Genomic Characterization of Non–Small-Cell Lung Cancer in African Americans by Targeted Massively Parallel Sequencing

Luiz H. Araujo, Cynthia Timmers, Erica Hlavin Bell, Konstantin Shilo, Philip E. Lammers, Weiqiang Zhao, Thanmozhi G. Natarajan, Clinton J. Miller, Jianying Zhang, Ayse S. Yilmaz, Tom Liu, Kevin Coombes, Joseph Amann, and David P. Carbone

Luiz H. Araujo, Cynthia Timmers, Erica Hlavin Bell, Konstantin Shilo, Weiqiang Zhao, Jianying Zhang, Ayse S. Yilmaz, Tom Liu, Kevin Coombes, Joseph Amann, and David P. Carbone, The Ohio State University Comprehensive Cancer Center, Columbus; Thanmozhi G. Natarajan and Clinton J. Miller, GenomOncology, Cleveland, OH; Philip E. Lammers, Meharry Medical College, Nashville, TN.

Published online ahead of print at www.jco.org on April 27, 2015.

Supported by a Long-Term International Fellowship from the Conquer Cancer Foundation of the American Society of Clinical Oncology (L.H.A.), a Landon Foundation–American Associate for Cancer Research INNOVATOR Award for International Collaboration in Cancer Research (L.H.A.), National Institutes of Health (NIH) Grant No. K12CA90625 (P.E.L.), NIH/National Cancer Institute (NCI) Grant No. 1RC1 CA146260-01, and Ohio State Cancer Center Support Grant No. (CCSG) NCI CA16058.

Presented in part at the 50th Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, May 30–June 3, 2014.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: 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.

© 2015 by American Society of Clinical Oncology

0732-183X/15/3317w-1966w/\$20.00

DOI: 10.1200/JCO.2014.59.2444

ABSTRACT

Purpose

Technologic advances have enabled the comprehensive analysis of genetic perturbations in non–small-cell lung cancer (NSCLC); however, African Americans have often been underrepresented in these studies. This ethnic group has higher lung cancer incidence and mortality rates, and some studies have suggested a lower incidence of epidermal growth factor receptor mutations. Herein, we report the most in-depth molecular profile of NSCLC in African Americans to date.

Methods

A custom panel was designed to cover the coding regions of 81 NSCLC-related genes and 40 ancestry-informative markers. Clinical samples were sequenced on a massively parallel sequencing instrument, and anaplastic lymphoma kinase translocation was evaluated by fluorescent in situ hybridization.

Results

The study cohort included 99 patients (61% males, 94% smokers) comprising 31 squamous and 68 nonsquamous cell carcinomas. We detected 227 nonsilent variants in the coding sequence, including 24 samples with nonoverlapping, classic driver alterations. The frequency of driver mutations was not significantly different from that of whites, and no association was found between genetic ancestry and the presence of somatic mutations. Copy number alteration analysis disclosed distinguishable amplifications in the 3q chromosome arm in squamous cell carcinomas and pointed toward a handful of targetable alterations. We also found frequent *SMARCA4* mutations and protein loss, mostly in driver-negative tumors.

Conclusion

Our data suggest that African American ancestry may not be significantly different from European/white background for the presence of somatic driver mutations in NSCLC. Furthermore, we demonstrated that using a comprehensive genotyping approach could identify numerous targetable alterations, with potential impact on therapeutic decisions.

J Clin Oncol 33:1966-1973. © 2015 by American Society of Clinical Oncology

INTRODUCTION

Sequencing technologies have enabled the comprehensive analysis of genetic perturbations in non–small-cell lung cancer (NSCLC)^{1–6} and leveraged the identification of therapeutic targets. Landmark examples include the discovery of epidermal growth factor receptor (*EGFR*) mutations and anaplastic lymphoma kinase (*ALK*) translocations in lung tumors that demonstrate outstanding sensitivity to specific kinase inhibitors.^{7,8}

It is well known that the prevalence of classic somatic mutations may vary in different settings,

according to clinical characteristics (age, sex, smoking status), tumor histology, and population of origin.^{7–9} For instance, *EGFR* mutations are found almost exclusively in lung adenocarcinomas and occur most frequently in tumors from never-smokers, women, and in Asian populations.^{7–9} The frequency of such mutations varies from approximately 10% of lung adenocarcinomas in North America and Europe to as high as 50% to 60% in Asia.^{10,11}

The molecular profile of NSCLC in African Americans (AAs) has been poorly explored thus far.^{12,13} This ethnic group comprises approximately 13% of the US population¹⁴ and presents higher

lung cancer incidence and mortality rates.^{15,16} Moreover, a lower frequency of *EGFR* mutations has been suggested in AAs.^{12,17} Although the reason for these differences is still a matter of debate, there is a clear need for a more comprehensive characterization of the somatic alterations in this group to better define novel targets for therapy.

