Characterization of Somatic Mutations in Air Pollution-Related Lung Cancer

Xian-Jun Yu, Min-Jun Yang, Bo Zhou, Gui-Zhen Wang, Yun-Chao Huang, Li-Chuan Wu, Xin Cheng, Zhe-Sheng Wen, Jin-Yan Huang, Yun-Dong Zhang, Xiao-Hong Gao, Gao-Feng Li, Shui-Wang He, Zhao-Hui Gu, Liang Ma, Chun-Ming Pan, Ping Wang, Hao-Bin Chen, Zhi-Peng Hong, Xiao-Lu Wang, Wen-Jing Mao, Xiao-Long Jin, Hui Kang, Shu-Ting Chen, Yong-Qiang Zhu, Wen-Yi Gu, Zi Liu, Hui Dong, Lin-Wei Tian, Sai-Juan Chen, Yi Cao, Sheng-Yue Wang, Guang-Biao Zhou

Keywords: Air pollution, Lung cancer, Whole genome sequencing, Exome sequencing, Exposure

Abstract

Air pollution has been classified as Group 1 carcinogenic to humans, but the underlying tumorigenesis remains unclear. In Xuanwei City of Yunnan Province, the lung cancer incidence is among the highest in China attributed to severe air pollution generated by combustion of smoky coal, providing a unique opportunity to dissect lung carcinogenesis of air pollution. Here we analyzed the somatic mutations of 164 non-small cell lung cancers (NSCLCs) from Xuanwei and control regions (CR) where smoky coal was not used. Whole genome sequencing showed that Xuanwei and CR NSCLCs had a mean of 68 and 22 mutated genes per tumor, respectively (p < 0.0001). We found 167 genes (including MYH3, GPR144, CACNA1E, RYR2, KRAS, TP53, ATRX, CACNA1E which had significantly higher mutation frequencies in Xuanwei than CR patients, and mutations in MYH3, GPR144, CACNA1E were associated with patients’ lifetime benzo(a)pyrene exposure. This study uncovers the mutation spectrum of air pollution-related lung cancers, and provides evidence for pollution exposure-genomic mutation relationship at a large scale.

1. Introduction

Air pollution is a significant environmental risk factor for lung cancer. For every increase of 1 μg/m³ in diameter (PM10) in the environment, the risk of lung cancer rises by 18%; for every elevation of 10 μg/m³ in PM smaller than 10 μm (PM10), the risk increases by 22% (Raaschou-Nielsen et al., 2013). Anthropogenic PM2.5 is associated with 220,000 lung cancer mortalities annually (Anenberg et al., 2010). Based on sufficient evidence of carcinogenicity, the International Agency for Research on Cancer (IARC) Working Group recently classified outdoor air pollution...
related PM as Group 1 carcinogenic to humans (Loomis et al., 2013). However, the carcinogenic mechanism of air pollution remains to be dissected using systematic approaches.

Xuanwei (XW) City in Yunnan Province of China (Fig. S1), provides an example of the epidemiological association between PM10, PM2.5 and lung cancer (Xiao et al., 2012; Cao and Gao, 2012; Mumford et al., 1987). This city and the neighboring Fuyuan (FY) county have a large deposit of smoky coal (Mumford et al., 1987). Until the 1970s, residents of these regions used smoky coal in unvented indoor fire pits for domestic cooking and heating, all processes that release high concentrations of PM10 and PM2.5. These airborne particles contain high concentrations of polycyclic aromatic hydrocarbons (PAHs) including benzo(a)pyrene (BaP) and polar compounds that are highly mutagenic (Mumford et al., 1987). Lung cancer incidence in XW is among the highest in China (Mumford et al., 1987; Xiao et al., 2012), and a reduction in lung cancer morbidity was noted in the 1990s after stove improvement in central XW, supporting the association between air pollution and lung cancer (Can et al., 2002). The findings in XW had been cited in the IARC monograph classifying indoor emissions from household combustion of coal as “carcinogenic to humans (Group 1)” (World Health Organization International Agency for Research on Cancer, 2010). However, the overall lung cancer incidence in this region has been increasing (Xiao et al., 2012; Chen, 2008), possibly due to pollutants generated by coal-burning industrial plants that moved into the area (Cao and Gao, 2012). In 2011, a survey (Li et al., 2011) of 52,833 residents living in 382 rural villages in XW/FY reported 363 subjects diagnosed with lung cancer, with the world age-standardized rate (ASR) of 426/100,000 in some regions of XW. Population in these highly polluted regions used smoky coal in unvented indoor cooking and heating, all processes that release high concentrations of coal-related PM as Group 1 carcinogenic to humans (Loomis et al., 2013). However, the carcinogenic mechanism of air pollution remains to be dissected using systematic approaches.

