Somatic Mutation Spectrum of Non–Small-Cell Lung Cancer in African Americans

A Pooled Analysis

Luiz H. Araujo, MD, MSc, Philip E. Lammers, MD, MSc,† Velmalia Matthews-Smith, MD,‡ Rosana Eisenberg, MD,§ Adriana Gonzalez, MD,‖ Ann G. Schwartz, PhD, MPH,¶ Cynthia Timmers, PhD,* Konstantin Shilo, MD,* Weiqiang Zhao, MD, PhD,* Thanemozhi G. Natarajan, PhD,# Jianying Zhang, PhD,†† Ayse Selen Yilmaz, MSc,*†††‡‡ Tom Liu, PhD,* Kevin Coombes, PhD,*†† and David P. Carbone, MD, PhD**

Introduction: The mutational profile of non–small-cell lung cancer (NSCLC) has become an important tool in tailoring therapy to patients, with clear differences according to the population of origin. African Americans (AAs) have higher lung cancer incidence and mortality than Caucasians, yet discrepant results have been reported regarding the frequency of somatic driver mutations. We hypothesized that NSCLC has a distinct mutational profile in this group.

Methods: We collected NSCLC samples resected from self-reported AAs in five sites from Tennessee, Michigan, and Ohio. Gene mutations were assessed by either SNaPshot or next generation sequencing, and *ALK* translocations were evaluated by fluorescence in situ hybridization.

Results: Two hundred sixty patients were included, mostly males (62.3%) and smokers (86.6%). Eighty-one samples (31.2%) were squamous cell carcinomas. The most frequently mutated genes were *KRAS* (15.4%), *epidermal growth factor receptor* (*EGFR*, 5.0%), *PIK3CA* (0.8%), *BRAF*, *NRAS*, *ERBB2*, and *AKT1* (0.4% each). *ALK* translocations were detected in two nonsquamous tumors (1.7%), totaling 61 cases (23.5%) with driver oncogenic alterations. Among 179 nonsquamous samples, 54 (30.2%) presented a driver alteration. The frequency of driver alterations altogether was lower than that reported in Caucasians, whereas no difference was detected in either *EGFR* or *KRAS* mutations. Overall survival was longer among patients with *EGFR* mutations.

Conclusions: We demonstrated that NSCLC from AAs has a different pattern of somatic driver mutations than from Caucasians. The majority of driver alterations in this group are yet to be described, which will require more comprehensive panels and assessment of noncanonical alterations.

Key Words: Lung neoplasms, Mutations, EGFR, African Americans, Ethnic groups.

(*J Thorac Oncol.* 2015;10: 1430–1436)

*Department of Internal Medicine, James Thoracic Center, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio; †Department of Internal Medicine, Meharry Medical College, Nashville, Tennessee; ‡Department of Internal Medicine, Baptist Center for Cancer Care, Columbus, Mississippi; §Department of Pathology, Vanderbilt-Ingram Cancer Center, Nashville, Tennessee; ‖Pathology Associates of Saint Thomas, Nashville, Tennessee; ¶Department of Internal Medicine, Karmanos Cancer Institute, Wayne State University, Detroit, Michigan; #GenomOncology, Cleveland, Ohio; ††Department of Biomedical Informatics, The Ohio State University, Columbus, Ohio; and ‡‡Department of Biomedical Informatics, Biomedical Informatics Shared Resource, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio.

Luiz H. Araujo and Philip E. Lammers contributed equally.

Disclosure: L.H.A. is supported by a Conquer Cancer Foundation of ASCO Long-Term International Fellowship (LIFe) and a Landon Foundation-AACR INNOVATOR Award for International Collaboration in Cancer Research. P.E.L. is supported by a NIH K12CA90625. This work was funded by a NIH/NCI 1RC1 CA146260-01, NCI R01CA60691, NCI R01CA87895, NCI P30CA022453, and the Ohio State Cancer Center Support Grant (CCSG), NCI CA16058. All other authors declare no conflict of interest.

Presented in part at the 2012 and 2014 Annual Meetings of the American Society of Clinical Oncology, Chicago, IL.

Address for correspondence: David P. Carbone, MD, PhD, James Thoracic Center, Department of Medicine, The Ohio State University Medical Center, 460 W 12th Ave, room 488, Columbus, OH 43210. E-mail: david.carbone@osumc.edu.

DOI: 10.1097/JTO.0000000000000650

ISSN: 1556-0864/15/1010-1430

The mutational profile of non–small-cell lung cancer (NSCLC) has become an important tool in tailoring therapy to patients.[1] Currently, several somatic mutations or rearrangements can be successfully targeted with specific agents such as tyrosine kinase inhibitors of the epidermal growth factor receptor (*EGFR*) and anaplastic lymphoma kinase (*ALK*) inhibitors.[2,3]

The prevalence of classic driver mutations may vary according to clinical and tumor characteristics—age, gender, smoking status, and histology—and has been thoroughly studied in large populations around the world.[2,4–6] For instance, while *EGFR* sensitizing mutations are found in approximately 10% of North American and European NSCLC patients, the frequency has been reported to be as high as 50% to 60% in East Asia.[6,7] On the other hand, the frequency of *EGF*R mutations in subsets of populations has been less extensively studied, particularly in African Americans (AAs).

