in 46% of tumors, *ALK* FISH positivity was seen in 14%, *KRAS* mutations were present in 6%, and 34% were “triple negative” for these gene mutations (Figure 1D). The tumor frequencies changed to *EGFR* 5%, *ALK* FISH 2% and *KRAS* 43% in the 129 patients with a smoking history (Figure 1D). The differences in distribution between all these smoking cohorts were significant ($p < 0.0001$), indicating that smoking history clearly affects the frequency of *EGFR* mutation, *ALK* FISH positivity and *KRAS* mutation in a given NSCLC population.

Out of all *KRAS* mutations identified, the most common were transversion mutations (G–T or G–C) including G12C (34 cases), G12V (17 cases), G12R (3 cases), G12A (2 cases) and G13C (1 case); while transition mutations (G–A) were less common (11 cases G12D and 2 cases G13D). In never smokers, 1/3 (25%) *KRAS* mutations were transition while in smokers 12/66 (18.1%) were transition ($p = 0.57$).

Frequency of *EGFR* mutation, *KRAS* mutation and *ALK* FISH by self-reported race

To analyze the effect of race on frequency of these tumor somatic genetic alterations, we subdivided our patients into their self-reported racial categories (Figure 2).

From the 272 White patients with a successful tumor *EGFR* analysis, the frequency of *EGFR* mutations was 18.4% (50/272). Out of the 50 Asian patients, the frequency was significantly higher at 62% (31/50; $p < 0.0001$ when compared to White). Out of the 22 Black patients, the frequency was 18.2% (4/22; $p = 1.0$ when compared to White).

From the 192 White patients with a successful tumor *ALK* FISH analysis, the frequency of *ALK* rearrangement was 7.81% (15/192). Out of the 27 Asian patients, the frequency was 14.8% (4/27; $p = 0.27$ when compared to White). Out of the 21 Black patients, the frequency was 0% (0/21; $p = 0.37$ when compared to White).

From the 161 White patients with a positive tumor *KRAS* analysis, the frequency of *KRAS* mutation was 41.6% (67/161). Out of the 21 Asian patients, the frequency was significantly

lower at 0% (0/21; $p < 0.001$ when compared to White). Out of the 21 Black patients, the frequency was again lower at 20% (3/15; $p = 0.16$ when compared to White).

In White patients, the frequencies of *EGFR* mutations decreased with increasing pack-year cohorts (Figure 2A). The frequencies of *ALK* FISH positivity also decreased with increasing pack-year cohorts (Figure 2B). The frequencies of *KRAS* mutations increased from never smokers to ever smokers (Figure 2C).

In Black patients, the frequencies of *EGFR* mutations (75% never smokers [all women], 0% 1–15 pack-years, 16.6% [20% women and 0% men, $p = 1$] 16–45 pack-years and 0% >45 pack-years) also showed a trend to decrease numerically with increasing pack-years of smoking (Figure 2A). *KRAS* mutations were only seen in the 16–45 pack-years (2/5, 40% [50% in women and 0% in men; $p = 1.0$] and >45 pack-year cohorts (1/6, 16.6% [all men]).

Interestingly, in Asian patients the frequencies of *EGFR* mutations were similar in never smokers and in the cohorts with less than 45 pack-year histories of smoking (Figure 2A). The frequencies of *EGFR* mutations were 63.6% (69.2% in women and 50% in men; $p = 0.39$) never smokers, 60% (all men) 1–15 pack-years, 71.4% (all men) 16–45 pack-years (Figure 2A). This high frequency only decreased to 33.3% (all men) in the 45 pack-year plus cohort (Figure 2A).

Out of the patient-tumor samples in which the patient reported the racial category as other, 11 patients self-reported Hispanic/Latino ethnicity. Out of these cases, the frequencies of *EGFR* mutations were 9.1% (1/11; $p = 0.70$ when compared to White), of *ALK* FISH positivity 33.3% (3/9; $p = 0.036$ when compared to White) and of *KRAS* mutations 0% (0/9; $p = 0.012$ when compared to White).

DISCUSSION

The frequencies of *EGFR*, *KRAS* and *ALK* gene aberrations in lung adenocarcinoma tumor samples are known to be associated with self-reported race, smoking status and sex (2). We performed a comprehensive analysis of a heterogeneous population of patients with predominantly advanced NSCLC who had tumor genotype performed as part of their routine clinical care, and confirmed the association of race and smoking status with the frequency of the aforementioned driver oncogene changes.