It should be noted that self-reported race on the basis of skin color is relatively inaccurate, mostly because of racial admixture.^{18,19} Indeed, although the genetic background of AAs is mostly derivable from African ancestors, a significant admixture with European genetic markers has been documented.^{18,19} For this reason, germline ancestry informative markers (AIMs) have been recommended to characterize the genetic origin within admixed populations such as AAs.²⁰⁻²²

Herein, we used a targeted massively parallel sequencing (MPS) approach to assess somatic mutations and copy number alterations (CNAs) in NSCLC samples resected from AAs. The main purpose was to comprehensively define the spectrum of somatic alterations in NSCLC samples resected from AAs and compare these results with historic data from whites. In addition, a panel of AIMs was applied to define the genetic ancestry in this cohort and interrogate a link between fractional global African ancestry and the frequency of somatic alterations. Ultimately, we aimed to define candidate driver alterations to guide targeted therapies in the present and near future.

METHODS

Patients

The study cohort comprised NSCLC samples resected from patients self-reported as AAs at the James Cancer Center of Ohio State University (Columbus, OH) between 1988 and 2011. Each sample was assigned a unique, unidentifiable code, and clinical data were reviewed and annotated. For comparison purposes, histologic subtypes were classified as either squamous (pure squamous cell carcinoma) or nonsquamous (including all other subtypes). Tumor slides were examined to confirm the samples' histology and adequacy for sequencing. To compare the frequency of mutations between NSCLC samples resected from AAs and whites, we identified cases reported as "White" in the lung adenocarcinoma and lung squamous cell carcinoma data sets from The Cancer Genome Atlas (TCGA). The institutional review board approved this project and waived the need for informed consent.

Genotyping

We identified 81 target genes relevant to NSCLC by mining existing databases, including the Catalogue of Somatic Mutations in Cancer (COSMIC)²³ and TCGA, as well as reference articles in the field.¹⁻⁵ A custom panel was designed using the HaloPlex Design Wizard software (Agilent Technologies, Santa Clara, CA) to cover the coding regions of the 81 selected genes, along with 40 previously established AIMs.²¹ The total panel covered 920,980 base pairs and included 44,234 amplicons. Libraries were constructed and indexed using the Agilent HaloPlex kit (Agilent Technologies). The indexed libraries were pooled and paired-end sequenced to $\times 1,000$ average coverage on an Illumina HiSeq 2500 (Illumina, San Diego, CA). The panel was validated using NSCLC cell lines and clinical samples with known mutation status. Variants were filtered using a multistep algorithm, and classic driver mutations commonly included in clinical sequencing platforms for NSCLC were nominated here as canonical. *ALK* translocations were evaluated in nonsquamous tumors using the Vysis *ALK* Break Apart FISH [fluorescent in situ hybridization] Probe Kit (Abbott Molecular, Des Plaines, IL). More information on methodologic aspects is provided in the Data Supplement.

Statistical Methods

The *t* test and Wilcoxon rank test were used for continuous end points (transversion rate and age, respectively) in comparisons between two indepen-

dent groups (male *v* female, stage I to II *v* III to IV, AA *v* white, mutant *v* wild type, squamous *v* nonsquamous). Fisher's exact test was used for proportions of mutations compared between two independent groups. One-way ANOVA model was used in comparisons among three or more groups for continuous measurement such as transversion rate compared among three smoking groups (former/current/never). The impact of ethnicity on the mutation rate was also analyzed by multivariable logistic regression with all other covariates present (sex, smoking status, histology, stage). Pearson's correlation coefficient between each pair of CNAs was calculated to determine similarities among samples and genes in heat maps. *P* values were adjusted for multiple comparisons by Holm's procedure in analyses involving tests with multiple genes (mutation status and CNA comparisons in squamous *v* nonsquamous). A *P* value of $\leq .05$ was considered statistically significant. Statistical analyses were performed using R version 3.0.1 (R Project for Statistical Computing), SAS 9.3 (SAS Institute, Cary, NC), and IBM SPSS version 22.0 (IBM, Armonk, NY). To estimate the genetic ancestry in each AA patient, we applied a model-based clustering algorithm using Structure software version 2.3.4 (Pritchard Lab, Stanford University, Stanford, CA),²⁴ as described in the Data Supplement.

