



Published in final edited form as:

*J Thorac Oncol.* 2015 February ; 10(2): 250–255. doi:10.1097/JTO.0000000000000420.

## Racial diversity of actionable mutations in non-small cell lung cancer

**Aliccia Bollig-Fischer, PhD, Wei Chen, PhD, Shirish M. Gadgeel, MD, Angela S. Wenzlaff, MPH, Michele L. Cote, PhD, Ann G. Schwartz, PhD, MPH, and Gerold Bepler, MD, PhD**  
Barbara Ann Karmanos Cancer Institute and Department of Oncology, Wayne State University, Detroit, Michigan, United States of America

### Abstract

**Introduction**—Lung cancer is the leading cause of cancer-related deaths in the United States. The reasons for higher incidence and poorer survival rates among black compared to white lung cancer patients have not been defined. We hypothesized that differential incidence of somatic cancer gene mutations may be a contributing factor. Previous genomic studies of non-small cell lung cancer (NSCLC) have not adequately represented black patients.

**Methods**—A MALDI-TOF mass-spectrometry approach was used to analyze tumor DNA for 214 coding mutations in 26 cancer genes previously identified in NSCLC. The samples included NSCLC from 335 white patients and 137 black patients. For 299 of these, normal matched DNA was available and analyzed.

**Results**—*EGFR* exon 19 deletions were only detected in female cases, with increased odds for black women compared to white women (odds ratio=3.914, 95% CI: 1.014–15.099, p=0.048). Beyond race, variations in mutation frequencies were seen by histology. *DDR2* alterations, previously described as somatic mutations, were identified as constitutional variants.

**Conclusions**—This study is among the largest comparing somatic mutations in black and white patients. The results point to the molecular diversity of NSCLC and raise new questions as to the importance of inherited alleles. Genomic tumor testing will benefit both populations, although the mutation spectrum appears to vary by sex, race, and histology.

### Keywords

non-small cell lung cancer; African American; mutations; oncogenes; tumor suppressors

### Introduction

Lung cancers are a heterogeneous group of tumors traditionally categorized by histology, with the majority classified as non-small cell lung cancer (85%, NSCLC)(1). NSCLC is further subdivided into adenocarcinomas (~45%), squamous cell carcinomas (~23%), and

---

**Corresponding author:** (Aliccia Bollig-Fischer, 4100 John R, Detroit, Michigan 48201; bollig@karmanos.org; ph: 313-577-8671, fax: 313-577-6200).

**Conflicts of Interest:** The authors do not have any relationships that could be construed as resulting in an actual, potential, or perceived conflict of interest.

large cell carcinomas (~3%), with other subtypes representing the remaining approximate 28% (1). Large-scale gene mutation profiling is transforming our knowledge of the heterogeneity of NSCLC(2). Genomic research now links specific oncogenes and recurring mutations to the disease phenotype and provides a rationale for use of molecularly targeted treatments that are improving lung cancer patients' outcomes (2).

Black Americans are more likely to develop lung cancer and at an earlier age compared to white individuals despite lower rates of smoking in black adolescents (3–7). Moreover, blacks with lung cancer show worse outcomes, including shorter overall survival and increased mortality, which persist when correcting for socioeconomic factors and unequal access to care (7). These data suggest diversity in disease etiology across populations. The specifics of disease etiology govern the mutational processes giving rise to particular mutation patterns or signatures in cancer (8). Along these lines, work by Alexandrov et al. demonstrates that a specific mutation profile in lung cancers is attributable to smoking (8). The same study also reports distinct mutation signatures in lung cancers not yet linked to a cause. It is predicted that more causally-linked cancer mutation signatures will emerge when both the numbers of specimens and diversity of patients under investigation increases. Thus, we hypothesized that the frequencies of specific mutations in NSCLC will vary between black and white populations. Previous studies have compared mutation frequencies according to race, but largely focused on a small subset of oncogenes – e.g., mutations in epidermal growth factor receptor (EGFR) and kirsten rat sarcoma viral oncogene homolog (KRAS) – and they yielded conflicting results (9–13). More comprehensive research is required to thoroughly investigate cancer gene mutation differences and to derive insight as to diagnostics and individualized treatment modalities that provide patient benefit. Here, we present findings from a large-scale genomic study, with a large proportion of black patients, to examine whether the landscape of cancer relevant mutation frequencies varies according to race.