2. Materials and Methods

2.1. Study Design

We sequenced the whole genomes of 14 non-small-cell lung cancers (NSCLCs) from HPR, and performed targeted exome sequencing of 2010 genes in additional 150 primary NSCLCs from HPR and control regions (CR) in the rest of Yunnan and Guangdong Province where the level of air pollution and lung cancer incidence was comparable to most parts of China (van et al., 2010; Chen, 2008). The mutation patterns of HPR and CR NSCLCs were compared, and the exposure–response relationship was analyzed (Fig. S1).

2.2. Patients

The study was approved by the Local Research Ethics Committees of all participating sites. Tumor and adjacent normal lung tissues and peripheral blood samples were obtained from 164 patients with previously untreated NSCLCs (Tables 1, S1, Fig. S1 and S2). The diagnosis and TNM stage were established as previously described (Brambilla et al., 2001; Goldstraw et al., 2007). The patients resided in their communities and rarely (or never) stayed in other regions for a long time, and had regular life routines and regular daily time-spent indoors and outdoors. The exposure dosages of the patients to BaP were estimated by historical measurements in various regions of China that used different fuels and their smoking histories (Sinton et al., 1995; Sullivan and Krieger, 2014). The most recent 10 years were excluded to allow for a hypothesized 10-year latency period between exposure and clinical recognition of lung cancer. Whole-genome sequencing (WGS) was performed in 14 HPR patients (Table S1 and Fig. S3), and exome sequencing was conducted in additional 65 HPR and 85 CR patients (Table 1).

| Characteristics | Total (n = 164) | HPR (n = 79) | CR (n = 85) |
|-----------------|----------------|-------------|-------------|
| Gender          |                |             |             |
| Male            | 101            | 42          | 59          |
| Female          | 61             | 37          | 24          |
| n.d.            | 2              | 0           | 2           |
| Age             |                |             |             |
| <65             | 122            | 61          | 61          |
| ≥65             | 40             | 18          | 22          |
| n.d.            | 2              | 0           | 2           |
| Median, range   | 56 [34, 78]    | 57 [36, 76] | 59 [34, 78] |
| Residence       |                |             |             |
| Xuanwei/Fuyuan  | 79             | 79          | 0           |
| Rest of Yunnan  | 24             | 0           | 24          |
| Guangdong       | 61             | 0           | 61          |
| Smoking         |                |             |             |
| Smoker          | 81             | 38          | 43          |
| Non-smoker      | 81             | 41          | 40          |
| n.d.            | 2              | 0           | 2           |
| Histology       |                |             |             |
| Adenocarcinoma  | 112            | 64          | 48          |
| Squamous-cell carcinoma | 46 | 14 | 32 |
| Adenosquamous carcinoma | 0 | 0 | 0 |
| Large-cell carcinoma | 6 | 0 | 0 |
| TNM stage       |                |             |             |
| IA              | 18             | 12          | 6           |
| IB              | 47             | 25          | 22          |
| IIA             | 13             | 2           | 11          |
| IIIB            | 22             | 9           | 13          |
| IIIB            | 28             | 13          | 15          |
| IV              | 15             | 9           | 6           |
| n.d.            | 5              | 0           | 5           |

n.d.: not determined.

2.3. Analytic Platforms

DNAs and RNAs were isolated from cancer or counterpart normal tissues, sequencing libraries were constructed, and sequenced using the Illumina Hiseq2000 platform. A SNP array using Illumina High Density Genome Wide Human 660WQuad_v1 was performed to detect somatic copy number alterations (SCNAs) throughout the genomes. Mutations of 2010 genes identified by WGS were screened in 150 additional NSCLCs by exome captured and sequencing. The protocols are detailed in eMethods.

