VOLUME 32 · NUMBER 2 · JANUARY 10 2014

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Integrative and Comparative Genomic Analysis of Lung Squamous Cell Carcinomas in East Asian Patients

Youngwook Kim, Peter S. Hammerman, Jaegil Kim, Ji-ae Yoon, Yoomi Lee, Jong-Mu Sun, Matthew D. Wilkerson, Chandra Sekhar Pedamallu, Kristian Cibulskis, Yeong Kyung Yoo, Michael S. Lawrence, Petar Stojanov, Scott L. Carter, Aaron McKenna, Chip Stewart, Andrey Y. Sivachenko, In-Jae Oh, Hong Kwan Kim, Yong Soo Choi, Kwhanmien Kim, Young Mog Shim, Kyu-Sik Kim, Sang-Yun Song, Kook-Joo Na, Yoon-La Choi, D. Neil Hayes, Jhingook Kim, Sukki Cho, Young-Chul Kim, Jin Seok Ahn, Myung-Ju Ahn, Gad Getz, Matthew Meyerson, and Keunchil Park

Author affiliations appear at the end of this article.

Published online ahead of print at www.jco.org on December 9, 2013.

Supported by Grant No. A111218-GM08 from the National Project for Personalized Genomic Medicine, Ministry for Health & Welfare, Republic of Korea; by Korea Biobank Network (provided tissue samples; distribution 2012-0006); by Grant No. NCI K08 CA163677 (P.S.H.); and by the Stephen and Alice Cutler Foundation.

Y.K., P.S.H., and J.K. contributed equally to this study. G.G., M.M., and K.P. are principal investigators who contributed equally to this study.

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

Corresponding author: Keunchil Park, MD, PhD, Division of Hematology/Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 50 Irwon-dong, Gangnam-gu, Seoul 135-710, Korea; e-mail: kpark@skku.edu.

© 2013 by American Society of Clinical Oncology

0732-183X/14/3202w-121w/$20.00

DOI: 10.1200/JCO.2013.50.8556

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

ABSTRACT

Purpose
Lung squamous cell carcinoma (SCC) is the second most prevalent type of lung cancer. Currently, no targeted therapeutics are approved for treatment of this cancer, largely because of a lack of systematic understanding of the molecular pathogenesis of the disease. To identify therapeutic targets and perform comparative analyses of lung SCC, we probed somatic genome alterations of lung SCC by using samples from Korean patients.

Patients and Methods
We performed whole-exome sequencing of DNA from 104 lung SCC samples from Korean patients and matched normal DNA. In addition, copy-number analysis and transcriptome analysis were conducted for a subset of these samples. Clinical association with cancer-specific somatic alterations was investigated.

Results
This cancer cohort is characterized by a high mutational burden with an average of 261 somatic exonic mutations per tumor and a mutational spectrum showing a signature of exposure to cigarette smoke. Seven genes demonstrated statistical enrichment for mutation: *TP53, RB1, PTEN, NFE2L2, KEAP1, MLL2,* and *PIK3CA*). Comparative analysis between Korean and North American lung SCC samples demonstrated a similar spectrum of alterations in these two populations in contrast to the differences seen in lung adenocarcinoma. We also uncovered recurrent occurrence of therapeutically actionable *FGFR3-TACC3* fusion in lung SCC.

Conclusion
These findings provide new steps toward the identification of genomic target candidates for precision medicine in lung SCC, a disease with significant unmet medical needs.

J Clin Oncol 32:121-128. © 2013 by American Society of Clinical Oncology

INTRODUCTION

Lung cancer is the leading cause of cancer-related death, resulting in an estimated 1.4 million deaths per year worldwide.[1] Lung squamous cell carcinoma (SCC) accounts for approximately 30% of overall lung cancer cases. This type of tumor typically occurs in large airways, and the etiologic role of smoking in lung SCCs is well established.[2]

In lung adenocarcinoma, it has been observed that patients harboring mutations in the epidermal growth factor receptor (*EGFR*) gene typically demonstrate dramatic responses to EGFR tyrosine kinase inhibitors. This provided the first demonstration that oncogenic alterations could serve as therapeutic biomarkers in lung cancer.[3,4] Subsequent large-scale clinical trials based on genotype-guided patient stratification have demonstrated a superior clinical benefit of molecularly targeted therapies compared with standard chemotherapy in patients with lung adenocarcinoma who have *EGFR* mutations and *ALK* fusions.[5-7] Other molecular inhibitors targeting *BRAF*, *ERBB2*, *RET*, and *ROS* alterations are currently being pursued in clinical trials with promising initial results.

Previous studies have demonstrated recurrent mutation of several genes in SCC, most notably *TP53*, *NFE2L2*, and *CDKN2A*[8-11] as well as copy-number alterations of *SOX2*, *CCND1*, *FGFR1*, and *CDKN2A*[12-14] in the pathogenesis of lung SCC. Recently, a large-scale comprehensive analysis of a large lung SCC cohort was reported by The Cancer

Genome Atlas (TCGA) Network that identified statistically recurrent somatic mutations in 18 genes that included *TP53*, *NFE2L2*, *KEAP1*, *PIK3CA*, *CDKN2A*, *MLL2*, *HLA-A*, and others.[15] However, in contrast to lung adenocarcinoma, in which molecularly targeted therapies are in routine clinical use to treat subsets of individuals with alterations in genes such as *EGFR* and *ALK*,[3-7] there are no approved targeted therapies for lung SCC. Moreover, subsets of patients with lung cancer with no apparent therapeutic biomarkers, such as EGFR and ALK, generally have worse prognosis, prompting the necessity of identifying driver oncogenic events for this subset of patients.[1,16]

In lung adenocarcinoma, the prevalence and distribution of specific oncogenic mutations vary widely among demographic groups and are especially pronounced when comparing individuals of mostly white and East Asian descent. For example, East Asians display a higher rate of *EGFR* mutations and a lower rate of *KRAS* mutations.[17,18] These demographic differences have had a substantial impact on the management of lung adenocarcinoma at regional centers and on the design of clinical trials.