## Materials and Methods

### Patients and Tissues

Biospecimen collection and outcomes data compilation were done according to the Helsinki Declaration and approved by the Wayne State University School of Medicine Institutional Review Board. Fresh-frozen or formalin fixed paraffin embedded (FFPE) specimens were collected from patients who underwent a surgical resection for diagnosed or suspected lung cancer. Frozen specimen procurement procedures were implemented to reduce the resection to freezing time interval to less than 30 minutes. The overall procurement period ranged from 1985 to 2012 from 5 different case series. Frozen and archived FFPE specimens were reviewed to verify NSCLC diagnosis and to determine that tumor cell content was equal or greater than 70%. Only tumors pathologically confirmed to be NSCLC were included in the analysis. DNA from adjacent normal lung tissue was available and analyzed from 299 of the 472 cases. Demographic and clinical outcomes data collected included: the dates of birth, diagnosis, and last follow-up or death; sex, race, tumor histology, pathological and clinical tumor stage (AJC staging manual, version 6.0) and self-reported smoking history (defined as

life-time never smoker for those who had smoked <100 cigarettes, former smoker for those who had quit cigarette smoking for more than one year, and smoker for all others).

### Genetic Analysis

Frozen NSCLC and normal tissue specimens (~100 mg) were pulverized using sterilized and frozen mortar and pestles and DNA was isolated using resin or phenol-based extraction methods. DNA was isolated from paraffin-embedded tissue using EZ1 Advanced magnetic bead technology, the EZ1 DNA de-paraffinization method and the EZ1 DNA Tissue kit (Qiagen, Valencia, CA). All sample DNA quantity and quality was assessed using a Nanodrop spectrophotometer and Quantifiler® assay (Life Technologies, Carlsbad, CA), a real-time polymerase chain reaction (PCR)-based approach for estimation of amplifiable DNA. A standard curve of known DNA concentrations was also analyzed for quantitation of each DNA sample.

Mutations were analyzed using the Sequenom MassARRAY® System employing Matrix-Assisted Laser Desorption/Ionization and Time of Flight (MALDI-TOF) mass spectrometry and the Sequenom LungCarta™ panel (Sequenom, San Diego, CA). The panel targets 214 sequence mutations in 26 oncogenes and tumor suppressors, which were previously identified in NSCLC.(14, 15) The genes include: AKT1, ALK, BRAF, DDR2, EGFR, EPHA3, EPHA5, ERBB2, FGFR4, JAK2, KRAS, MAP2K1, STK11, MET, NOTCH1, NRAS, NRF2, NTRK1, NTRK2, NTRK3, PIK3CA, PTCH1, PTEN, PTPN11, PTPRD and *TP53*. The coding mutation types include synonymous and nonsynonymous nonsense and missense point mutations, transversions and transitions, and short insertions and deletions(Supplementary Table 1). The methodology starts with PCR of small DNA segments (<100 base pairs) encompassing the DNA mutation sites in a multiplex reaction. A follow-up extension reaction adds a single base to extension primers, the nature of the base pair added to a primer affects its mass and time-of-flight; thus, a wild-type and mutation sequence can be differentiated and quantified. The sensitivity of the approach allows for detection of a mutation that represents 10% of the sample. The assay precision and accuracy for both FFPE and frozen specimen DNA were validated for as little as 60 ng by us and others (16); however, typically 480 ng of sample DNA was analyzed for this study.

### Statistical Analysis

Descriptive statistics were provided for patients' demographic and clinical characteristics. Multivariable logistic regression was used to estimate the odds of having the genetic alteration, adjusted for potential confounders such as race (Black, reference: White), age (as continuous variable), sex (Female, reference: Male), tumor stage (II, III or IV, reference: I), histology (Squamous, other NSCLC, reference: Adenocarcinoma) and smoking status (Never, unknown, reference: Ever). All statistical tests were traditional two sided at a significance level of 0.05. Given the nature of this exploratory study, p values were not adjusted for multiple testing. SNP allele frequency within each subgroup of race was compared to the National Institutes of Health Heart, Lung and Blood Institute (NHLBI) Grand Opportunity (GO) Exome Sequencing 6,500 Project (ESP) database(17) using Fisher's exact test. The statistical software R 3.0.0 [The R foundation for statistical computing; <http://www.r-project.org/foundation/>] was used for all analyses.