RESULTS

Sequencing Results

Ninety-nine NSCLC samples resected from AA patients were identified, comprising 31 squamous and 68 nonsquamous tumors (Table 1 and Data Supplement). Most patients were males (64%) and

Table 1. Comparison of Clinical Characteristics and Sequencing Results Between the African American Cohort From OSU and White Patients Included in TCGA

Characteristic	African Americans (N = 99)		Whites (N = 283)		P
	No.	%	No.	%	
Clinical characteristics					< .001
Median age, years	61		68		
Range	41-76		40-86		
Sex					.035
Male	63	63.6	145	51.2	
Female	36	36.4	138	48.8	
Histology					.184
Squamous	31	31.3	111	39.2	
Nonsquamous	68	68.7	172	60.8	
Stage					.880
I-II	75	79.8	229	80.9	
III-IV	19	20.2	54	19.1	
Smoking status					.282
Current/former	86	94.5	244	90.4	
Never	5	5.5	26	9.6	
Sequencing results					
<i>KRAS</i> status					.760
Mutant	16	16.2	51	18.0	
Wild type	83	83.8	232	82.0	
<i>EGFR</i> status					1.00
Mutant	5	5.1	17	6.0	
Wild type	94	94.9	266	94.0	
Any driver mutations					.163
Mutant	24	24.2	90	31.8	
Wild type	75	75.8	193	68.2	

Abbreviations: OSU, Ohio State University; TCGA, The Cancer Genome Atlas.

smokers (76% current, 11% former, 5% never-smokers, and 8% unknown). We detected 227 nonsilent variants in the coding sequences of the 81 selected genes after using a multistep filtering process. The median and mean mutant genes per sample were 2 and 2.4 (range, 0 to 11), with 78 tumors presenting at least one mutation. The mean number of mutant genes per sample was not significantly different between squamous and nonsquamous tumors (mean values, 2.0 [standard deviation (SD), 1.7] and 2.6 [SD, 2.1], respectively; $P = .185$). Of the 81 genes selected, 51 were mutated at least in one patient case, with a median and mean of 1 and 2.9 patient cases per gene (range, 0 to 33).

The most frequently mutated genes were *TP53* (33 patient cases), *LRP1B* (19), *KRAS* (16), *CSMD1* (14), *KEAP1* (14), *SMARCA4* (7), *MLL2* (7), *EYS* (7), *EGFR* (6), *APC* (6), *EPHA3* (6), *NOTCH1* (6), and *ROS1* (6). Mutations in *LRP1B*, *KRAS*, *CSMD1*, *SMARCA4*, and *EGFR* tended to occur in nonsquamous tumors. Conversely, *KEAP1*, *MLL2*, *EYS*, *NFE2L2*, and *CDKN2A* mutations tended to predominate in squamous cell carcinomas (Figs 1A and 1B). The majority of the variants detected are located in conserved and likely active domains, suggesting a functional role for variants that passed the filtering process. Canonical driver mutations were found in *KRAS* (16 samples), *EGFR* (5), *PIK3CA* (1), and *NRAS* (1). In addition, one sample was positive for an *ALK* translocation by FISH, totaling 24 patient cases harboring classic driver alterations, with no overlap (Fig 1C). In line with recent observations,^{2,6} the transversion rate tended to be lower among tumors resected from former (mean, 0.42 [SD, 0.28]; $P = .116$) and never-smokers (mean, 0.32 [SD, 0.28]; $P = .069$), in comparison with current smokers (mean, 0.57 [SD, 0.29]). Additional details are provided in the Data Supplement.

Comparison With Whites

To compare our data for AAs with that of whites, we analyzed the TCGA database and identified 283 NSCLC samples resected from whites. Except for an older median age at diagnosis (68 v 61 years; $P < .001$) and a higher frequency of females (48.8% v 36.4%; $P = .035$) in TCGA, there were no significant differences among major clinical and tumor characteristics, including histology ($P = .184$), tumor stage ($P = .880$), and smoking status ($P = .282$). In a direct comparison, the frequencies of driver mutations in either *KRAS* (18% v 16%; $P = .760$), *EGFR* (6% v 5%; $P = 1.00$), or any driver mutations (32% v 24%; $P = .163$) were not significantly different between whites and AAs, respectively (Table 1). Combining these two data sets, we confirmed that ethnicity was not significantly associated with the status of *KRAS*, *EGFR*, or any driver mutation. In a multivariable analysis, smoking status and histology were the only factors that were significantly associated with *KRAS* and *EGFR* status, whereas histology was associated with any driver mutation (Data Supplement).