EGFR mutated NSCLCs have been known to be enriched in Asian populations, never smokers and women since the original reports of these mutations first occurred in 2004–2005 (6;7;13;14). In European-based practices, with mostly White patients, the overall frequency of *EGFR* mutations in lung adenocarcinoma samples is below 15–20% (15); while the frequency in Asian-based practices exceeds 35–45% (6;16). Most United States-based practices see patients of multiple self-reported racial categories. In one of the largest series of tumors sequenced for *EGFR* mutations the Thoracic Oncology service at Memorial Sloan-Kettering Cancer Center (MSKCC) reported that the frequency of tumor *EGFR* mutations was 17.5% in 2736 White patients, 55.1% in 136 Asian/Pacific patients and in 20.8% in 77 Black patients (17). Our own data corroborates these observations and the distribution of patient ethnicities that had their NSCLC tumors genotyped. The consistency of these results highlights the expected frequency of *EGFR* mutations by self-reported race. The effect of smoking history is the main determinant of the frequency of *EGFR* mutations in most NSCLC populations (7;18). In United States-based centers, *EGFR* mutated tumors comprise around 40% of never smokers (19) while *EGFR* mutated tumors comprise more than 55–70% of Asian-based NSCLC never smoker cohorts (18;20). The frequency of *EGFR* mutated NSCLC, in mostly White patients, is known to decreased with increasing pack-years of smoking (17). As an example, in the aforementioned MSKCC data of 3026

patients, the frequency of *EGFR* mutations changed from 42.5% in never smokers to below 8.5% in patients with more than 26 pack-years (17). Interestingly, even cohorts with 1–10 pack-years had frequencies of *EGFR* mutations below 28% (17;21). Our data with White and Black patients is similar to these numbers. We observed in our White patients with NSCLC a decreasing frequency with increasing smoking dose (from ~35% in never smokers, to around 20% in 1–15 pack-years, to ~15% in 16–45 pack-years to below 7% in the cohort with >45 pack-years). In Asian patients, the effects of smoking on decreasing the frequency of *EGFR* mutations were only noticeable in the >45 pack-year smoking cohort; and the number of *EGFR* mutated NSCLCs exceed 60% in all other groups. Although female sex is more frequent than male sex in patients with *EGFR* mutated NSCLC, this association is linked to smoking status in many Asian countries (18). Indeed, in our series none of the Asian women had smoking histories. In both our Asian and White patient populations, *EGFR* mutations were numerically more frequent in women than men. Female sex, for unknown reasons, is known to be one of the most robust risk factors in case controls studies of the risk of *EGFR* mutated NSCLC (22).

ALK gene rearrangements also have been known to be enriched in never smokers since the original reports of these inversions or translocations first occurred in 2007–2009 (8;23). In both Asian and White populations, the overall frequency of *ALK* rearrangements is reported to be around 5% (24;25). However, the frequency can be as high as 20% in never smokers with lung adenocarcinomas (4;8;26). An effect of race on the frequency of *ALK* rearrangements has not been routinely reported (24;25). Our cohort corroborates these prior observations. Near 18% of Asian and White NSCLCs in our series harbored *ALK* rearrangements. The frequency of *ALK* FISH positivity decreased with incremental pack-years history of smoking (~20% in never smokers to <2% in the cohort with >45 pack-years). Our small number of cases precludes conclusions regarding the true frequencies of *ALK* rearrangements in Black patients and in patients of Hispanic/Latino ethnicity with NSCLC.

As opposed to *EGFR* mutations and *ALK* rearrangements, *KRAS* mutations are known to be associated with increasing smoking histories (9). In the largest published series of *KRAS* genotyped NSCLCs (2529 cases), the frequency of *KRAS* mutations increased from 6.4% in never smokers to near 40% in tumors from patients with more than 16 pack-years history of smoking (17). Our series closely mirrors these observations. In White patients, *KRAS* mutated NSCLCs accounted for ~10% of never smokers with NSCLC but increased to near 45–50% of tumors in smoking cohorts. *KRAS* mutations in smokers are more frequently transversion mutations (G–C or G–T), which are though to be a marker of a smoking carcinogen signature (17). In our series, transversion *KRAS* mutations accounted for over 80% of mutations in smokers. *KRAS* mutations are not common in Asian patients, even those with significant smoking histories (9). The reason(s) for the lack of *KRAS* mutated NSCLC in Asians with NSCLC are not clear. In our series, none of our Asian patients genotyped for this oncogene had *KRAS* mutations.