In lung SCC, no large studies have yet been completed to determine the spectrum of genomic alterations in Asians. The prevalence of smoking among Koreans is still high (> 40% among men), and SCC accounts for 25% to 30% of lung cancer cases.[19] To identify therapeutic targets in lung SCC, we probed somatic alterations in tumor samples from Korean patients and compared the results with data from the completely nonoverlapping TCGA cohort,[15] which consisted primarily of patients from academic centers in the United States. Here, we present the multidimensional analysis of genomic alterations in more than 100 Korean patients with lung SCC by using whole-exome and transcriptome sequencing and single nucleotide polymorphism array analysis.

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

PATIENTS AND METHODS

Study Population

The study included an institutional review board–approved retrospective analysis of 104 Korean patients with lung SCC. Information on sex, age, tumor stage, smoking record, and overall survival was extracted from the clinical database for this cohort (Data Supplement). None of these patients had previously been subject to genomic characterization, and no patients from these hospitals were part of the TCGA study of lung SCC.

Sample Preparation, Sequencing Data Analysis, and Statistical Methods

Analysis of sequencing data was performed as previously described.[15,20] Detailed methods for sample preparation and sequencing analysis can be found in the Data Supplement (for statistical analysis, each statistical method used is described in the text where the data are presented). Enrichment analysis of specific somatic mutations and copy-number alterations in somatic copy-number alteration (SCNA) clusters was performed by using Fisher's exact test. Exclusivity and co-occurrence of mutations were examined by using one-tailed Fisher's exact tests. For multiple test adjustment, false discovery rate was controlled by calculating *q* values.[21] Other statistical methods are described in the main text and the Data Supplement.

RESULTS

Patient Cohort Description

To provide comprehensive understanding of genetic abnormalities occurring in lung SCC, we used massively parallel paired-end sequencing technology[22] to perform whole-exome solution-based hybrid capture sequencing of 104 matched tumor/normal pairs. These samples were from individuals who underwent resection of a lung SCC from 2005 to 2012. The median age of patients in the study was 65 years with a median of 29.0 months of follow-up (Korean columns in Table 1). The cohort included five self-reported never-smokers, 52 former smokers, and 47 current smokers; most patients had extensive smoking histories. The cohort included 43 stage I, 41 stage II, 18 stage III, and two stage IV lung SCC tumors, all of which were primary resection specimens. Clinical and histopathologic information for the cohort set used in the study is provided in the Data Supplement. In addition to whole-exome sequencing, we also conducted transcriptome sequencing and copy-number analysis for a subset of this patient cohort (Data Supplement).

Table 1. Clinical Data Summary

Variable	Korean (n = 104) No.	Korean (n = 104) %	TCGA[15] (n = 178) No.	TCGA[15] (n = 178) %
Age at surgery, years				
Median	65		68.0	
Range	43-86		40.0-85.0	
Sex				
Male	100	96.2	131	73.6
Female	4	3.8	47	26.4
Smoking status				
Never-smoker	5	4.8	7	3.9
Former smoker, pack-years	52	50	137	77
≤ 20	7	6.7		
> 20	45	43.3		
Current smoker, pack-years	47	45.2	28	15.7
≤ 20	4	3.8		
> 20	43	41.3		
N/A	0	0	6	3.4
Median follow-up, months	29		15.8	
Tumor stage				
I	43	41.3	97	54.5
II	41	39.4	38	21.3
III	18	17.3	38	21.3
IV	2	1.9	3	1.7
N/A	0	0	2	1.1
T stage				
T1	24	23.1	35	19.7
T2	63	60.6	116	65.2
T3	12	11.5	14	7.9
T4	5	4.8	12	6.7
N/A	0	0	1	0.6
N stage				
N0	70	67.3	117	65.7
N1	20	19.2	39	21.9
N2	14	13.5	17	9.6
N3	0	0	5	2.8

Abbreviations: N/A, not applicable; TCGA, The Cancer Genome Atlas.

Detection of Somatic Aberrations

Analysis of whole-exome sequencing data from our cohort identified a total of 29,353 point mutations and 935 insertion or deletion events across 104 tumor/normal pairs in coding regions of the genome. Mean sequencing depth in the target regions was 62.8X. Of the

29,353 point mutations, 21,550 resulted in an amino acid change in the corresponding proteins. Individuals displayed a mean point mutation rate of 8.71 mutations per Mb (range, 1.19 to 36.2 mutations per Mb). Among nonsynonymous substitutions, transitions and transversions at CpG sites were the most commonly observed mutational context with a rate of 16.18 per Mb and 9.53 per Mb, respectively. Mutations at Cp(A/C/T) sites were the second most commonly observed context at 7.68 per Mb (Data Supplement). Hierarchical clustering of tumors based on mutational spectra is shown in the Data Supplement.