## Results

### Patient Characteristics

Descriptive statistics for the 472 NSCLC patients – 137 black individuals and 335 white individuals – who contributed specimens to this study are listed in Table 1. The proportions for histology, sex, stage and smoking history are given stratified by race; median age and pack years smoked by race are also indicated.

### Frequency of Cancer Gene Mutations According to Race

An initial overview of the data from mass spectrometry analysis showed that for 180 of the 472 (38%) NSCLC specimens a mutation was detected (characteristics and mutations identified for each specimen are provided in the on-line Supplementary Data file). Comparing frequencies for individual mutations across race showed no significant differences for point mutations; however, there was a higher frequency for the *EGFR* E746-T751>S deletion in exon 19 in tumor specimens from black patients compared to white patients ( $p=0.076$ ; Table 2). When the combined frequency of all variations of deletions detected at E746 (i.e., E746-T751 and E746-T750) were considered, we observed that these mutations occurred exclusively in females and more often in black women than in white women (OR 3.914; 95% CI [1.014–15.099];  $p=0.048$ ), after adjusting for age, stage, histology, race and smoking status (Table 3). Across all mutation types, tumors from women were 80% more likely to carry at least one mutation than those from men after adjustment for smoking, age, race, stage and histologic type (OR 1.815; 95% CI [1.200–2.747];  $p=0.005$ ) (Supplemental Table 2).

We also observed a race difference when we considered the numbers of specimens showing no mutation in black and white groups – 68% versus 59%, respectively. Having adjusted for covariates including smoking, sex, stage, histology and age, which are known determinates for the over-all extent of genomic mutation in cancer specimens (18), the probability of having a mutation was decreased for the black patients (OR 0.582; 95% CI [0.363–0.933];  $p=0.025$ ) (Supplementary Table 2). Figure 1 summarizes the number of single and multiple mutations observed in NSCLC specimens from black and white patients. Given the small numbers of multiple mutations, correlations cannot be tested for or meaningfully reported; however, mapping of co-occurring mutations displays a greater diversity of known mutations in tumors from white patients, which is consistent with overall significantly more known mutations having been identified in NSCLC specimens from white patients.

### Distinguishing Constitutional Polymorphisms from Somatic Mutations

In an overview of the data from mass spectrometry analysis, we identified that there were significant differences across race for genetic alterations in tumors that are also present in germline DNA, representing previously described single nucleotide polymorphisms (SNPs) in *c-met* (*MET*), *notch 1* (*NOTCH1*) and *serine/threonine kinase 11* (*STK11*). The differentially detected SNPs included *NOTCH* p.V1671I and *STK11* p.Y272Y (both increased in blacks;  $p<0.0001$ ), and *MET* p.N375S (increased in whites;  $p=0.0117$ ). Using the National Institutes of Health Heart, Lung and Blood Institute (NHLBI) Grand Opportunity (GO) Exome Sequencing 6,500 Project (ESP) database,(17) we observed that

all three SNPs demonstrated similar differences in comparing black and white frequencies in the general U.S. population. However, *NOTCH1* p.V1671I (p=0.01), and *MET* p.N375S (p=0.076) appear to be more frequent in the cancer patient population (Supplementary Table 3). All of these SNPs are in the National Center for Biotechnology Information (NCBI) SNP database (dbSNP) although they have previously been highlighted as mutations in the literature or COSMIC database (15, 19–22).

We proceeded to analyze matched normal lung DNA from 299 cases to confirm these SNPs and to identify other variants that might be inherited. Our efforts identified 7 constitutional polymorphisms in the dataset, described in Table 4. Included among them was *EGFR* p.R776H, an activating mutation only recently identified as inherited in a single case study of a nonsmoking mother-daughter pair diagnosed with lung cancer (20). In the present study we identified it in normal and corresponding malignant tissue from a single white patient in our cohort of 472 (0.2%) patients. Also of note, two *DDR2* SNPs were detected in matched normal and tumor specimens: *DDR2* p.L63V in a white patient and *DDR2* p.G505S in a black patient. Both variants were previously characterized as activating mutations in NSCLC (15).

