RESEARCH   **Open Access**

Identification and validation of *PROM1* and *CRTC2* mutations in lung cancer patients

Yanqi He[1†], Yalun Li[1†], Zhixin Qiu[1], Bin Zhou[2], Shaoqin Shi[3], Kui Zhang[4], Yangkun Luo[5], Qian Huang[6] and Weimin Li[1*]

Abstract

Background: Genetic alterations could be responsible lung cancer, the leading cause of worldwide cancer death.

Methods: This study investigated gene mutations in a Han Chinese family of lung cancer using the whole genome exome sequencing and subsequent Sanger sequencing validation and then confirmed alteration of prominin 1 (*PROM1*) and cyclic AMP-response element binding protein-regulated transcription co-activator2 (*CRTC2*) in blood samples of 343 sporadic lung cancer patients vs. 280 healthy controls as well as in 200 pairs of lung cancer and the corresponding normal tissues using PCR-restriction fragment length polymorphism and directed DNA sequencing of PCR products.

Results: The data showed *PROM1* (p. S281R) and *CRTC2* (p. R379C) mutations, in 5 and 2 cases of these 343 sporadic lung cancer patients, respectively. Notably, these mutations were absent in the healthy controls. Furthermore, in the 200 lung cancer and the matched normal tissues, *PROM1* mutation occurred in 3 patients (i.e., one squamous cell carcinoma and two adenocarcinomas) with a mutation frequency of 1.5%, while *CRTC2* mutation occurred in 5 patients (two squamous cell carcinomas and three adenocarcinomas) with a mutation frequency of 2.5%.

Conclusions: The data from the current study demonstrated novel *PROM1* and *CRTC2* mutations, which could promote lung cancer development.

Keywords: Lung cancer, Gene mutation, *PROM1*, *CRTC2*, Whole genome exome sequencing

Background

Lung cancer is a significant worldwide health problem, accounting for 13% (1.6 million) of the total cancer cases and 18% (1.4 million) cancer deaths [1]. In China, data from the most recently retrospective sampling survey showed that the rate of lung cancer death rate was $30.83/10^5$ ($41.34/10^5$ for men and $19.84/10^5$ for women) [2]. The overall five year survival rate of lung cancer remains approximately 15% worldwide, driven by the majority of cases diagnosed at advanced stages, resulting in non-resectable lesions [1]. Tobacco smoking is the predominant risk factor for both small cell lung cancer and non-small cell lung cancer (NSCLC) [3]. However, to date, the defined pathogenesis of lung cancer remains to be determined because only a small fraction of tobacco

smokers have developed lung cancer, but a certain percentage of lung cancer patients are never smokers, indicating that there is individual variation in cancer susceptibility in the general population and there are other factors, such as genetic factors or host contribution to a predisposition of lung cancer development [4]. Another risk factor in lung cancer development is familial aggregation and history, which associated with 73% increase (95% confidence interval [95%CI]: 50%-100%) in lung cancer risk in women and 71% increase (95%CI: 49-96%) in men [5]. Individuals with the first-degree relative lung cancer had a 1.51-fold increase in lung cancer risk compared to individuals without a family history (95%CI: 1.39-1.63) [6]. Thus, lung cancer risk models using epidemiologic data have been developed and the most parsimonious models for both ever and never smokers include a family history as a lung cancer risk [7].

Indeed, lung cancer can be aggregated in the family and inherited and appears to be the result of an interaction of multiple genes and environmental factors. Thus, search

* Correspondence: weimi003@yahoo.com
†Equal contributors
[1]Department of Respiratory Medicine, West China Hospital of Sichuan University, 37 Guoxue Xiang, Chengdu, Sichuan 610041, China
Full list of author information is available at the end of the article

for a gene or genes associated with susceptibility is urgently needed. Previous studies showed that several tumor suppressor genes were inactivated in lung cancer patients, including *TP53* [8] *RB1* [9,10], and *PTEN* [11], while infrequent activating mutations or amplifications of *PIK3CA*, *EGFR* and *KRAS*, and *MYC* did occur in certain lung cancer patients [12]. Recent genome-wide association studies (GWAS) identified some loci that associated with lung cancer development [13,14]. Chromosomal region 15q24-25.1, containing nicotinic acetylcholine receptor sub-unit genes, has been associated with increased risk of lung cancer in ever smokers [13,14]. Linkage analysis in families with aggregation of lung cancer showed a region on chromosome 6q23-25 associated with risk of lung cancer [15]. These data plus data among individuals with a family history of lung cancer indicate that the relative risk of lung cancer associated with markers in this region is much higher in familial cases compared to the relative risk observed among sporadic cases [16,17].