We applied the MutSig method to identify genes that demonstrated evidence of statistical selection for mutation in our cohort.[20,23-25] MutSig is a computational algorithm that takes into account gene length, composition, silent-to-nonsilent mutation ratios, and other factors, and it identified six genes that displayed significant enrichment for mutations as defined by $q < 0.1$, with robust evidence of expression in lung SCC: *TP53, RB1, PTEN, NFE2L2, KEAP1,* and *MLL2* (Fig 1 and Data Supplement). A second MutSig analysis, in which only genes annotated in the Catalogue of Somatic Mutations in Cancer (COSMIC) database were considered, rediscovered *TP53, RB1,* and *PTEN* and additionally nominated *PIK3CA* as statistically significant (Data Supplement). Additional genes with annotated roles in cancer that did not reach statistical significance but did display recurrent mutations were *CHD7* (nine mutations), *KRAS* (two), *NF1* (13), *EGFR* (two), *CDKN2A* (three), *IBTK* (three), *TIE1* (five), *NOTCH1* (seven), *NOTCH3* (nine), *FBXW7* (four), *FGFR1* (two), *FGFR2* (two), *KDM6A* (five), and *APC* (four; Fig 1). Of note, one of two patients with an *EGFR* mutation demonstrated a canonical exon 19 deletion and the other demonstrated D347N. Both *KRAS* mutations resulted in G12C substitutions.

SCNAs were inferred from sequencing data by read-depth analysis (Data Supplement). We used GISTIC 2.0 analysis[26,27] to identify statistically recurrent peaks of amplification and deletion and identified 19 regions of focal amplification and 25 regions of deletion with a $q < 0.25$ (Data Supplement). The most significantly amplified regions were 3q26, an amplicon including *SOX2*; 8p11 (*FGFR1, WHSC1L1*); 7p11 (*EGFR*); 11q13 (*CCDN1*); and 4q12 (*KDR, KIT, PDGFRA*). Deletion peaks tended to be more broad and included 2q37, 4q35 (*CASP3*), 9p21 (*CDKN2A*), and 10q23 (*PTEN*). The copy-number variations were also confirmed and validated by using high-density single nucleotide polymorphism arrays (Data Supplement).

Hierarchical clustering of the cohort based on the SCNAs demonstrated three clusters (Data Supplement): Cluster 1 was defined by a paucity of SCNAs (depicted in red in Fig 1C), Cluster 2 was defined by both broad and focal 3q amplification (depicted in green in Fig 1C), and Cluster 3 was defined by broad 3q amplification and SCNAs in cyclin genes and receptor tyrosine kinase–related genes (depicted in magenta in Fig 1C; Fisher's exact test: *CCND1*, $P < .03$; *CCND2*, $P < .001$; *ERBB2*, $P < .04$; *FGFR1*, $P < .02$; *EGFR, FRS2,* and *PDGFRA,* all $P < .001$). Cluster 1 demonstrated a lower mutation rate compared with the other clusters (Data Supplement). This cluster also contained some samples with lower calculated tumor purity, which may have contributed to a lower prevalence of genomic alterations. Cluster 2 exhibited near exclusivity with *RB1* mutations (Fisher's exact test: $P < .001$) and enriched in *PIK3CA*-activating mutations ($P < .007$). Virtually all the patients in the third cluster (magenta in Fig 1C) possessed *TP53* mutations (Fisher's exact test: $P < .001$; Data Supplement).

Identification of Recurrent Structural Variants by Using RNA Sequencing

Paired-end RNA sequencing was performed on a subset of 26 individuals from the original cohort. We first used these data to validate our mutation calls by calling the mutations in the RNA sequencing data. After controlling for coverage, we validated more than 72% of coding single nucleotide variants and 86% of events in genes depicted in Figure 1 (Data Supplement). Analysis of these data identified a total of 296 gene fusions involving 244 distinct fusion partners (Data Supplement).[28] No fusions involving the *ALK, RET,* or *ROS1* kinase genes were observed. We observed recurrent fusions involving the *CASP1* gene (four), *IGF2BP2* (two), and *KRT5* (two). We identified single-fusion events involving the *FGFR3* and *GSK3β* kinases in the 26 initial individuals surveyed.

Detailed analysis of the *FGFR3* event (in sample J9) demonstrated an in-frame fusion of exon 8 of *TACC3* with exon 17 of *FGFR3* (Fig 2A), a fusion event recently reported as oncogenic and potentially therapeutically targetable in glioblastoma multiforme, bladder cancer, and other types of cancer, including prostate cancer.[29-33] We extended analysis to an additional cohort of 40 patients with lung SCC and identified an additional individual (J1) with an *FGFR3-TACC3* fusion by reverse transcriptase polymerase chain reaction (Fig 2B). We conducted genomic DNA polymerase chain reaction and Sanger sequencing and identified the genomic breakpoints and joining points for these two samples containing *FGFR3-TACC3* fusion transcripts (Data Supplement). Given these observations, we further probed the TCGA lung SCC RNA sequencing data set[15] and identified four additional patients who had tumors containing *FGFR3-TACC3* in-frame fusion transcripts of 178 tumor RNA specimens (Data Supplement), further confirming that these fusion events are recurrent in lung SCC. All six patients with fusion transcripts involved the omission of the last exon of *FGFR3*, which was instead fused in-frame to various exons of *TACC3* and accompanied by an increase in *FGFR3* transcript level (Data Supplement).