### Mutations Associated with Histological Subtype

According to our data, mutation frequencies were lower in squamous cell carcinoma than in adenocarcinoma (OR 0.583; 95% CI [0.374–0.909]; p=0.017). Further investigation of the frequency of somatic mutations in our 472 patient cohort according to histology highlighted potential differences specifically for *PIK3CA* p.E545K (increased rate in squamous cell carcinoma, p=0.029) and *KRAS* p.G12C/V (increased rate in adenocarcinoma, p<0.001 and p=0.016) (Supplementary Table 4). In general, adenocarcinomas exhibited greater numbers and higher diversity of mutations compared to squamous cell carcinomas (Figure 2). It is noted that the level of diversity for co-occurring mutations discovered in squamous cell carcinoma appears to resemble the relative level of diversity observed for black patients. Any observation of similarity is potentially misleading and not consistent with a histology bias in the study population; where to the contrary, squamous cell cancer is significantly less represented in our black patient cohort compared to adenocarcinomas (21% versus 67%, p<0.001).

### Discussion

Certain epidemiological data establishing disparities in NSCLC occurrence in black versus white patients are also suggestive of differences in disease etiology (3–7), and provide reasonable support for the notion that mutation profiles from the two populations will be different. Our study analyzed a broad variety of sequence mutations in NSCLC-linked genes and showed that the frequencies of coding mutations were within previously reported ranges for NSCLC (adenocarcinoma and squamous cell combined: *KRAS* 11%, *EGFR* 8% and *PIK3CA* 2%); and their frequencies were similar for black and white groups. The exception to this rule is the frequency of small deletions in *EGFR* exon 19 at p.E746. Results showed the deletion was only detected in lung tumors from women, and with increased frequency in lung tumors from black women. The data suggest that women and black women more

specifically may benefit from testing for this lesion, which is a predictive biomarker of response to anti-EGFR therapy. The clinical importance of a small deletion in *EGFR* exon 19 is further underscored by early evidence linking the aberration to longer progression free survival relative to other EGFR mutations in metastatic cancer treated with first-line EGFR tyrosine kinase inhibitors (23).

The types and or signatures of mutations that define cancer subtypes are attributable to causative mutational processes and evidence surrounding small deletions like p.E746–751 or p.E746–750 identified in *EGFR* exon 19 points specifically to defects in the DNA mismatch repair process (24). Furthermore, significant differences in the numbers of specimens for black patients showing no coding mutation, 68% versus 59% for whites, indicates that further genomic analyses should be done to more comprehensively characterize the range of mutation types occurring in NSCLC of black patients. For example, analysis of a larger gene set or copy number variation analysis may increase the percentage of tumors where an oncogene mutation is discovered. Whole genome sequencing of tumors from black patients may also identify novel genetic alterations in this population.

Our study is only the second to report a rare EGFR single nucleotide germ line variant previously reported in a single case study to be oncogenic (20). Also, ours is the first report of the discovery of *DDR2* variants in both normal and NSCLC tissue from individual patients. Specifically, the *DDR2* p.G505S and *DDR2* p.L63V variants are reported to be activating mutations in the cancer literature — mutations that supposedly render a cancer cell sensitive to the anti-growth effects of the targeted drug dasatinib (15). It cannot be determined from the present study what bearing these inherited *DDR2* variants may have had on cancer initiation and outcomes. Their discovery brings to mind ongoing efforts to understand to what degree inherited *EGFR* p.T790M mutations lead to an increased risk of developing lung cancer (25–27); the same consideration must be given to *DDR2* variants.

Finally, our analyses also honed in on histology-associated differences for individual mutation frequencies, adding to the growing body of evidence and emerging understanding that histological subtypes are distinguished by the mutations they harbor (2, 28, 29). It was observed that adenocarcinomas exhibited greater numbers and greater diversity of mutations, and more often co-occurring mutations compared to squamous cell carcinomas (Figure 2). It should not be interpreted that squamous cell carcinoma is less apt to harbor driver mutations; on the contrary the spectrum of cancer driver gene mutations associated with squamous cell histology is anticipated to be different. Thus, pointing to a need for future work to engage expanded genomic analyses to identify the full spectrum of cancer gene lesions in lung cancer subtypes. The same notions apply to our analyses across race, where fewer mutations were identified in black patient samples.