Genetic Ancestry

The mean proportions of African, European, and Amerindian genetic background were 0.65, 0.25, and 0.10, respectively. These data confirm a predominance of African ancestry, but also an evident admixture and within-group variability. To evaluate a possible link between the genetic ancestry and the presence of somatic mutations, we explored the mean proportions of African, European, and Amerindian backgrounds between patients with or without driver mutations. Among patients with *KRAS* mutations, the mean proportions of African, European, and Amerindian ancestry were 0.67, 0.23, and 0.10, without a statistically significant difference in comparison with

Fig 1. (A) Most frequently mutated genes in non-small-cell lung cancer (NSCLC) tumors resected from African Americans and (B) percentage according to tumor histology. Each bar represents a cancer gene as labeled in the x-axis, whereas the y-axis indicates the (A) frequency or (B) percentage of patient cases that presented a mutation in each gene. (C) Classic mutations are depicted as pie charts for NSCLC and nonsquamous patient cases. (*) P value $< .05$.

Fig 2. (A) Triangular plots of the genomic proportions of African, European, and Amerindian ancestry in 99 patients self-reported as African Americans. Each point represents an individual patient. African Americans present a predominantly African genetic background, along with European admixture. Patient cases harboring canonical (B) *EGFR*, (C) *KRAS*, or (D) any driver mutation are considerably sparse within the groups (in gold), suggesting that the genetic ancestry is not associated with the presence of specific somatic mutations.

patients without *KRAS* mutations (*P* values of .628, .522, and .738, respectively). Likewise, no difference was observed according to *EGFR* mutations or any canonical driver mutations. These data are represented in triangular plots in Figure 2. We also explored potential differences according to other gene mutations, but no evident trend was detected (Data Supplement).

Candidate Driver Genes

To evaluate potential driver genes in this cohort, we compared the frequency of mutant genes in driver-positive and driver-negative samples. Among the most frequently altered genes, *KEAP1*, *SMARCA4*, *NOTCH1*, *NFE2L2*, *PIK3CA*, *CDKN2A*, and *MTOR* showed higher frequencies in driver-negative tumors, although statistical significance was not achieved (Fig 3A). In the case of *SMARCA4*, we were able to integrate genetic data with protein expression, as assessed by immunohistochemistry (Figs 3B and 3C). Ten samples (12.5%) had low or negative *SMARCA4* protein (also known as BRG1) expression (scores 0 to 1), nine of which were driver-negative cases. Four cases were explained by truncating, nonsense mutations in *SMARCA4* (Fig 3D); all truncating mutations were associated with low/absent protein expression ($P < .001$). Conversely, some missense

mutations were found in samples with high expression (score +3). These variants passed the filtering steps and are located in conserved regions of the protein. According to COSMIC, similar mutations have been found in colorectal (G1146S and A1186V), renal (A1186G), and endometrial carcinomas (R1520H). The functional consequence of these missense mutations is currently unknown. Interestingly, two samples with loss of *SMARCA4* expression presented truncating mutations in *ARID1A*. *SMARCA4* mutations (seven samples) and *ARID1A* mutations (five samples) were mutually exclusive, and no CNA was detected for *SMARCA4*.

CNAs

We detected a mean of 2.4 and median of 1.0 CNAs per sample, including both focal and segmental alterations. In an unsupervised analysis, we observed two distinct sample clusters, the first enriched by nonsquamous and the second by squamous cell carcinomas (Data Supplement). To elucidate the genetic differences between these histologic subtypes, we applied a direct comparison in the gene level and demonstrated that only three genes were statistically significant: *PIK3CA*, *SOX2*, and *TP63* (adjusted *P* values of .0156, $< .001$, and .0038, respectively). These genes are all located in the 3q chromosome