Comparative Mutational Analysis of Lung SCCs in East Asian and Mostly White Patients From the United States

By using the set of genomic alterations in this lung SCC cohort, we investigated the association among the spectrum of somatic mutations with clinical parameters. Most of the major somatic mutations did not have a statistical association with overall survival (Data Supplement). We did observe that some mutations or SCNAs were significantly associated with poor survival (Data Supplement). The biologic significance of alterations in genes with apparent prognostic value, such as *MLL2* mutations, requires further experimental analysis and validation in additional independent cohorts.

Given that no systematic comparison has yet been reported for lung SCC among different ethnicities, we compared the spectra of mutational and SCNAs of this cohort to a similarly sized cohort recently reported by the TCGA,[15] consisting mostly of white patients from the United States (Table 1). Both cohorts displayed a similarly high rate of mutations of *TP53* (73% Korean, 81% TCGA; Fisher's exact $P > .16$), mutations in the *NFE2L2/KEAP1* genes (33% and 31%; $P > .85$), and those in *PIK3CA/PTEN* (24% and 24%) and *NOTCH1* (7% and 8%; $P > .9$). Both cohorts also demonstrated enrichment for CpG transition and transversion mutations. SCNAs involving *SOX2, FGFR1, KDR/PDGFRA/KIT,* and *PTEN* were also observed at comparable frequencies. One notable difference among

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

© 2013 by American Society of Clinical Oncology

Fig 1. Clinicopathologic characteristics and genetic aberrations across 104 lung squamous cell carcinomas. (A) Rate of synonymous and nonsynonymous mutations, expressed in mutations per megabase (Mb) of covered target sequence in association with clinical parameters for the samples. In the bottom panel, the first row indicates smoking status, the second row sex, and the third row cancer stage. (B) Heatmap representation of individual mutations present in 104 lung squamous cell carcinoma samples in association with information from (A). It displays the mutational types in a given sample and in a given gene in a two-dimensional matrix. (Left) Percentages of mutations, representing at least one mutational event in each gene. (Right) List of recurrently mutated genes. *FGFR, RAS, SWI/SNF, NOTCH,* and *WNT* mini-gene sets are shown as single tracks because there was complete exclusivity of mutations within these gene families. *FGFR** indicates conglomerated representation of *FGFR1 + FGFR3, RAS**: *HRAS + NRAS + KRAS, SWI/SNF**: *ARID1A + ARID1B + PBRM1, NOTCH**: *NOTCH1 + NOTCH3, WNT**: *APC + CTNNB1*. (C) Somatic copy-number alterations (SCNAs) of samples in association with (A) and (B). Significant SCNAs are shown, highlighting the samples with threshold copy-number changes specified in the lower bottom panel. Indel, insertions and deletions.

the data sets was somatic alterations in cell cycle regulatory genes, including *CDKN2A* and *RB1*. In the TCGA cohort, *CDKN2A* mutations were observed in 15% of patients and homozygous deletions in 33%. In contrast, *CDKN2A* mutations and copy-number alterations were infrequent in the Korean cohort (4% and 5%, respectively; $P < .007$ and $P < .001$, respectively); however, mutations of *RB1* were more common in the Korean cohort (15%) than in the TCGA cohort (4%; Fisher's exact $P < .005$).

Both our data and the TCGA data set[15] confirmed previous reports suggesting that the two most common oncogenic mutations in lung adenocarcinomas, *KRAS* and *EGFR*,[34] are rare in lung SCCs. We identified only one case with an activating *EGFR* mutation in this cohort, an exon 19 deletion, and no cases of L861Q, which was observed in the TCGA cohort.[15] We noted two cases of *KRAS* G12C but no codon 12 or 13 *RAS* events were reported by TCGA (Data Supplement). Alterations in other *RAS* family members such as *NRAS* and