The effects and associations of race, sex and smoking history on the frequencies of *EGFR*, *KRAS* and *ALK* gene aberrations in lung adenocarcinoma confirmed in this report have important clinical implications. In *EGFR* mutated NSCLCs with classic *EGFR* exon 19 deletions or the L858R mutation, clinical responses can be achieved in more than 70% of patients with the use of the reversible EGFR TKIs gefitinib or erlotinib (2); and the evidence-based use of these drugs is now restricted to *EGFR* mutated NSCLCs in the first line treatment of advanced tumors (5;27). In *ALK* rearranged NSCLCs, the multitargeted ALK TKI crizotinib also leads to clinical responses in most patients; and the evidence-based use of this drug is only approved for metastatic *ALK* rearranged NSCLCs (4;24;28). Although *KRAS* mutated NSCLCs don't have a targetable inhibitor to date, recent advances

have shown promise for the use of cytotoxic chemotherapies with MEK inhibitors – among other strategies – for these recalcitrant tumors (29;30). And, in addition, since *KRAS* mutations are usually mutually exclusive with other driver genetic changes, the identification of *KRAS* mutated NSCLC may help define a subset of NSCLCs that may not require additional tumor genotyping for novel driver actionable oncogenes (31).

In never smokers with lung adenocarcinomas, most tumors will have *EGFR* mutations or *ALK* rearrangements (>80% of Asian never smoker-related NSCLCs, and in >60% of White/Black never smoker-related NSCLCs) and will be able to reap the palliative benefits of approved oral TKIs against these driver oncogenes. Since never smokers seem to be groups that are also enriched for other driver oncogenic aberrations that in preclinical models can be inhibited by currently available or in development TKIs (such as *ROS1* rearrangements, *ERBB2* mutations, *RET* rearrangements and *BRAF* mutations) it is expected that additional biomarker-based antineoplastics will be approved for nearly all never smokers with advanced NSCLC (31;32). Never smokers with NSCLC may be an ideal group to have their tumors genotyped using targeted next-generation sequencing to obtain somatic mutational events, rearrangements and copy number changes (33). However, it is clear that many of the aforementioned mutations or gene rearrangements also occur in smokers and it is essential to genotype all lung adenocarcinomas (and maybe all NSCLCs) to guarantee complete identification of all cases with a clinically-relevant driver oncogene. In patients with a significant history of smoking and NSCLC, the frequency of *KRAS* mutated tumors is exorbitantly high in White patients (in our report close to 50%); highlighting both the urgent unmet need of identifying palliative therapeutic strategies that can target *KRAS* mutations and the need to increase the number of clinical trials that are specific for *KRAS* mutated NSCLC (30).

In summary, we confirm prior observations that the frequencies of somatic *EGFR*, *KRAS*, and *ALK* gene abnormalities from United States-based academic medical practices reflect the diverse ethnicity and smoking patterns of the tested population. The knowledge of these frequencies may help other medical practices appreciate the expected results from introduction of routine tumor genotyping techniques into their day-to-day care of NSCLC.

Acknowledgments

We would like to thank all current and former members of the Thoracic Oncology Clinic at Beth Israel Deaconess Medical Center, and our patients.

This work was funded in part through a fellowship from the American Society of Clinical Oncology Conquer Cancer Foundation (DBC), an American Cancer Society grant (RSG 11-186 to DBC), a Lung Cancer Foundation of America-International Association for the Study of Lung Cancer grant (to DBC), and National Institutes of Health (NIH) grant CA090578 (to DBC).

Funding/Grant Support: This work was funded in part through a fellowship from the American Society of Clinical Oncology Conquer Cancer Foundation (DBC), an American Cancer Society grant (RSG 11-186 to DBC), a Lung Cancer Foundation of America-International Association for the Study of Lung Cancer grant (to DBC), and National Institutes of Health (NIH) grant CA090578 (to DBC).