Fig 3. (A) Candidate driver genes with higher frequency of mutations in driver-negative versus driver-positive patient cases. (B, C, and D) Ten samples (12.5%) had low SMARCA4 expression by immunohistochemistry (IHC), nine of which occurred in driver-negative patient cases. Four of these patient cases were explained by truncating mutations in *SMARCA4*, and two nonoverlapping patient cases had truncating mutations in *ARID1A*. All truncating mutations were associated with low protein expression ($P < .001$), whereas some missense mutations were found in samples with high expression (score 3). ATP, adenosine triphosphate; HSA, helicase/SANT-associated; QLQ, glutamine-leucine-glutamine; WT, wild type.

arm and were frequently amplified in squamous tumors, whereas no alteration was detected in nonsquamous tumors (Figs 4A through 4C). We also demonstrated that these genes were highly correlated, suggesting that the 3q segment was amplified in most instances, instead of there being isolated gene amplifications. A similar pattern was observed for genes located in the 8p chromosome arm (*FGFR1* and *WHSC1L1*), which was compatible with broader segmental amplifications. Subsequently, we compared the most frequent CNAs between driver-positive and driver-negative samples and showed that *SOX2*, *TP63*, *PIK3CA*, *FGFR1*, *MCL1*, *FGFR3*, *TNFAIP3*, *MLL2*, and *WHSC1L1* were more frequently amplified in driver-negative samples (Fig 4D). The only gene with higher frequency of deletions was *RB1*, although none achieved statistical significance. Although less frequent, *ERBB2* (three samples) and *MET* (one sample) amplifications were reported in driver-negative cases, and may be involved as drivers in these tumors carcinogenesis. Selected samples with amplification in

major oncogenes were validated using the OncoScan FFPE Assay Kit (Affymetrix; Santa Clara, CA; Data Supplement).

Pathway Analysis

Frequently altered genes code for proteins that are involved in several NSCLC-related pathways. Among our samples, 67 had at least one alteration (mutations or CNAs) in survival and proliferative pathways, including the receptor tyrosine kinase pathway in 23 samples, RAS/MAPK/ERK in 27, and phosphatidylinositol 3-kinase in 29 (Fig 5A). We also detected alterations in the WNT pathway, which is involved in proliferation as well as cell differentiation. These pathways comprise known oncogenes and tumor suppressor genes that play pivotal roles in NSCLC and are potential targets for cancer therapy.²⁵ In our results, we observed a significant pattern of mutual exclusivity among alterations involving oncogenes in these pathways (Fig 5B). Moreover, 45% of the tyrosine and 40% of the serine/threonine kinase

Fig 4. (A) Clustering analysis of copy number alterations. The clinical samples were grouped according to histology (nonsquamous samples on the left, squamous on the right) and sorted by the presence of drivers. The blue arrows on the right side are pointing to the commonly amplified 3q and 8p chromosome arms. (B) *PIK3CA*, *SOX2*, and *TP63*—all in the 3q arm—are more frequently amplified in squamous than in nonsquamous tumors. (C) The segmental origin of *PIK3CA*, *SOX2*, and *TP63* in the 3q arm and *WHSC1L1* and *FGFR1* in the 8p arm is supported by an unsupervised correlation analysis. (D) Alterations with higher frequencies in driver-negative patient cases. Amp, amplification; del, deletion.

mutations were located in kinase domains (Data Supplement), which is compatible with their putative gain-of-function activity. Genes involved in cell cycle control were frequently altered, with 52 samples presenting at least one alteration. A mutual exclusivity pattern was observed for the most common alterations in both RB and TP53 pathways (Fig 5B). Other genes of interest included *KEAP1* and *NFE2L2* in the redox pathways, and the *NOTCH* family in stem-cell differentiation.

DISCUSSION

This in-depth analysis provides important insights into the biology of NSCLC in AAs and confirms potential therapeutic targets in this population. To the best of our knowledge, this is the first study to use an MPS technology to genotype lung cancer in this subgroup and integrate data from DNA sequencing (including mutations and CNAs), FISH, and immunohistochemistry. As opposed to the traditional direct sequencing used in other studies, MPS can simultaneously determine the mutational status of multiple genes in a single reaction, with significantly higher coverage depth and accuracy rates.^{26,27} Furthermore, our panel has the advantage of covering all of the exons of each of the 81 genes included, whereas other multiplex technologies focus only on known hotspots and therefore lack the ability to assess less frequent, noncanonical alterations.^{28,29}