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

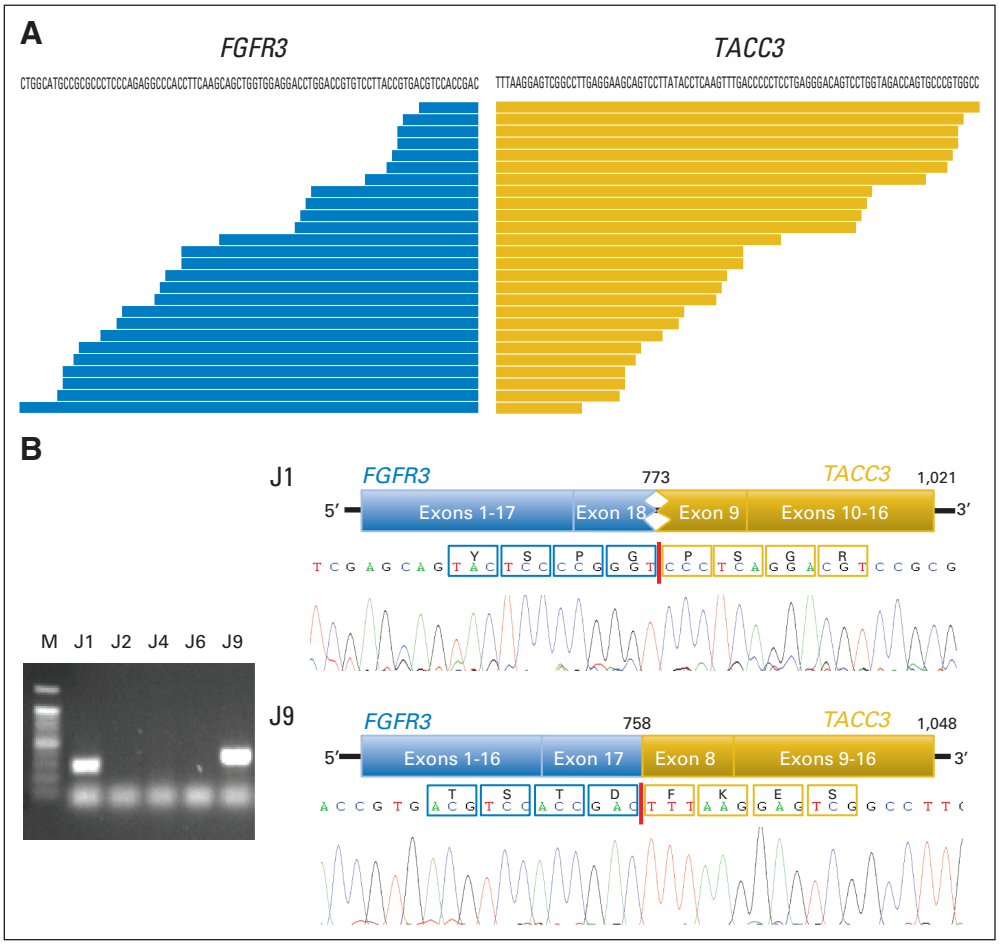

Fig 2. Recurrent *FGFR3-TACC3* fusion identified in lung squamous cell carcinoma. (A) *FGFR3-TACC3* fusion identified from RNA sequencing; 26 split-reads that span the *FGFR3-TACC3* junction are shown. (B) Left: *FGFR3-TACC3* fusion-specific polymerase chain reaction from complementary DNA derived from patients with lung squamous cell carcinoma. Right: sequencing chromatogram from two patient samples spanning the fusion junction. Shown in box are the translational reading frames and designation of amino acids. M, DNA marker. J1, patient identification number.

HRAS were virtually absent in this cohort, although putative loss-of-function events in negative regulators of *RAS* such as *RASA1* and *NF1* were observed at appreciable frequencies (7% and 13%, respectively; summarized in the Data Supplement).

Closer examination of mutational profiles revealed core signaling pathways implicated in discrete functional categories. As in the TCGA cohort analysis, a transcription factor pathway regulating redox stress comprising *NFE2L2*, *KEAP1*, and *CUL3* aberrations was altered in 40% of the patients (Fig 3). Another pathway implicated was deregulated chromatin remodeling with alterations in chromatin remodeling enzymes, including histone demethylases and *SWI-SNF* complexes identified in more than 49% of patients (Fig 3). Some members in this group of core pathways demonstrated mutually exclusive mutational patterns (Fig 3). An exception to this exclusivity relationship was the strict co-occurrence of *MLL2* and *KDM6A* mutations (Fisher's exact $P < .001$). It is of note that members of these two core pathways were also likely to interact functionally, given the observation that mutations in *NFE2L2* and *ARID1A* or *ARID1B* co-occur ($P < .001$; Data Supplement). Squamous cell differentiation pathways involving inactivating mutations of *NOTCH* genes were also frequently altered[15,20] (Fig 3). *FGFR* kinase amplification and mutations were frequent and were identified in 18 individuals, and *PI3K* was the most commonly altered serine/threonine kinase family with mutational events in 17 individuals (Fig 3).

DISCUSSION

This study has identified a potentially targetable fusion gene in lung SCCs, *FGFR3-TACC3*, which defines a new specific subset for potential targeted therapy in lung SCC and was recently reported by others.[32,33] Previously, FGFR small-molecule inhibitors were shown to possess therapeutic activity in lung SCC cells with *FGFR* family kinase amplification and mutation.[14,35] Clinical activity and safety of BGJ398 in *FGFR1*-amplified lung SCC and AZD4547 in various types of solid tumors have recently been demonstrated.[36,37] The FGFR antagonists are also under clinical investigation in lung SCCs with mutations or amplification of *FGFR1*, *FGFR2*, or *FGFR3* as more diversified inclusion criteria (clinical trials NCT01761747 and NCT01795768). Preclinical model studies of *FGFR3-TACC3* in glioblastoma and bladder cancer demonstrated encouraging utility of FGFR inhibitors in the treatment of cancer cells harboring the fusion gene.[29-31] Thus, it is likely that clinical trials could be expanded to test the efficacy of these agents in patients with *FGFR3* fusions in lung SCC with concurrent development of companion diagnostics. Despite apparently modest demographic influences in lung SCC, this study also points out the value of analyzing different cohorts, since the TCGA study initially missed the identification of the *FGFR3* fusion events.[15]