In summary, this study designed to test for differences in known somatic coding mutation frequencies across populations, presents a number of novel and unexpected findings. The important significant findings associated with race support the need for further study to molecularly define NSCLC in diverse patient populations. The particular discovery that *DDR2* variants (*DDR2* p.G505S and *DDR2* p.L63V) classify as inherited alleles forces the need for further evaluation of their occurrence and oncogenic role in NSCLC.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to thank the patients who consented to provide tumor material that was used in this study.

**Sources of Funding:** U.S. Department of Defense grant W81XWH-11-1-0050, the Herrick Foundation and NIH grants/contracts P30CA022453, R01CA060691, and R01CA87895. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## References

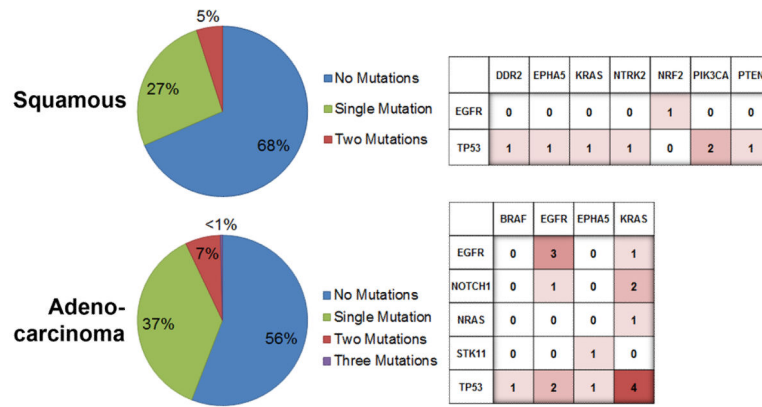
1. Howlander, NNA.; Krapcho, M.; Neyman, N.; Aminou, R.; Waldron, W.; Altekruze, SF.; Kosary, CL.; Ruhl, J.; Tatalovich, Z.; Cho, H.; Mariotto, A.; Eisner, MP.; Lewis, DR.; Chen, HS.; Feuer, EJ.; Cronin, KA.; Edwards, BK. SEER Cancer Statistics Review, 1975–2008. 2011. [cited based on November 2010 SEER data submission, posted to the SEER web site, 2011; Available from: [http://seer.cancer.gov/csr/1975\\_2008](http://seer.cancer.gov/csr/1975_2008)
2. A genomics-based classification of human lung tumors. *Science translational medicine*. 2013; 5:209ra153.
3. Anderson C, Burns DM. Patterns of adolescent smoking initiation rates by ethnicity and sex. *Tobacco control*. 2000; 9(Suppl 2):II4–8. [PubMed: 10841585]
4. Hardy D, Liu CC, Xia R, Cormier JN, Chan W, White A, et al. Racial disparities and treatment trends in a large cohort of elderly black and white patients with nonsmall cell lung cancer. *Cancer*. 2009; 115:2199–211. [PubMed: 19365824]
5. Gadgeel SM, Severson RK, Kau Y, Graff J, Weiss LK, Kalemkerian GP. Impact of race in lung cancer: analysis of temporal trends from a surveillance, epidemiology, and end results database. *Chest*. 2001; 120:55–63. [PubMed: 11451816]
6. Schwartz AG, Swanson GM. Lung carcinoma in African Americans and whites. A population-based study in metropolitan Detroit, Michigan. *Cancer*. 1997; 79:45–52. [PubMed: 8988725]
7. DeSantis C, Naishadham D, Jemal A. Cancer statistics for African Americans, 2013. *CA Cancer J Clin*. 2013; 63:151–66. [PubMed: 23386565]
8. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature*. 2013; 500:415–21. [PubMed: 23945592]
9. Leidner RS, Fu P, Clifford B, Hamdan A, Jin C, Eisenberg R, et al. Genetic abnormalities of the EGFR pathway in African American Patients with non-small-cell lung cancer. *J Clin Oncol*. 2009; 27:5620–6. [PubMed: 19786660]
10. Cote ML, Haddad R, Edwards DJ, Atikukke G, Gadgeel S, Soubani AO, et al. Frequency and type of epidermal growth factor receptor mutations in African Americans with non-small cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2011; 6:627–30.
11. Reinersman JM, Johnson ML, Riely GJ, Chitale DA, Nicastri AD, Soff GA, et al. Frequency of EGFR and KRAS mutations in lung adenocarcinomas in African Americans. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2011; 6:28–31.
12. El-Telbany A, Ma PC. Cancer genes in lung cancer: racial disparities: are there any? *Genes & cancer*. 2012; 3:467–80. [PubMed: 23264847]
13. Bauml J, Mick R, Zhang Y, Watt CD, Vachani A, Aggarwal C, et al. Frequency of EGFR and KRAS mutations in patients with non small cell lung cancer by racial background: do disparities exist? *Lung Cancer*. 2013; 81:347–53. [PubMed: 23806795]
14. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008; 455:1069–75. [PubMed: 18948947]