References

1. Cheng L, Alexander RE, Maclennan GT, Cummings OW, Montironi R, Lopez-Beltran A, et al. Molecular pathology of lung cancer: key to personalized medicine. *Mod Pathol*. 2012; 25(3):347–369. [PubMed: 22282308]
2. Gaughan EM, Costa DB. Genotype-driven therapies for non-small cell lung cancer: focus on EGFR, KRAS and ALK gene abnormalities. *Ther Adv Med Oncol*. 2011; 3(3):113–125. [PubMed: 21904575]

3. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* 2012; 13(3):239–246. [PubMed: 22285168]
4. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med.* 2010; 363(18):1693–1703. [PubMed: 20979469]
5. Keedy VL, Temin S, Somerfield MR, Beasley MB, Johnson DH, McShane LM, et al. American Society of Clinical Oncology provisional clinical opinion: epidermal growth factor receptor (EGFR) Mutation testing for patients with advanced non-small-cell lung cancer considering first-line EGFR tyrosine kinase inhibitor therapy. *J Clin Oncol.* 2011; 29(15):2121–2127. [PubMed: 21482992]
6. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst.* 2005; 97(5):339–346. [PubMed: 15741570]
7. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A.* 2004; 101(36):13306–13311. [PubMed: 15329413]
8. Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J Clin Oncol.* 2009; 27(26):4247–4253. [PubMed: 19667264]
9. Riely GJ, Marks J, Pao W. KRAS mutations in non-small cell lung cancer. *Proc Am Thorac Soc.* 2009; 6(2):201–205. [PubMed: 19349489]
10. Nguyen KS, Sanford RA, Huberman MS, Goldstein MA, McDonald DM, Farquhar M, et al. Patterns of care for non-small-cell lung cancer at an academic institution affiliated with a national cancer institute-designated cancer center. *J Oncol Pract.* 2012; 8(1):57–62. [PubMed: 22548013]
11. Jackman DM, Miller VA, Cioffredi LA, Yeap BY, Janne PA, Riely GJ, et al. Impact of epidermal growth factor receptor and KRAS mutations on clinical outcomes in previously untreated non-small cell lung cancer patients: results of an online tumor registry of clinical trials. *Clin Cancer Res.* 2009; 15(16):5267–5273. [PubMed: 19671843]
12. Costa DB, Nguyen KS, Cho BC, Sequist LV, Jackman DM, Riely GJ, et al. Effects of erlotinib in EGFR mutated non-small cell lung cancers with resistance to gefitinib. *Clin Cancer Res.* 2008; 14(21):7060–7067. [PubMed: 18981003]
13. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med.* 2004; 350(21):2129–2139. [PubMed: 15118073]
14. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science.* 2004; 304(5676):1497–1500. [PubMed: 15118125]
15. Rosell R, Moran T, Queralt C, Porta R, Cardenal F, Camps C, et al. Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med.* 2009; 361(10):958–967. [PubMed: 19692684]
16. Mitsudomi T, Yatabe Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci.* 2007; 98(12):1817–1824. [PubMed: 17888036]
17. Dogan S, Shen R, Ang DC, Johnson ML, D’Angelo SP, Paik PK, et al. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin Cancer Res.* 2012; 18(22):6169–6177. [PubMed: 23014527]
18. Li C, Fang R, Sun Y, Han X, Li F, Gao B, et al. Spectrum of oncogenic driver mutations in lung adenocarcinomas from East asian never smokers. *PLoS One.* 2011; 6(11):e28204. [PubMed: 22140546]
19. Janne PA, Wang X, Socinski MA, Crawford J, Stinchcombe TE, Gu L, et al. Randomized phase II trial of erlotinib alone or with carboplatin and paclitaxel in patients who were never or light former