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

© 2013 by American Society of Clinical Oncology

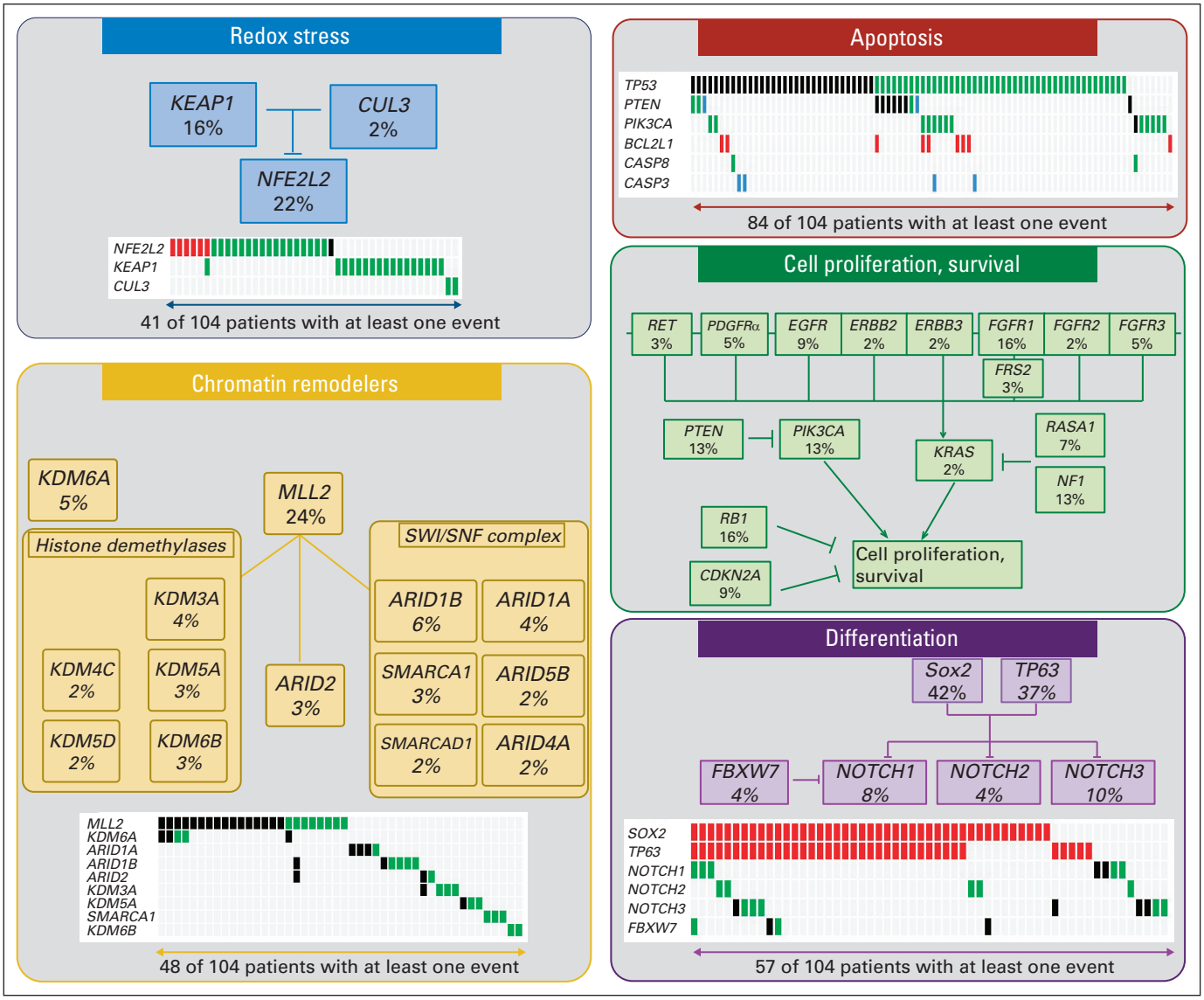

Fig 3. Diagram of integrated lung squamous cell carcinoma pathways. Proposed diagram integrating five different core pathways of lung squamous cell carcinoma. Green, substitution mutations; black, truncating mutations; red, amplifications; blue, deletions.

In contrast to lung adenocarcinomas in which the frequency of genomic alterations in oncogenes differs among ethnic groups,[34,38] in lung SCC we found that the mutational landscape of two different population groups with similar sized samples had overall congruent mutational frequencies in major oncogenic drivers and tumor suppressors. To the best of our knowledge, this is the first and largest comparison of the integrative genomic analysis of a given cancer type with more than 100 individuals between populations of Asians and mostly white patients from the United States. Although the frequency of mutations in *RB1* and *CDKN2A* genes were significantly different between the two cohorts, frequencies of other major genes, including *TP53*, *NFE2L2*, *KEAP1*, *PIK3CA*, *PTEN*, *MLL2*, and *NOTCH1*, were similar to each other. Thus, in lung SCC, in which most of the patients had been chronically exposed to heavy doses of tobacco, the subtle genetic and inherent biologic influences seem to be largely overridden by heavy mutational burdens from exogenous carcinogens. It has been shown that several cancer types, including melanoma, display highly increased mutational burden with exogeneous carcinogens linked to their tumorigenesis. Systematic comparative genomic profiling of this and other cancer types with high mutational rates in patients in different ethnic groups will help address the generalizability of the findings.