Given that AAs have higher lung cancer incidence and mortality,^{15,16} we hypothesized that AAs could present with different, perhaps more aggressive subtypes of NSCLC. Previous studies have reported discrepant results concerning the frequency of somatic mutations in AAs and are limited by the smaller number of samples, smaller number of genes analyzed, depth of analysis,^{12,17,30-33} and the lack of clinical data.³⁴ In our cohort, only 24 tumors (of 99) were found to have canonical driver alterations, including mutations in *KRAS*, *EGFR*, *PIK3CA*, *NRAS*, and an *ALK* fusion. Nevertheless, there was not a significant difference in the frequency of driver mutations in comparison with white patients on the basis of data in TCGA. Corroborating this observation, we have not detected a significant association between African ancestry and the presence of specific driver mutations. Taken together, these findings suggest that AA ethnicity—and African ancestry per se—may not be associated with a distinct mutational profile in NSCLC. Among the limitations, the current study was not designed to address patient outcomes data, and our cohort predominantly included smokers with early-stage disease. For this reason, our genomics results may not reflect the mutation frequencies in specific subgroups, such as AA never-smokers. Additional studies that target this subset may be needed to better define specific mutation patterns. The mechanisms underlying the clear difference in frequency of abnormalities in the Asian population are still unknown.

Fig 5. (A) Gene alterations in central pathways in non-small-cell lung cancer. Survival/proliferation pathways comprise the receptor tyrosine kinase (RTK), RAS/MAPK/ERK, phosphatidylinositol 3-kinase (PI3K), and some crosstalk with the WNT pathway. The canonical RB1 and TP53 pathways are illustrated controlling cell cycle and apoptosis. (B) Major alterations in oncogenes and cell cycle genes are shown, with a significant pattern of mutual exclusivity. CNAs, copy number alterations.

Apart from classic driver mutations, we were able to assess multiple noncanonical somatic mutations and CNAs in this group, and had a special interest in finding possible candidate driver alterations that could be clinically targetable at present or in the near future. We found numerous alterations in otherwise driver-negative samples that could potentially be used to guide therapy. For instance, amplifications in *PIK3CA*, *FGFR1*, and *PDGFRA* have been used as biomarkers to test targeted agents, especially in squamous cell carcinomas of the lung.^{35,36} In addition, we report four samples with either *ERBB2* or *MET* amplifications, which are potential drivers in lung adenocarcinomas.⁶ Altogether, we detected 67 samples with at least one alteration in genes that are involved in survival pathways, 42 of which are reported as gain-of-function changes in major oncogenes. These changes were mostly mutually exclusive and could have immediate implications for therapy.

The proteins encoded by *SMARCA4* and *ARID1A* are key components in the SWI/SNF (SWItch/Sucrose NonFermentable) complex, which regulates the transcription of multiple genes through chromatin remodeling. These genes are commonly mutated in numerous cancer types^{2,6} and have been implicated as plausible tumor suppressors in NSCLC.^{37,38} Moreover, germline mutations in *SMARCA4* have been associated with pediatric atypical teratoid/rhabdoid tumors and small-cell carcinoma of the ovary (hypercalcemic type).³⁹⁻⁴¹ In NSCLC, approximately 10% to 30% of patients are estimated to have loss of *SMARCA4* expression, which has been associated with poor prognosis.^{42,43} These patient cases are partially explained by the presence of truncating mutations in NSCLC

(approximately 5%), whereas epigenetic silencing has been proposed as an alternative mechanism.^{44,45} In our cohort, we detected 10 patient cases (12.5%) with deficient expression, four of which were associated with nonsense mutations. Interestingly, we found two samples with *SMARCA4* loss of expression and truncating mutations in *ARID1A*. This finding led us to suggest that *ARID1A* mutation could be causally related to the *SMARCA4* loss of expression. Supporting this hypothesis, it has been demonstrated in vitro that *ARID1A* knockdown leads to lower expression and lower incorporation of *SMARCA4* into the SWI/SNF complex.⁴⁶ If our observation is supported by future work, it could have important implications for drug development, given that different causes of *SMARCA4* loss may require distinct therapies. In addition, we detected three samples with *SMARCA4* missense mutations with high protein expression. The possible acquired functional consequences of these mutations are currently unknown and may require further investigation.