- smokers with advanced lung adenocarcinoma: CALGB 30406 trial. *J Clin Oncol.* 2012; 30(17): 2063–2069. [PubMed: 22547605]
20. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med.* 2009; 361(10):947–957. [PubMed: 19692680]
 21. Varghese AM, Sima CS, Chaft JE, Johnson ML, Riely GJ, Ladanyi M, et al. Lungs don't forget: Comparison of the KRAS and EGFR mutation profile and survival of collegiate smokers and never smokers with advanced lung cancers. *J Thorac Oncol.* 2013; 8(1):123–125. [PubMed: 23242442]
 22. Matsuo K, Ito H, Yatabe Y, Hiraki A, Hirose K, Wakai K, et al. Risk factors differ for non-small-cell lung cancers with and without EGFR mutation: assessment of smoking and sex by a case-control study in Japanese. *Cancer Sci.* 2007; 98(1):96–101. [PubMed: 17054433]
 23. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature.* 2007; 448(7153): 561–566. [PubMed: 17625570]
 24. Ou SH, Bartlett CH, Mino-Kenudson M, Cui J, Iafrate AJ. Crizotinib for the treatment of ALK-rearranged non-small cell lung cancer: a success story to usher in the second decade of molecular targeted therapy in oncology. *Oncologist.* 2012; 17(11):1351–1375. [PubMed: 22989574]
 25. Shaw AT, Engelman JA. ALK in Lung Cancer: Past, Present, and Future. *J Clin Oncol.* 2013; 31(8):1105–1111. [PubMed: 23401436]
 26. Cardarella S, Ortiz TM, Joshi VA, Butaney M, Jackman DM, Kwiatkowski DJ, et al. The introduction of systematic genomic testing for patients with non-small-cell lung cancer. *J Thorac Oncol.* 2012; 7(12):1767–1774. [PubMed: 23154547]
 27. Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, et al. Non-small cell lung cancer. *J Natl Compr Canc Netw.* 2012; 10(10):1236–1271. [PubMed: 23054877]
 28. Camidge DR, Doebele RC. Treating ALK-positive lung cancer—early successes and future challenges. *Nat Rev Clin Oncol.* 2012; 9(5):268–277. [PubMed: 22473102]
 29. Chen Z, Cheng K, Walton Z, Wang Y, Ebi H, Shimamura T, et al. A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. *Nature.* 2012; 483(7391):613–617. [PubMed: 22425996]
 30. Janne PA, Shaw AT, Pereira JR, Jeannin G, Vansteenkiste J, Barrios C, et al. Selumetinib plus docetaxel for KRAS-mutant advanced non-small-cell lung cancer: a randomised, multicentre, placebo-controlled, phase 2 study. *Lancet Oncol.* 2013; 14(1):38–47. [PubMed: 23200175]
 31. Oxnard GR, Binder A, Janne PA. New targetable oncogenes in non-small-cell lung cancer. *J Clin Oncol.* 2013; 31(8):1097–1104. [PubMed: 23401445]
 32. Yasuda H, Figueiredo-Pontes LL, Kobayashi S, Costa DB. Preclinical Rationale for Use of the Clinically Available Multitargeted Tyrosine Kinase Inhibitor Crizotinib in ROS1-Translocated Lung Cancer. *J Thorac Oncol.* 2012; 7(7):1086–1090. [PubMed: 22617245]
 33. Wagle N, Berger MF, Davis MJ, Blumenstiel B, Defelice M, Pochanard P, et al. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov.* 2012; 2(1):82–93. [PubMed: 22585170]

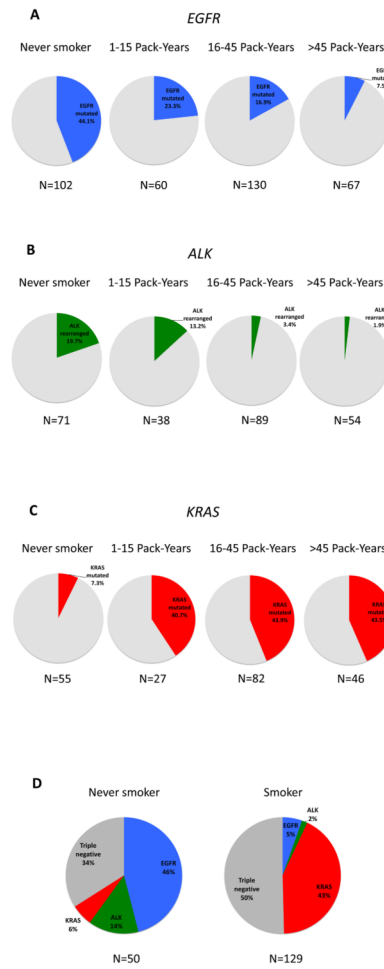


FIGURE 1. Frequencies of *EGFR* mutations (A), *ALK* FISH positivity (B) and *KRAS* mutations within increasing cohorts of smoking history (from never smokers to >45 pack-years history of smoking). Concurrent frequency of *EGFR* mutations, *ALK* FISH positivity and *KRAS* mutations in never smokers and smokers (D) successfully tested for all three tumor genotypes.

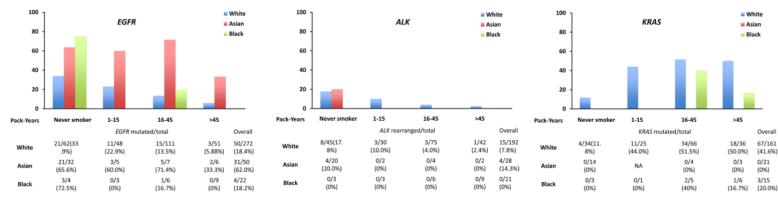


FIGURE 2. Frequencies of *EGFR* mutations (A), *ALK* FISH positivity (B) and *KRAS* mutations (C) within increasing cohorts of smoking history (from never smokers to >45 pack-years history of smoking) by self-reported race (White, Asian or Black).