Although most SCCs occur in smokers, a small subset of never-smoking patients (five of 104 for this study; seven of 178 for TCGA) still succumb to the disease. We did not capture information regarding exposure to second-hand smoke or environmental toxins such as asbestos in this study, and smoking status was self-reported. We believe that larger-scale genomic studies specifically targeting these subsets will help to better elucidate underlying biologic mechanisms associated with the never-smoker lung SCC subgroup. In this sense, it is of note that ongoing large-scale cancer consortium studies, including that of TCGA, currently aim to delineate genomic aberrations in more than 500 patients for each subtype of cancer in the coming years. The sample size used in this study (n = 104; n = 178 for TCGA) and the lower degree of sequencing coverage compared with that reported

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

in TCGA limit the ability to perform comparative analyses of less common clinical features or genomic alterations with sufficient statistical power. Given the complexities of the cancer genomic landscape, it is likely that expansion of patient samples in the future will help to further clarify subtle genetic differences in lung SCC contributed by ethnicity. It should also be noted that demographic heterogeneity within the TCGA cohort was not well annotated, which prevented the possible exclusion of patients of Asian descent in this cohort for our comparative analyses.

In summary, this study has comparatively described the recurrent somatic mutations in lung SCCs in a large cohort of East Asians. It revealed the commonality in the pathogenesis of lung SCC across ethnicities and identified several targetable oncogenic mutations in a disease with large unmet medical needs.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** Matthew D. Wilkerson, GeneCentric Diagnostics (C); Matthew Meyerson, Bayer Pharmaceuticals (U), Foundation Medicine (C) **Stock Ownership:** Matthew Meyerson, Foundation Medicine **Honoraria:** None **Research Funding:** None **Expert Testimony:** None **Patents:** Michael S. Lawrence, MutSigCV; Matthew Meyerson, patent US20060147959 **Other Remuneration:** Michael S. Lawrence, MutSigCV

AUTHOR CONTRIBUTIONS

Conception and design: Youngwook Kim, Myung-Ju Ahn, Matthew Meyerson, Keunchil Park
Provision of study materials or patients: In-Jae Oh, Hong Kwan Kim, Yong Soo Choi, Kwhanmien Kim, Young Mog Shim, Kyu-Sik Kim, Sang-Yun Song, Kook-Joo Na, Yoon-La Choi, Jhingook Kim, Sukki Cho, Young-Chul Kim, Keunchil Park
Collection and assembly of data: Youngwook Kim, Peter S. Hammerman, Jaegil Kim, Ji-ae Yoon, Yoomi Lee, Yeong Kyung Yoo, In-Jae Oh, Hong Kwan Kim, Yong Soo Choi, Kwhanmien Kim, Young Mog Shim, Kyu-Sik Kim, Sang-Yun Song, Kook-Joo Na, Jhingook Kim, Sukki Cho, Young-Chul Kim, Jin Seok Ahn, Gad Getz, Matthew Meyerson, Keunchil Park
Data analysis and interpretation: Youngwook Kim, Peter S. Hammerman, Jaegil Kim, Ji-ae Yoon, Yoomi Lee, Jong-Mu Sun, Matthew D. Wilkerson, Chandra Sekhar Pedamallu, Kristian Cibulskis, Yeong Kyung Yoo, Michael S. Lawrence, Petar Stojanov, Scott L. Carter, Aaron McKenna, Chip Stewart, Andrey Y. Sivachenko, Yoon-La Choi, D. Neil Hayes, Gad Getz, Matthew Meyerson, Keunchil Park
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. World Health Organization: Cancer Fact Sheet No. 297. http://www.who.int/mediacentre/factsheets/fs297/en/
2. Kenfield SA, Wei EK, Stampfer MJ, et al: Comparison of aspects of smoking among the four histological types of lung cancer. Tob Control 17:198-204, 2008
3. Paez JG, Jänne PA, Lee JC, et al: EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. Science 304:1497-1500, 2004
4. 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
5. Kwak EL, Bang YJ, Camidge DR, et al: Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 363:1693-1703, 2010
6. Maemondo M, Inoue A, Kobayashi K, et al: Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 362:2380-2388, 2010
7. Pao W, Chmielecki J: Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. Nat Rev Cancer 10:760-774, 2010
8. Okudela K, Suzuki M, Kageyama S, et al: PIK3CA mutation and amplification in human lung cancer. Pathol Int 57:664-671, 2007
9. Shibata T, Ohta T, Tong KI, et al: Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proc Natl Acad Sci U S A 105:13568-13573, 2008
10. Kan Z, Jaiswal BS, Stinson J, et al: Diverse somatic mutation patterns and pathway alterations in human cancers. Nature 466:869-873, 2010
11. Hammerman PS, Sos ML, Ramos AH, et al: Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. Cancer Discov 1:78-89, 2011
12. Bass AJ, Watanabe H, Mermel CH, et al: SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet 41:1238-1242, 2009
13. Ramos AH, Dutt A, Mermel C, et al: Amplification of chromosomal segment 4q12 in non-small cell lung cancer. Cancer Biol Ther 8:2042-2050, 2009
14. 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
15. Cancer Genome Atlas Research Network: Comprehensive genomic characterization of squamous cell lung cancers. Nature 489:519-525, 2012
16. Eberhard DA, Johnson BE, Amler LC, et al: Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. J Clin Oncol 23:5900-5909, 2005
17. Shigematsu H, Lin L, Takahashi T, et al: Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 97:339-346, 2005
18. Mok TS, Wu YL, Thongprasert S, et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 361:947-957, 2009
19. National Cancer Information Center: Cancer Information Service. http://www.cancer.go.kr/ncic/cics_b/01/013/1640957_5873.html
20. Stransky N, Egloff AM, Tward AD, et al: The mutational landscape of head and neck squamous cell carcinoma. Science 333:1157-1160, 2011
21. Benjamini Y, Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B (Methodol) 57:289-300, 1995
22. Bentley DR, Balasubramanian S, Swerdlow HP, et al: Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53-59, 2008
23. Chapman MA, Lawrence MS, Keats JJ, et al: Initial genome sequencing and analysis of multiple myeloma. Nature 471:467-472, 2011
24. Cibulskis K, McKenna A, Fennell T, et al: ContEst: Estimating cross-contamination of human samples in next-generation sequencing data. Bioinformatics 27:2601-2602, 2011
25. Cibulskis K, Lawrence MS, Carter SL, et al: Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 31:213-219, 2013
26. Beroukhim R, Getz G, Nghiemphu L, et al: Assessing the significance of chromosomal aberrations in cancer: Methodology and application to glioma. Proc Natl Acad Sci U S A 104:20007-20012, 2007
27. Mermel CH, Schumacher SE, Hill B, et al: GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 12:R41, 2011
28. Berger MF, Lawrence MS, Demichelis F, et al: The genomic complexity of primary human prostate cancer. Nature 470:214-220, 2011
29. Parker BC, Annala MJ, Cogdell DE, et al: The tumorigenic FGFR3-TACC3 gene fusion escapes