3. Results

3.1. Somatic Mutation Profile in HPR NSCLCs

For WGS, cancer DNAs were sequenced to an average of 65.74× (range, 61.02×–74.64×) and normal controls of 43.16× (30.07×–78.70×) coverage. Point mutations, indels, somatic structural variations, and somatic copy number alterations (SCNAs) were found throughout the 14 NSCLC genomes (Fig. S3, Fig. 1 and Tables S1, S2, S3). We reported a mean of 12.75 somatic genomic mutations/Mb, 8.16 exonic mutations/Mb, and 289 exonic mutations/tumor (Table S1). Only an average of 0.74% mutations was found in coding sequences (CDS) (Fig. S4A, B). Using capillary sequencing, 331/361 (91.7%) mutations in CDS were validated. Smoker, non-smokers, squamous cell carcinoma (SCC) and adenocarcinoma (AD) patients had approximately equal mutations in their genomes and CDS (Fig. S4C). The numbers of non-synonymous mutations, synonymous mutations (Fig. 1A), chromosomal rearrangements (Table S2) and somatic gene rearrangements (Fig. 1B) were not associated with smoking status or age. Within predicted promoter regions,
there is a positive correlation between GC content and somatic mutation rate (Fig. S4D).

We analyzed the CDS mutations and reported a mean of 158 non-recurring and 31 recurring (defined as mutated in at least two samples) mutated genes per tumor (Fig. 1C and Table S4). Three genes, MUC16, RYR2 and TP53, were mutated in 7 (50%) of the 14 patients. CSMD3, RYR1, TTN and ZNF831 were mutated in 5 (35.7%) of the patients. There were 18 (including XIRP2 and ANK2), 65 (including EGFR and KRAS), 338 (including CACNA1E, CACNA1S, CACNA2D1 and RYR3) and 2209 genes mutated in 4 (28.57%), 3 (21.43%), 2 (14.29%), and 1 (7.14%) of the 14 patients, respectively (Table S4). Among the 428 recurrently mutated genes, 381 ones had a ratio of non-synonymous to synonymous mutations (dN/dS) > 2. These genes fall into 18 categories (Table S4 and Fig. 1D), with calcium signaling and ion channel-related genes as the three most frequently mutated gene categories.

### 3.2. Somatic Copy Number Alterations (SCNAs)

A total of 479 SCNA segments and a mean of 34 copy number variations (CNVs) per tumor were detected (Table S3 and Fig. S5A). We identified 5 regions of significant CNVs: copy loss in 13q12.3-q34 (containing 13 genes including BRCA2, ERCC5, and RB1), 4p16.1-p13 (containing 6 genes), 4q22.1-q35.2 (containing 18 genes including CASP3, EGF, and FGF2), and 3p24.3-12.2 (containing 17 genes including TGFBR2 and SETD2); copy gain in 1q21.1-q44 (containing 35 genes including ABL2, IL6R, and MCL1) (Table S3).

Some genes including CYP1A1, CYP1B1, CAT, and ERCCI, affect PAH metabolism, detoxification, PAH-DNA adduct formation and repair (Irigaray and Belpomme, 2010). We assessed 23 genes involved in these processes, and reported that 20 ones had copy loss in 12/14 patients, and 1 gene (EPHX1) had copy gain in 4/14 patients at the DNA level (Fig. S5B). By quantitative real-time RT-PCR analysis of 17 of these genes, we found that 13 genes were down-regulated in at
least 7 patients, with down-regulated CYP3A4 seen in 14 patients, decreased CAT, CYP1A1, and NAT in 13 patients, and down-regulated GSTM1 in 12 cases (Fig. S5C). CNV was also frequently seen in DNA repair genes in the patients, with a mean of 114 copy loss and 17 copy gain genes per tumor (Table S3).

3.3. Genomic Rearrangements

We used the CREST method (Wang et al., 2011) to detect and map the breakpoints of the somatic rearrangements among the 14 HPR patients. Totally, 992 chromosomal rearrangements including 573 (57.8%) gene rearrangements and 419 (42.2%) purely intergenic events were identified, with a mean of 71 genomic rearrangements per tumor (Table S2 and Fig. 1B). We identified six previously unreported intrachromosomal in-frame fusion transcripts: ARHGEF10-IMMP2L, COL13A1-DLD, PCDH15-SOX5, CACNA1B-FAF1, MCF2L2-PHF3, and ABCG8-Corf55 in 5 patients (Fig. S6, A-F), and five previously unreported intrachromosomal in-frame fusion transcripts (PLCB1-CRL5, DOCK2-TEMN2, SOX5-ST8S1A1, TADA2B-TBR1D19 and NIN2-NTF3) in four patients (Fig. S7, A-E). Sanger sequencing of PCR products using genomic DNA of the samples was conducted, and 99/108 (91.7%) tested structural variations were validated.