15. Hammerman PS, Sos ML, Ramos AH, Xu C, Dutt A, Zhou W, et al. Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer discovery*. 2011; 1:78–89. [PubMed: 22328973]
16. Beadling C, Heinrich MC, Warrick A, Forbes EM, Nelson D, Justusson E, et al. Multiplex mutation screening by mass spectrometry evaluation of 820 cases from a personalized cancer medicine registry. *J Mol Diagn*. 2011; 13:504–13. [PubMed: 21726664]
17. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP). Seattle, WA: Oct 8. 2013 (URL: <http://evs.gs.washington.edu/EVS/>)
18. Tomasetti C, Vogelstein B, Parmigiani G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc Natl Acad Sci U S A*. 2013; 110:1999–2004. [PubMed: 23345422]
19. Krishnaswamy S, Kanteti R, Duke-Cohan JS, Loganathan S, Liu W, Ma PC, et al. Ethnic differences and functional analysis of MET mutations in lung cancer. *Clin Cancer Res*. 2009; 15:5714–23. [PubMed: 19723643]
20. van Noesel J, van der Ven WH, van Os TA, Kunst PW, Weegenaar J, Reinten RJ, et al. Activating germline R776H mutation in the epidermal growth factor receptor associated with lung cancer with squamous differentiation. *J Clin Oncol*. 2013; 31:e161–4. [PubMed: 23358982]
21. Launonen V, Avizienyte E, Loukola A, Laiho P, Salovaara R, Jarvinen H, et al. No evidence of Peutz-Jeghers syndrome gene LKB1 involvement in left-sided colorectal carcinomas. *Cancer Res*. 2000; 60:546–8. [PubMed: 10676634]
22. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*. 2011; 39:D945–50. [PubMed: 20952405]
23. Lee VH, Tin VP, Choy TS, Lam KO, Choi CW, Chung LP, et al. Association of exon 19 and 21 EGFR mutation patterns with treatment outcome after first-line tyrosine kinase inhibitor in metastatic non-small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2013; 8:1148–55.
24. Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology*. 2010; 138:2073–87 e3. [PubMed: 20420947]
25. Gazdar A, Robinson L, Oliver D, Xing C, Travis WD, Soh J, et al. Hereditary Lung Cancer Syndrome Targets Never Smokers with Germline EGFR Gene T790M Mutations. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2014; 9:456–63.
26. Vikis H, Sato M, James M, Wang D, Wang Y, Wang M, et al. EGFR-T790M is a rare lung cancer susceptibility allele with enhanced kinase activity. *Cancer Res*. 2007; 67:4665–70. [PubMed: 17510392]
27. Oxnard G, Heng J, Rainville I, Dahlberg S, Sable-Hunt A, Wiesner G, et al. Investigating hereditary risk from T790M (inherit): a prospective multicentered study of risk associated with germ-line EGFR T790M. *J Clin Oncol*. 2013; 31
28. An SJ, Chen ZH, Su J, Zhang XC, Zhong WZ, Yang JJ, et al. Identification of enriched driver gene alterations in subgroups of non-small cell lung cancer patients based on histology and smoking status. *PLoS One*. 2012; 7:e40109. [PubMed: 22768234]
29. Gao WM, Mady HH, Yu GY, Siegfried JM, Luketich JD, Melhem MF, et al. Comparison of p53 mutations between adenocarcinoma and squamous cell carcinoma of the lung: unique spectra involving G to A transitions and G to T transversions in both histologic types. *Lung Cancer*. 2003; 40:141–50. [PubMed: 12711114]





**Figure 1.** Overall rates of detection & patterns of multi-mutations by race. The percentages of specimens where a single mutation or multiple mutations were discovered are presented in the pie graphs (*left*). The frequencies for the co-occurrence of mutated genes are displayed in the charts (*right*).