Table 1Baseline characteristics of patients and tumors successfully tested for *EGFR* mutations.

	<i>EGFR</i> tested (n=359)		
	<i>EGFR</i> mutated (n=86)	<i>EGFR</i> WT (n=273)	<i>p</i> -value
Age at the time of biopsy median(range)	65(33–90)	65(29–88)	1.00
Women n (%)	64(74.4)	158 (71.2)	0.006
Race n (%)			
White	50(58.1)	222(81.3)	Ref.
Asian	31(36.1)	18(6.59)	<0.0001
Black	4(4.65)	10(3.66)	<0.0001
Other	1(1.16)	19(6.96)	<0.0001
Smoking status n (%)			
Current smoker	4(4.65)	58(21.3)	<0.0001
Former smoker	37(43.0)	158(57.9)	<0.0001
Never smoker	45(52.3)	57(20.9)	Ref.
Stage n (%)			
I	3(3.49)	20(7.33)	0.21
II	6(6.98)	16(5.86)	1.00
III	7(8.14)	38(13.9)	0.14
IV	70(81.4)	190(72.9)	Ref.
Histology n (%)			
Adenocarcinoma	84(97.7)	226(82.8)	Ref.
Squamous cell carcinoma	0(0)	10(3.66)	0.07
NSCLC (NOS)	2(2.33)	35(12.8)	0.002
Others	0(0)	2(0.73)	1.00

NSCLC, non-small-cell lung cancer; NOS, not otherwise specified.

Table 2Baseline characteristics of patients and tumors successfully tested for *ALK* FISH.

	<i>ALK</i> tested (n=252)		
	<i>ALK</i> rearranged (n=23)	<i>ALK</i> FISH negative (n=229)	<i>p</i> -value
Age at the time of biopsy median(range)	56(29–80)	65(33–90)	0.0003
Women n (%)	11(47.8)	143(62.5)	0.18
Race n (%)			
White	15(65.2)	177(77.3)	Ref.
Asian	4(17.4)	24(10.5)	0.28
Black	0(0)	21(9.17)	0.37
Other	4(17.4)	7(3.06)	0.01
Smoking status n (%)			
Current smoker	1(4.35)	52(22.7)	0.002
Former smoker	8(34.8)	120(52.4)	0.42
Never smoker	14(60.9)	57(24.9)	Ref.
Stage n (%)			
I	1(4.35)	17(7.42)	1.00
II	0(0)	12(5.24)	0.61
III	2(8.70)	27(11.8)	0.75
IV	20(87.0)	173(75.5)	Ref.
Histology n (%)			
Adenocarcinoma	21(91.3)	190(83.0)	Ref.
Squamous cell carcinoma	1(4.35)	7(3.06)	0.58
NSCLC (NOS)	1(4.35)	29(12.7)	0.33
Others	0(0)	3(1.31)	1.00

NSCLC, non-small-cell lung cancer; NOS, not otherwise specified.

Table 3Baseline characteristics of patients and tumors successfully tested for *KRAS* mutations.

	<i>KRAS</i> tested (n=207)		
	<i>KRAS</i> mutated (n=71)	<i>KRAS</i> WT (n=136)	<i>p</i> -value
Age at the time of biopsy median(range)	65(45–88)	65(29–89)	0.0003
Women n (%)	49(69.0)	83(61.0)	0.29
Race n (%)			
White	67(94.4)	94(69.1)	Ref.
Asian	0(0)	21(15.4)	0.0001
Black	3(4.23)	12(8.82)	0.17
Other	1(1.41)	9(6.68)	0.053
Smoking status n (%)			
Current smoker	20(28.2)	19(14.0)	0.0001
Former smoker	47(66.2)	66(48.5)	0.0001
Never smoker	4(5.63)	51(37.5)	Ref.
Stage n (%)			
I	6(8.45)	7(5.15)	0.36
II	5(7.04)	8(5.88)	0.76
III	10(14.1)	16(11.8)	0.65
IV	50(70.4)	105(77.2)	Ref.
Histology n (%)			
Adenocarcinoma	65(91.6)	108(79.4)	Ref.
Squamous cell carcinoma	0(0)	9(6.62)	0.03
NSCLC (NOS)	6(8.45)	17(12.5)	0.36
Others	0(0)	2(1.47)	0.53

NSCLC, non-small-cell lung cancer; NOS, not otherwise specified.