3.4. Genomic Signatures

In 12/14 (85.7%) HPR patients, the C:G → A:T transversions were the most frequent nucleotide substitutions (Fig. S8A). The percentage of transversions ranged from 17.7% to 58.8% (Table S5; p = 0.008). In males, the percentage of A:T → C:G mutations were much higher than in females (Table S5; p = 0.008); in 2 of the 6 males, the most prevalent mutation was A:T → T:A (Fig. S8B). The most frequently observed dinucleotide change was GG → TT/CC → AA (p < 0.0001) (Fig. S8B). Among the trinucleotide alterations, XpCpG → XpApG was detected in 13/14 patients. CpCpX → CpApX and XpCpG → XpTpG were also detected (Fig. S8C).

3.5. Mutations in Calcium Signaling Genes

Gene Ontology Analysis showed that motor activity, calcium ion binding, extracellular matrix structural constituent, ion binding, cation binding, and metal ion binding activity were affected (Table S6). Pathway analysis using single-nucleotide variation and indel data revealed that the differentially altered genes were significantly enriched in 46 KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways, including pathways involved in focal adhesion, calcium signaling and pathogenesis of NSCLC (p < 0.0001) (Table S6). Indeed, some of the most frequently mutated genes (Table S4) were related to calcium signaling, e.g., RYR1-3, which encode RYRs calcium release channels located on the membrane of the endoplasmic reticulum (Lanner et al., 2010); ANK2, which is required for cardiac sinoatrial node Ca2+ homeostasis; XIRP2, which is a target of the calcium-dependent transcription factor MEF2A; CACNAIE (or Cav2.3) (Soong et al., 1993), CACNA1S and CACNA2D1, which encode plasma membrane calcium channel subunits (Pallone et al., 2012).

3.6. HPR NSCLCs Have Much More Mutations Than CR Patients

To unveil the difference in somatic mutations between HPR and CR NSCLCs (Table 1), exons of 2010 of the 2637 (76.2%) mutated genes found in WGS were captured and sequenced (Table S7), and the results were confirmed by sequencing of PCR products using according primers (Table S8). Cancer DNAs were sequenced to an average of 140 × (54.8 ± 261 ×) and normal controls of 145.2 × (46.6 ± 218.5 ×) coverage. We found that the 79 HPR and 85 CR NSCLCs had a mean of 68 and 22 mutated genes per tumor (Table 2), respectively (p < 0.0001). The HPR NSCLCs had much more mutations, e.g., the total number of mutated genes (Fig. 2A), recurrent mutated genes (Fig. 2B), and genes mutated in more than 3%–10% tumor samples (Fig. 2C, Table 2), than the CR patients. In HPR, the number of genes mutated in > 3% samples was 785, more than 3 times higher than that in CR (213 genes; Table 2). There were 59 genes which were mutated in > 10% tumor samples in HPR NSCLCs, but only 6 genes (TP53, EGFR, KRAS, COL22A1, PAPPA2, TMEM132C) were mutated in > 10% tumor samples in CR (Fig. 2D). Among the 1529 recurrent mutated genes (Table S7), 167 genes (including TP53, RYR2, KRAS, CACNAIE, XIRP2) had statistically significantly higher mutation rates in HPR than CR patients, but no gene had significantly higher mutation rate in CR than HPR patients. One gene, TMEM132C, was mutated in 7.59% HPR and 14.12% CR samples, respectively (p = 0.182; Fig. 3C).

3.7. Comparison of Somatic Mutations in HPR and CR NSCLCs

We compared the mutation patterns in HPR and CR NSCLCs. Among the total 80 mutations (46 mutations in HPR and 34 ones in CR NSCLCs) of TP53, 9 common mutations (G154V, R158L, A159P, E171*, Y220C, G244C, G245C, G245V, and E285K) were seen in both regions (Fig. 3A). However, most of the TP53 mutations in HPR were different from those in CR cases. HPR NSCLCs had equal EGFR mutation rate to CR cases (p = 0.135). However, CR non-smokers had higher EGFR mutation rate (40%) than smokers (16.3%; p = 0.0158), while in HPR non-smokers (43.9%) had equal mutation rate to smokers (31.6%; p = 0.26). In HPR, 14/43 (32.6%) mutations occurred in G719, while in CR only 1/25 (4%) alterations was observed in this amino acid: 9 (20.9%) mutations occurred in S768 in HPR cases, but no mutation was found in CR patients (Fig. 3A). Mutations in RYR2, COL22A1, PAPPA2 and TMEM132C in HPR NSCLCs were distinct to those found in CR patients (Fig. 3). Interestingly, 16/19 (84.2%) KRAS mutations in HPR NSCLCs and 10/11 (90.9%) KRAS mutations in CR patients were found in G12 (Fig. 3A), suggesting that NSCLCs of the two regions also had a few similar points in mutation patterns.