In summary, we demonstrated a relatively low frequency of classic driver mutations in NSCLC among AAs. However, these results are comparable with the data reported for whites and suggest that AA ethnicity may not be a surrogate marker for the presence of specific driver mutations. These data were corroborated by a lack of association between genetic ancestry and oncogenic driver mutations. In addition, we demonstrated that using an in-depth genotyping approach could identify multiple genetic alterations. More importantly, many of these abnormalities could be immediately used to guide therapy decisions and patient enrollment onto clinical trials.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Luiz H. Araujo, Phillip E. Lammers, David P. Carbone

Financial support: David P. Carbone

Administrative support: Joseph Amann

Collection and assembly of data: Luiz H. Araujo, Erica Hlavin Bell, Konstantin Shilo, Weiqiang Zhao, Thanemozhi G. Natarajan, Clinton J. Miller, Tom Liu, Joseph Amann

Data analysis and interpretation: Luiz H. Araujo, Cynthia Timmers, Erica Hlavin Bell, Konstantin Shilo, Weiqiang Zhao, Thanemozhi G. Natarajan, Clinton J. Miller, Jianying Zhang, Ayse S. Yilmaz, Tom Liu, Kevin Coombes, Joseph Amann

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

- Ding L, Getz G, Wheeler DA, et al: Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069-1075, 2008
- Imielinski M, Berger AH, Hammerman PS, et al: Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150:1107-1120, 2012
- Cancer Genome Atlas Research Network: Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 489:519-525, 2012
- Govindan R, Ding L, Griffith M, et al: Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* 150:1121-1134, 2012
- Seo JS, Ju YS, Lee WC, et al: The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res* 22:2109-2119, 2012
- Cancer Genome Atlas Research Network: Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511:543-550, 2014
- Jänne PA, Engelman JA, Johnson BE: Epidermal growth factor receptor mutations in non-small-cell lung cancer: Implications for treatment and tumor biology. *J Clin Oncol* 23:3227-3234, 2005
- Lynch TJ, Bell DW, Sordella R, et al: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350:2129-2139, 2004
- Cappuzzo F, Hirsch FR, Rossi E, et al: Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* 97:643-655, 2005
- Rosell R, Moran T, Queralt C, et al: Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med* 361:958-967, 2009
- Shi Y, Au JS, Thongprasert S, et al: A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *J Thorac Oncol* 9:154-162, 2014
- Bauml J, Mick R, Zhang Y, et al: Frequency of EGFR and KRAS mutations in patients with non-small cell lung cancer by racial background: Do disparities exist? *Lung Cancer* 81:347-353, 2013
- Yamaguchi N, Vanderlaan PA, Folch E, et al: 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. *Lung Cancer* 82:31-37, 2013
- US Census Bureau: The Black Population: 2010. <http://www.census.gov/prod/cen2010/briefs/c2010br-06.pdf>
- Ward E, Jemal A, Cokkinides V, et al: Cancer disparities by race/ethnicity and socioeconomic status. *CA Cancer J Clin* 54:78-93, 2004
- DeSantis C, Naishadham D, Jemal A: Cancer statistics for African Americans, 2013. *CA Cancer J Clin* 63:151-166, 2013
- Phelps MA, Stinchcombe TE, Blachly JS, et al: Erlotinib in African Americans with advanced non-small cell lung cancer: A prospective randomized study with genetic and pharmacokinetic analyses. *Clin Pharmacol Ther* 96:182-191, 2014
- Zakharia F, Basu A, Absher D, et al: Characterizing the admixed African ancestry of African Americans. *Genome Biol* 10:R141, 2009
- Tang H, Jorgenson E, Gadde M, et al: Racial admixture and its impact on BMI and blood pressure in African and Mexican Americans. *Hum Genet* 119:624-633, 2006
- Yaeger R, Avila-Bront A, Abdul K, et al: Comparing genetic ancestry and self-described race in African Americans born in the United States and in Africa. *Cancer Epidemiol Biomarkers Prev* 17:1329-1338, 2008
- Bastos-Rodrigues L, Pimenta JR, Pena SD: The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann Hum Genet* 70:658-665, 2006
- Rosenberg NA, Pritchard JK, Weber JL, et al: Genetic structure of human populations. *Science* 298:2381-2385, 2002
- Forbes SA, Tang G, Bindal N, et al: COSMIC (the Catalogue of Somatic Mutations in Cancer): A resource to investigate acquired mutations in human cancer. *Nucleic Acids Res* 38:D652-D657, 2010
- Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 155:945-959, 2000
- Buettner R, Wolf J, Thomas RK: Lessons learned from lung cancer genomics: The emerging concept of individualized diagnostics and treatment. *J Clin Oncol* 31:1858-1865, 2013
- Mardis ER: A decade's perspective on DNA sequencing technology. *Nature* 470:198-203, 2011
- Tran B, Dancy JE, Kamel-Reid S, et al: Cancer genomics: Technology, discovery, and translation. *J Clin Oncol* 30:647-660, 2012
- Dias-Santagata D, Akhavanfard S, David SS, et al: Rapid targeted mutational analysis of human tumours: A clinical platform to guide personalized cancer medicine. *EMBO Mol Med* 2:146-158, 2010
- MacConaill LE, Campbell CD, Kehoe SM, et al: Profiling critical cancer gene mutations in clinical tumor samples. *PLoS One* 4:e7887, 2009
- Cote ML, Haddad R, Edwards DJ, et al: Frequency and type of epidermal growth factor receptor mutations in African Americans with non-small cell lung cancer. *J Thorac Oncol* 6:627-630, 2011
- Leidner RS, Fu P, Clifford B, et al: Genetic abnormalities of the EGFR pathway in African American patients with non-small-cell lung cancer. *J Clin Oncol* 27:5620-5626, 2009
- Yang SH, Mechanic LE, Yang P, et al: Mutations in the tyrosine kinase domain of the epidermal growth factor receptor in non-small cell lung cancer. *Clin Cancer Res* 11:2106-2110, 2005
- Krishnaswamy S, Kanteti R, Duke-Cohan JS, et al: Ethnic differences and functional analysis of MET mutations in lung cancer. *Clin Cancer Res* 15:5714-5723, 2009
- Reinersman JM, Johnson ML, Riely GJ, et al: Frequency of EGFR and KRAS mutations in lung adenocarcinomas in African Americans. *J Thorac Oncol* 6:28-31, 2011
- Rooney M, Devarakonda S, Govindan R: Genomics of squamous cell lung cancer. *Oncologist* 18:707-716, 2013
- Weiss J, Sos ML, Seidel D, et al: Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci Transl Med* 2:62ra93, 2010
- Wu JN, Roberts CW: ARID1A mutations in cancer: Another epigenetic tumor suppressor? *Cancer Discov* 3:35-43, 2013
- Hoffman GR, Rahal R, Buxton F, et al: Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc Natl Acad Sci U S A* 111:3128-3133, 2014
- Schneppenheimer R, Frühwald MC, Gesk S, et al: Germline nonsense mutation and somatic inactivation of SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome. *Am J Hum Genet* 86:279-284, 2010
- Witkowski L, Carrot-Zhang J, Albrecht S, et al: Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcaemic type. *Nat Genet* 46:438-443, 2014
- Jelinic P, Mueller JJ, Olvera N, et al: Recurrent SMARCA4 mutations in small cell carcinoma of the ovary. *Nat Genet* 46:424-426, 2014
- Reisman D, Glaros S, Thompson EA: The SWI/SNF complex and cancer. *Oncogene* 28:1653-1668, 2009
- Fukuoka J, Fujii T, Shih JH, et al: Chromatin remodeling factors and BRM/BRG1 expression as prognostic indicators in non-small cell lung cancer. *Clin Cancer Res* 10:4314-4324, 2004
- Thompson KW, Marquez SB, Reisman D: A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1: Letter. *Cancer Res* 74:4946-4947, 2014
- Oike T, Ogiwara H, Tominaga Y, et al: A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. *Cancer Res* 73:5508-5518, 2013
- Helming KC, Wang X, Wilson BG, et al: ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med* 20:251-254, 2014

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST**Genomic Characterization of Non–Small-Cell Lung Cancer in African Americans by Targeted Massively Parallel Sequencing**

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Luiz H. Araujo

No relationship to disclose

Cynthia Timmers

No relationship to disclose

Erica Hlavin Bell

No relationship to disclose

Konstantin Shilo

No relationship to disclose

Phillip E. Lammers

Consulting or Advisory Role: Pfizer, Boehringer Ingelheim

Weiqiang Zhao

No relationship to disclose

Thanemozhi G. Natarajan

Employment: GenomOncology

Clinton J. Miller

Employment: GenomOncology

Jianying Zhang

No relationship to disclose

Ayse S. Yilmaz

No relationship to disclose

Tom Liu

No relationship to disclose

Kevin Coombes

No relationship to disclose

Joseph Amann

No relationship to disclose

David P. Carbone

Consulting or Advisory Role: Genentech/Roche, Bristol-Myers Squibb, Boehringer Ingelheim, Novartis, Pfizer, Merck, GlaxoSmithKline

Research Funding: Bristol-Myers Squibb (Inst)