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

© 2013 by American Society of Clinical Oncology

miR-99a regulation in glioblastoma. J Clin Invest 123:855-865, 2013

30. Singh D, Chan JM, Zoppoli P, et al: Transforming fusions of FGFR and TACC genes in human glioblastoma. Science 337:1231-1235, 2012

31. Williams SV, Hurst CD, Knowles MA: Oncogenic FGFR3 gene fusions in bladder cancer. Hum Mol Genet 22:795-803, 2013

32. Wu YM, Su F, Kalyana-Sundaram S, et al: Identification of targetable FGFR gene fusions in diverse cancers. Cancer Discov 3:636-647, 2013

33. Majewski IJ, Mittempergher L, Davidson NM, et al: Identification of recurrent FGFR3 fusion genes in lung cancer through kinome-centred RNA sequencing. J Pathol 230:270-276, 2013

34. Ding L, Getz G, Wheeler DA, et al: Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455:1069-1075, 2008

35. Dutt A, Ramos AH, Hammerman PS, et al: Inhibitor-sensitive FGFR1 amplification in human non-small cell lung cancer. PLoS One 6:e20351, 2011

36. Andre F: Results of a phase I study of AZD4547, an inhibitor of fibroblast growth factor receptor (FGFR), in patients with advanced solid tumors. 104th Annual Meeting of the American Association for Cancer Research, Washington, DC, April 6-10, 2013 (abstr LB-145)

37. Wolf J, LoRusso PM, Camidge RD, et al: A phase I dose escalation study of NVP-BGJ398, a selective pan FGFR inhibitor in genetically preselected advanced solid tumors. 103th Annual Meeting of the American Association for Cancer Research, Chicago, IL, March 31-April 4, 2012 (abstr LB-122)

38. Sun Y, Ren Y, Fang Z, et al: Lung adenocarcinoma from East Asian never-smokers is a disease largely defined by targetable oncogenic mutant kinases. J Clin Oncol 28:4616-4620, 2010

Affiliations

Youngwook Kim, Ji-ae Yoon, Yoomi Lee, Yeong Kyung Yoo, and Keunchil Park, Samsung Medical Center, Samsung Biomedical Research Institute; Jong-Mu Sun, Hong Kwan Kim, Yong Soo Choi, Kwhanmien Kim, Young Mog Shim, Yoon-La Choi, Jhingook Kim, Jin Seok Ahn, Myung-Ju Ahn, and Keunchil Park, Samsung Medical Center, Sungkyunkwan University School of Medicine; Sukki Cho, Seoul National University Bundang Hospital, College of Medicine, Seoul National University, Seoul; In-Jae Oh, Kyu-Sik Kim, Sang-Yun Song, Kook-Ju Na, and Young-Chul Kim, Chonnam National University Hwasun Hospital, Jeonnam, Republic of Korea; Peter S. Hammerman, Jaegil Kim, Chandra Sekhar Pedamallu, Kristian Cibulskis, Michael S. Lawrence, Petar Stojanov, Scott L. Carter, Aaron McKenna, Chip Stewart, Andrey Y. Sivachenko, Gad Getz, and Matthew Meyerson, Broad Institute of Harvard and MIT, Cambridge; Matthew Meyerson, Harvard Medical School; Peter S. Hammerman and Chandra Sekhar Pedamallu, Dana-Farber Cancer Institute; Gad Getz, Massachusetts General Hospital Cancer Center, Boston, MA; and Matthew D. Wilkerson and D. Neil Hayes, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

Acknowledgment

We thank William Hahn, MD, PhD (Dana-Farber Cancer Institute) for his generous advice and input; Macrogen for sequencing efforts; Theragen Bio for quality control efforts; C. Maeng, H. Jung, S. Park, J. Cha, W. Shin, and E. Lee (Samsung Medical Center) for their valuable input into the project; Y.S. Lee for her excellent graphical help; and PGM21 and its director, H. Kim, MD, for their continued support of the project.

Downloaded from ascopubs.org by Massachusetts Institute of Technology (MIT) on May 16, 2024 from 018.029.117.013
Copyright © 2024 American Society of Clinical Oncology. All rights reserved.

© 2013 by American Society of Clinical Oncology