3.8. Association Between BaP Exposure and Gene Mutation

Lifetime exposure to BaP was calculated by applying air concentration (ng/m³) reported for various regions of China that used different fuels for cooking and heating (Sinton et al., 1995), an average inhalation rate of 20 m³/day, and the duration of exposure. Smoking a pack of filtered cigarettes per day was assigned a BaP exposure of 0.4 µg/day (Sullvivan and Krieger, 2014). The median value of BaP exposure in

Table 2

|                      | HPR (n = 79) | CR (n = 85)* | p (HPR vs CR) |
|----------------------|-------------|-------------|---------------|
|                      | Total       | S (n = 38)  | NS (n = 41)   |
|                      |             |             | p             |
| Mutation/Mb (mean)   | 15.90       | 18.31       | 13.66         | 0.818       |
|                      |             |             |               |             |
| G:C → T:A substitution| 54.79%     | 54.03%      | 55.73%        | 0.114       |
| Mutated genes/tumor  | 67.99       | 74.87       | 61.61         | 0.930       |
| Non-silent mutations/tumor | 72.66 | 81.24       | 64.71         | 0.895       |
| Genes mutated in >3% samples | 785      | 703         | 617           | 0.062       |

S, smoker; NS, non-smoker.

* The smoking status of two CR patients was unknown.
the HPR was 151.0 mg, five times as high as in the CR (30.1 mg) (Fig. 4A). By logistic regression, we found that the mutation frequencies of 70 genes (including RYR2, MYH3, GPR144, RBCC1, NRXN2, KLHL1, TCHH, ARAP1, COL13A1, CUX2, ZNF800, KCNT1, XIRP2, CACNA1E and TP53) were associated with BaP exposure (p < 0.05; Table S9, Fig. 4B and Fig. S9).

4. Discussion

In this study, we examined the cancer genomes of HPR NSCLCs to identify the genomic mutation profile associated with prolonged exposure to smoky coal pollutants. Of the 14 WGS patients, there was a mean of 12.75 somatic genomic mutations/Mb, 8.16 exonic mutations/Mb, and 289 exonic mutations/tumor. Among the 2010 genes sequenced by targeted exome sequencing, the HPR patients had 68 mutated genes/tumor, 3 times higher than that in CR cases. Previous studies demonstrated that smoker NSCLCs bear more somatic mutations than never-smokers (Imielinski et al., 2012; Govindan et al., 2012). However, in HPR, the smokers and non-smokers harbored equal numbers of mutations and gene rearrangements in their genome. In CR, stages III–IV cancers had more mutations in 6 genes (KRAS, MYH13, TNR, ADAMS520, PXDL1 and SEZ6L) than stages I–II tumors, while patients ≥65 years harbored more mutations in 4 genes (RYR2, COL22A1, ADAMS12 and ZFPM2) than patients <65 years; in HPR, only ACVR2A had more mutations in stages III–IV cancers than stages I–II tumors, and ADCY7 had a higher mutation rate in patients ≥65 years than those <65 years (Table S7 sheet 4). These results demonstrate the genotoxic effect of air pollution and the urgent need to attenuate pollution.

PAHs are important carcinogens in PM2.5 and PM10 (Zielinska et al., 2010; Mumford et al., 1987). A variety of enzymes metabolize PAHs to more polar and water-soluble metabolites to be excreted from the body. However, during the course of metabolism, some unstable and reactive intermediates are formed, which can bind to DNA to form bulky DNA adducts (Hecht, 2012; DeMarini et al., 2001). At the same time, the cells constantly deal with the formation of DNA adducts by DNA repair processes to eliminate these alterations so that mutation does not occur (Irigaray and Belpomme, 2010). We showed that in the WGS NSCLCs, genes responsible for PAH detoxification (GSTM1, GSTP1, GSTT1) were mainly copy loss (Fig. S5B) or down-regulated (Fig. S5C), while genes involved in PAH activation (CYP1B1 in particular; Fig. S5C) were mainly up-regulated. DNA repair genes were mainly copy loss or mutated (Table S3). Mutations in DNA repair pathways have also been implicated in the production of chromosomal translocations (Aplan, 2006). Therefore, the events in PAH metabolism and DNA repair genes may pave the way to genomic mutations and chromosomal translocations, and may represent an essential step to allow accumulation of significant mutations to initiate malignant transformation.