**Figure 2.** Overall rates of detection & patterns of multi-mutations by histology. The percentages of specimens where a single mutation or multiple mutations were discovered are presented in the pie graphs (*left*). The frequencies for the co-occurrence of mutated genes are displayed in the charts (*right*).

**Table 1**

## Patient Characteristics

		<b>Race:</b>	<b>Black (n=137)</b>	<b>White (n=335)</b>
Age	Median		60.8	66.1
	Range(min, max)		(29.3,84.7)	(25.8,83.7)
Gender	Female		80 (58%)	116 (35%)
	Male		57 (42%)	219 (65%)
Histology	Adenocarcinoma		92 (67%)	162 (48%)
	Squamous cell		29 (21%)	133 (40%)
	NSCLC*		16 (12%)	40 (12%)
Variables Stage	I		75 (55%)	219 (65%)
	II		17 (12%)	70 (21%)
	III or IV		45 (33%)	46 (14%)
Smoking history	Ever		123 (90%)	273 (81%)
	Never		13 (9%)	22 (7%)
	Unknown		1 (1%)	40 (12%)
Pack years	Median		38	50
	Range(min, max)		(1, 130)	(1,300)

\* NSCLC not sub-categorized in pathology review.

**Table 2**

Frequency and significance\* for specific mutations identified in tumor tissues among 472 patients by race

gene	variant	Black	White	Pvalue
EGFR	delE746-T751>S	2%	0%	0.076
EGFR	G719C	1%	0%	0.084
EGFR	delE746-A750	4%	2%	0.116
TP53	R158L	1%	3%	0.189

\* Top genes shown, ranked according to smallest Pvalue of significance.

**Table 3**

Logistic regression for odds ratio of having *EGFR* mutations delE746-T751>S or delE746-A750 discovered only in female samples.

Variable	Odds ratio	Lower 95% CI	Upper 95% CI	P value
Race Black	3.914	1.014	15.099	0.048
Age	1.056	0.993	1.123	0.081
StageII	1.666	0.367	7.560	0.508
StageIII,IV	0.572	0.125	2.611	0.471
HistologyNSCLC	0.000	0.000	Inf	0.991
HistologySquamous	0.618	0.106	3.605	0.593
Smoke Never	7.070	2.016	24.790	0.002
Smoke Unknown	9.281	1.286	66.973	0.027

The model was built in the 196 female-only group.

Reference groups: stageI, adenocarcinoma or smoker.

**Table 4**

Description of constitutional polymorphisms and frequencies in NSCLC specimens

Gene variant	dbSNP# or Germ line Reference	Cosmic# or Mutation Reference	Frequency*		Pvalue
			Black	White	
<b>DDR2 (G505S)</b>	rs115169993	Hammerman 2011 (activating)	1(1%;CI:0%,4%)	0(0%;CI:0%,1%)	0.2903
<b>DDR2 (L63V)</b>	rs144594252	Hammerman 2011 (activating)	0(0%;CI:0%,3%)	1(0%;CI:0%,2%)	1
<b>EGFR (R776H)</b>	van Noesel 2013	COSM22940	0(0%;CI:0%,3%)	1(0%;CI:0%,2%)	1
<b>MET (N375S)</b>	rs33917957	COSM28925; Krishnaswamy 2009 (inhibitor resistant)	1(1%;CI:0%,4%)	19(6%;CI:4%,9%)	0.0117
<b>NOTCH1 (V1671I)</b>	rs2229968	COSM33750	12(9%;CI:5%,15%)	0(0%; CI:0%,1%)	<0.0001
<b>STK11 (F354L)</b>	Launonen 2000	COSM1360	0(0%;CI:0%,3%)	5(1%;CI:1%,3%)	0.3278
<b>STK11 (Y272Y)</b>	rs9282859	COSM29005	21(15%;CI:10%,22%)	1(0%;CI:0%,2%)	<0.0001

\* among 472 patients (137 African American, 335 European ancestry) with Wilson's 95% CI