AAs have a higher incidence of lung cancer and tend to have poorer outcomes,[8] emphasizing the importance to

confirm targets for therapy in this high-risk population. Prior smaller studies have found discrepant frequencies of *EGFR* mutations in this group, ranging from 2% to 19%.[9–15] In a recent publication by Bollig-Fischer et al.,[16] *EGFR* exon 19 deletions were only detected in females, with a trend toward higher frequency in samples derived from AAs than from Caucasians. In addition, the authors observed a lower frequency of overall mutations among AAs compared with Caucasians, although important alterations like *ALK* rearrangements were not assessed.[16]

In this article, we describe the mutational spectrum of NSCLC in a pooled data set from AA patients, combining multiplex sequencing strategies and *ALK* testing. Our hypothesis was that NSCLC diagnosed in AAs has a different mutational profile than NSCLC diagnosed in Caucasians in North America.

PATIENTS AND METHODS

Patient Selection

Self-reported AA patients diagnosed with any stage NSCLC between 1988 and 2011 were retrospectively identified from three sites in Tennessee, Vanderbilt-Ingram Cancer Center (64 patients), Nashville General Hospital at Meharry Medical College (41 patients), and West Tennessee Healthcare (27 patients); one site in Michigan, Karmanos Cancer Institute, Wayne State University (29 patients); and one site in Ohio, Ohio State University (99 patients). Twenty samples from Wayne State overlap with the prior publication by Bollig-Fischer et al.[16] The patient search was based on tissue availability, and the clinical data correspond to available data from routine care delivered and documented in the medical records at each institution. Clinical characteristics were annotated, and tumor samples were evaluated for known recurrent driver mutations. For comparison purposes, histological subtypes were classified as either squamous (pure squamous cell carcinoma) or nonsquamous (including all other subtypes). Study data were managed using REDCap electronic data capture tools hosted at Vanderbilt. To compare the data of AA patient with that of Caucasians diagnosed with NSCLC, we analyzed the lung adenocarcinoma[17] and lung squamous cell carcinoma[18] data sets from The Cancer Genome Atlas (TCGA).[19] Clinical and mutational data from TCGA were downloaded from the publicly available repository and filtered to exclusively include cases reported as "White." Data were then obtained from this cohort for the same driver mutations tested in the panel used in this work, as described below. The institutional review board at each participating center approved this project and waived the need for informed consent.

Tumor Genotyping

After routine pathology review, unstained slides were cut from formalin-fixed paraffin-embedded tissues and used for DNA extraction and sequencing. Classic driver mutations in *EGFR*, *KRAS*, *BRAF*, *NRAS*, *AKT1*, *PIK3CA*, *PTEN*, *ERBB2*, and *MEK1* were evaluated by either SNaPshot/sizing assays in 161 patients (sites in TN and MI) or by next generation sequencing (NGS) in 99 samples (site in OH). SNaPshot/sizing assays were performed in a Clinical Laboratory Improvement Amendments-certified laboratory (Vanderbilt-Ingram Cancer Center) and followed standard protocols for sample selection and genotyping.[20] SNaPshot technique involved multiplex polymerase chain reaction, single-base extension, and analysis by capillary electrophoresis.[20] Sizing assays were performed using established protocol to detect *EGFR* exon 19 deletions, *EGFR* exon 20 insertions, and *HER2* exon 20 insertions.[20,21] The NGS platform consisted of targeted gene sequencing using a NSCLC-specific panel, performed at Ohio State University. Libraries were constructed using the Agilent Haloplex kit (Agilent Technologies, Santa Clara, CA) and sequenced on an Illumina HiSeq 2500 (Illumina, San Diego, CA). To enable a combined analysis with SNaPshot, NGS results were filtered to exclusively include the same base substitutions and insertions/deletions covered by that panel, as summarized in Supplemental Table 1 (Supplemental Digital Content 1, http://links.lww.com/JTO/A866). The mutations evaluated herein have been characteristically defined as somatic events in NSCLC. Because no paired normal tissue was used, these mutations were considered to be somatic. More NGS methodological details and sequencing metrics are provided in Supplemental Methods (Supplemental Digital Content 2, http://links.lww.com/JTO/A867). *ALK* translocations were evaluated in nonsquamous tumors (whenever adequate tissue was available after NGS/SNaPshot), using the Vysis ALK break apart fluorescence in situ hybridization (FISH) probe kit (Abbott, Abbott Park, IL).[22]

Statistical Methods

Fisher's exact test was used to compare the proportion of mutations between the two ethnic groups or between any two subgroups defined by clinical characteristics such as gender, smoking status, stage, and histology. Wilcoxon rank sum test was used to compare median age between any two ethnic (Caucasian versus AA) or mutant groups (mutant versus wild type). In multivariate analysis, logistic linear regression models were performed to adjust the mutation rate of *KRAS*/*EGFR*/Any driver for all other covariates mentioned above. To minimize the selection bias due to confounding clinical covariates, propensity score matching was performed on AAs and Caucasians by the optimal method, and the propensity score was obtained from the logistic regression of the binary response of AA or Caucasian on five covariates: age, gender, smoking, stage, and histology. The *KRAS*/*EGFR*/Any driver mutation rate was compared between the two matched ethnic groups. Survival curves were estimated using the Kaplan–Meier method and compared according to clinical covariates and molecular markers using log-rank test and Cox proportional hazards model (for hazard ratio estimate). *P* values of less than or equal to 0.05 were considered statistically significant. Given the exploratory nature of the current study, *p* values were not adjusted for multiple testing. Statistical analyses were performed using the statistical software R version 3.0.1 (R Foundation for Statistical Computing, Vienna, Austria), SAS 9.3 (SAS, Cary, NC), and IBM SPSS version 22.0 (IBM, Armonk, NY).