Fig. 2. Comparison of mutations in HPR lung cancer with those in CR NSCLCs. (A): HPR NSCLCs bore more mutated genes than CR patients. (B): Comparison of recurrent mutated genes in HPR with those in CR NSCLCs. (C): Genes mutated in >5% tumor samples from HPR and CR. (D): Mutations in 59 genes whose mutation frequencies are >10% of HPR NSCLCs. Tumors are arranged from left to right in the top track.
PAHs are associated with the C:G → A:T transversions in nucleotides (Ruggeri et al., 1993; Eisenstadt et al., 1982), and recent studies in cell lines showed that BaP can induce this type of nucleotide substitutions (Olivier et al., 2014). We found that the C:G → A:T substitutions were the most frequent nucleotide substitutions in 12/14 patients (Fig. S8), and exome sequencing confirmed the prevalent of C:G → A:T transversions in HPR NSCLCs (Table 2), indicating that PAHs were the main carcinogens for these patients. However, in 2/14 cases the most frequent nucleotide changes were A:T → T:A transitions (Fig. S8), suggesting that there might be other pollutants that caused this signature in the genomes.

Some genes, e.g., TP53, EGFR, and Kras, have high frequency of mutations in lung cancer (The Cancer Genome Atlas Research Network, 2012; Imleriński et al., 2012), and ethnic and sex-related differences in mutation spectrum are noted (Dearden et al., 2013; Kosaka et al., 2004). We showed that the mutation pattern of TP53, EGFR, and Kras in CR NSCLCs (Figs. 2D and 3) was in consistence with previous report in Asian patients (Dearden et al., 2013), and HPR patients also had high mutation rates in these genes (Fig. 2D). Some genes, e.g., COL22A1, PAPPA2, TNR, TMEM132C, ADAMTS20, BAI3, CPS1, and OTOA, had high mutation frequencies in both regions (Table S7). Mutations in most genes, e.g., TP53, COL22A1, PAPPA2, CA1NA1E,
MYH3, NRXN2, RYR2, XIRP2, and TEMEM132C, distributed throughout the entire genome were either missense or nonsense in nature; on the contrary, some genes, e.g., EGF and KRAS, had mutation hot spots (Fig. 3). The results indicated that although NSCLCs from HPR and CR had distinct mutation patterns in many genes, they did show some similar points in some somatic mutations.

Alterations in Ca²⁺ TRP, ORA1 and RYR channels have been identified in cancer (Monteith et al., 2012; Ho et al., 2013; Love et al., 2012), and overexpression of CACNA1E are correlated with relapse in Wilms’ tumors (Natrajan et al., 2006), while CACNA2D1 plays a role in maintaining the properties of tumor-initiating cells in hepatocellular carcinoma (Zhang et al., 2013). Interestingly, Olivier et al. (2014) found that treatment of cells with BaP for 6 days leads to mutations in CACNA1C and CACNA1G. We showed that in the 79 HPR NSCLCs, calcium signaling-related genes RYR2, RYR1, XIRP2, CACNA1E and ANK2 had high frequency mutations (29.1%–17.7%), compared to 1.2%–8.2% mutation rates in CR patients. In HPR NSCLCs, RYR1 and RYR2 had mutations of loss of function patterns, because the mutations were distributed throughout the entire genome and were either missense or nonsense in nature. CACNA1E–FAF1 fusion (Fig. 5D) could also damage CACNA1B’s Ca²⁺ channel function, because its C-terminal ion transmission and calcium channel domains were deleted. The 23 CACNA1E mutations in 15/79 (19%) HPR lung cancers were missense mutations and distributed throughout the entire genome. Among them, 10/23 (43.5%) mutations were found in the amino acid 119–546 region, and 10/15 (67%) patients had one mutation in ion transmission, PKD channel or calcium channel domains (Fig. 3B). These mutations may interfere with the function of the calcium channel and the intracellular Ca²⁺ concentration, the essential second messenger that can regulate nearly every aspect of cellular functions. Further investigation should be conducted to characterize the “driver mutation” aspects of these genes.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2015.04.003.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgment

This project has been funded by the National Natural Science Funds for Distinguished Young Scholar to G.B.Z. (81425025), National Key Program for Basic Research (2012CB910800), National Natural Science Foundation of China (81171925, 81201537), and grant from the Shanghai Municipal Science and Technology Commission (13431500000). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Anenberg, S.C., Horowitz, L.W., Tong, D.Q., West, J.J., 2010. An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling. Environ. Health Perspect. 118, 1189–1195.
Aplan, P.D., 2006. Causes of oncogenic chromosomal translocation. Trends Genet. 22, 46–55.
Brambilla, E., Travis, W.D., Colby, T.V., Corrin, B., Shimosato, Y., 2001. The new World Health Organization classification of lung tumours. Eur. Respir. J. 18, 1059–1068.
Cao, Y., Gao, H., 2012. Prevalence and causes of air pollution and lung cancer in Xuanwei City and Fuyuan County, Yunnan Province, China. Front. Med. 6, 217–220.
Chen, Z., 2008. The Third Chinese National Retrospective Surveys for the Causes of Death. Peking Union Medical College Press, Beijing.
Dearden, S., Stevens, J., Wu, Y.L., Blowes, D., 2013. Mutation incidence and coincidence in non-small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). Ann. Oncol. 24, 2371–2376.
DeMarini, D.M., Landi, S., Tian, D., Hanley, N.M., Li, X., Hu, F., Roop, B.C., Mass, M.J., Keohavong, P., Gao, W., Olivier, M., Hainaut, P., Mumford, J.L., 2001. Lung tumor KRAS and TP53 mutations in nonsmokers reflect exposure to PAH-rich coal combustion emissions. Cancer Res. 61, 6679–6681.
Eisenstadt, E., Warren, AJ, Porter, J., Atkins, D., Miller, J.H., 1982. Carcinogenic epitopes of benzo[a]pyrene and cyclopenta[d]pyrene induce base substitutions via specific transversions. Proc. Natl. Acad. Sci. 79, 1945–1949.
Goldstraw, P., Crowley, J., Chansky, K., Giroux, D.J., Groome, P.A., Rami-Porta, R., Postmus, P.E., Rusch, V., Sobin, L., 2007. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J. Thorac. Oncol. 2, 706–714.
Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N.D., Karchi, K.L., Maher, C.A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Rose, R., Omirz, D., Xiong, D., You, M., Dooling, D.J., Watson, M., Mardis, E.R., Wilson, R.K., 2012. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 150, 1121–1134.
Hecht, S.S., 2012. Lung carcinogenesis by tobacco smoke. Int. J. Cancer 131, 2724–2732.
Ho, A.S., Kannan, K., Roy, D.M., Morris, L.G.T., Ganly, I., Katabi, N., Ramaswami, D., Walsh, L.A., Eng, S., Huse, J.T., Zhang, J., Dolgalev, I., Huberman, K., Heguy, A., Viale, A., Drobnjak, M., Leversha, M.A., Rice, C.E., Singh, B., Iyer, N.G., Leemans, C.R.,
Bloomena, E., Ferris, R.L., Seethala, R.R., Gross, B.E., Liang, Y., Sinha, R., Peng, L., Raphael, B.J., Turcan, S., Gong, Y., Schultz, N., Kim, S., Choi, A., Shah, J.P., Sander, C., Lee, W., Chan, T.A., 2013. The mutational landscape of adenoid cystic carcinoma. Nat. Genet. 45, 791–798.

Imelinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capellotti, M., Sivachenko, A., Sougnez, C., Auclair, D., Lawrence, M.S., Stojanov, P., Cibulskis, K., Choi, K., de la L., Sharifnia, T., Brooks, A., Greulich, H., Banerji, S., Zander, T., Seidel, D., Leenders, F., Ansell, S., Ludwig, C., Engel-Riedel, W., Stoeblenh, E., Wolf, J., Goparju, C., Thompson, K., Winckler, W., Kwiatkowski, D., Johnson, B.E., Jänne, P.A., Miller, V.A., Pao, W., Travis, W.D., Pass, H.J., Gabriel, S.B., Lander, E.S., Thomas, R.K., Garraway, L.A., Getz, G., Meyerson, M., 2012. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150, 1107–1120.

Irigaray, P., Belpomme, D., 2010. Basic properties and molecular mechanisms of exogenous chemical carcinogens. Carcinogenesis 31, 135–148.

Kosaka, T., Yatabe, E., Endoh, K., Takahashi, T., Mitsuomota, T., 2004. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res. 64, 8919–8923.

Lan, Q., Chapman, R.S., Schreinemachers, D.M., Tian, L., He, X., 2002. Household stove improvement and risk of lung cancer in Xuanwei, China. J. Natl. Cancer Inst. 94, 826–835.

Lanner, J.T., Georgiou, D.K., Joshi, A.D., Hamilton, S.L., 2010. Ryandione receptors: structure, expression, molecular details, and function in calcium release. Cold Spring Harb. Perspect. Biol. 2, a003996.

Li, J.H., Zhang, Y.S., Li, Y., Yin, G.Q., Li, Y.B., Ning, B.F., Guo, J.M., 2011. Descriptive study on improvement and risk of lung cancer in Xuanwei, China. J. Natl. Cancer Inst. 118, 847–852.

Loomis, D., Grosse, Y., Lauby-Secretan, B., IARC, 2013. The carcinogenicity of outdoor air pollution. Lancet Oncol. 14, 1262–1263.

Love, C., Sun, Z., Jänne, P.A., Miller, V.A., Pao, W., Travis, W.D., Pass, H.J., Gabriel, S.B., Lander, E.S., Thomas, R.K., Garraway, L.A., Getz, G., Meyerson, M., 2012. Modelling mutational landscapes of human cancers in vitro. Sci. Rep. 4, 4482.

Li, J.H., Zhang, Y.S., Li, Y., Yin, G.Q., Li, Y.B., Ning, B.F., Guo, J.M., 2011. Descriptive study on improvement and risk of lung cancer in Xuanwei, China. J. Natl. Cancer Inst. 94, 8919–8923.

van, D.A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ. Health Per. 813–822.

Soong, T.W., Stea, A., Hudson, C.D., Dubel, S.J., Vincent, S.R., Snutch, T.P., 1992. The genetic landscape of mutations in the p53 gene. Proc. Natl. Acad. Sci. U.S.A., 1013–1017.

Sinton, J.E., Smith, K.R., Hu, H.S., Liu, J.Z., 1995. Indoor Air Pollution Database for China. World Health Organization, Geneva (Switzerland) (WHO/EHG/95.8).

Sullvivan, J.B., Krieger, G.R., 2014. Clinical Environmental Health and Toxic Exposures. Lippincott Williams & Wilkins, Philadelphia, PA.

The Cancer Genome Atlas Research Network, 2012. Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 525–531.

van, D.A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ. Health Per. 813–822.

Wang, J., Müllighan, C.G., Easton, J., Roberts, S., Healey, S.L., Ma, J., Rusch, M.C., Chen, K., Harris, C.C., Ding, L., Holmfeldt, L., Payne-Turner, D., Fan, X., Wei, L., Zhao, D., Obenauer, J.C., Naeve, C., Mardis, E.R., Wilson, R.K., Downing, J.R., Zhang, J., 2011. CREST maps somatic structural variation in cancer genomes with base-pair resolution. Nat. Methods 8, 652–658.

World Health Organization International Agency for Research on Cancer, 2010. Indoor air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). Lancet Oncol. 14, 813–822.

World Health Organization International Agency for Research on Cancer, 2010. Indoor air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). Lancet Oncol. 14, 813–822.

Wang, Y., Xing, B., Zhang, Z., 2013. 1B50-1, a mAb raised against recurrent tumor antigens, targets liver tumor-initiating cells by binding to the calcium channel alpha2delta1 subunit. Cancer Cell 23, 541–550.

Zhao, W., Wang, L., Han, H., Jin, K., Lin, N., Guo, T., Chen, Y., Cheng, H., Lu, F., Fang, W., Wang, Y., Xing, B., Zhang, Z., 2013. 1B50-1, a mAb raised against recurrent tumor cells, targets liver tumor-initiating cells by binding to the calcium channel alpha2delta1 subunit. Cancer Cell 23, 541–550.

Zielinska, B., Samy, S., McDonald, J.D., Seagrave, J., 2010. Atmospheric transformation of diesel emissions. Res. Rep. Health Eff. Inst. 147, 5–60.