Thus, in this study, we detected gene mutations in a Han Chinese family of lung cancer using the whole genome exome sequencing [18] and subsequent Sanger sequencing validation and then confirmed these gene alterations in blood samples of 343 sporadic lung cancer patients vs. 280 healthy controls as well as in 200 pairs of lung cancer and the corresponding normal tissues using PCR-restriction fragment length polymorphism and directed DNA sequencing of PCR products.

Results

Detection of PROM1 T/G and CRTC2 G/A mutations in members of lung cancer family using whole genome Exome sequencing

In this study, we performed whole genome Exome sequencing on genomic DNA samples of the four affected and unaffected relatives in this lung cancer family (Table 1). To identify potential genetic variants associated with lung cancer development, we generated an average of 4.9 Gb of DNA sequence with 50X average coverage per individual as single-end, 90-bp reads and the fraction of effective bases on target was about 50% with a minimum 54-fold of average sequencing depth on the target. At this depth of coverage, more than 97% of the targeted bases were sufficiently covered to pass the thresholds of variant calling. We then compared the variants with the Han Chinese Beijing SNPs from dbSNP132 and the 1000 Genome Project. This generated a total of 57 genetic variants (including 55 non-synonymous SNPs and 2 splice) that were shared by the two patients, but absent in the two healthy members, which were predicted to potentially have a functional impact on the gene expression. After that, we directly sequenced PCR-products to validate these 57 variants and obtained accuracy of 67% (38/57) for those variants. We found that mutations of *PROM1* and *CRTC2*

in two lung cancer patients of family cases, but absent in the two healthy members (Table 1).

Validation of genetic variants

To confirm the 38 validated genetic variants, we performed PCR-restriction fragment length polymorphism (PCR-RFLP) analysis of blood samples from 343 unrelated sporadic lung cancer patients and 280 healthy controls (Table 2). This cohort of 343 lung cancer patients had a mean age of 62.04 years old (ranged between 30 and 84 years) with 63.56% of men and 36.44% women, whereas the 280 control individuals had a mean age of 59.98years old (ranged between 35 and 80 years) with 69.64% men and 30.64% women. Age and gender between cases and controls were well balanced. We found mutations in *PROM1* and *CRTC2* in these two family NSCLC cases (Figure 1 and Table 1). *PROM1* mutation was a T to G change, resulting in an S281R amino acid change. The *PROM1* mutation was validated in 5 samples, including 3 adenocarcinomas (ADC) and 2 squamous carcinomas (SCC), with mutation rate of 1.4% (Table 3). Moreover, *CRTC2* mutation was validated with a G to A change, resulting in an R379C amino acid change (Table 3) in two lung cancer samples, including two ADCs, with mutation rate of 0.5%. However, both gene mutations were absent in these 280 controls.

In addition, we further verified *PROM1* and *CRTC2* mutations in an additional 200 pairs of lung cancer and the matched normal tissue specimens (Table 4). We found *PROM1* mutation in 3 pairs of lung cancers and normal tissues, including 1 SCC and 2 ADCs with mutation frequency being 1.5% and *CRTC2* mutation in 5 pairs of lung cancers and normal tissues, including 3 ADCs and 2 SCCs with mutation frequency being 2.5%. All mutations were further confirmed by direct DNA sequencing of PCR products (Figure 1 and Table 3).

Discussion

In the current study, we sequenced and compared the whole genome coding regions of genes in two lung cancer patients and two healthy family members and then filtered the benign changes using public databases, including the 1000 Genome Project and dbSNP132. Use of the second-generation sequencing technique produced a high level of coverage with higher accuracy and allows more regions of a genome to be sequenced in very cost effective manner. We successfully identified *PROM1* and *CRTC2* mutations. Our data are novel and mutation of these two genes had not been previously reported in lung cancer. After that, we confirmed our data in blood samples of 343 lung cancer patients, but not in 280 healthy controls. These two gene mutations were not reported in the 1000 Genomes data and dbSNP132.

PROM1 is localized at chromosome 4p15.32 and contains 27 exons to code CD 133 protein. To date, the

Table 1 Clinicopathological features of the lung cancer family with *PROM1* and *CRTC2* mutations

Patient ID	Gender	Age (years)	Alcohol consumption	Tobacco smoking	Pack years	Medical history	Histological type	TNM (Stage)	Tumor differentiation	*PROM1* mutation		*CRTC2* mutation	
										Nucleotide change	Amino acid change	Nucleotide change	Amino acid change
II4	Male	65	Frequent	Former	1095	COPD, bronchiectasis	Adenocarcinoma	T1aN0M0	Moderate	841 T > G	S281R	1135G > A	R379C
III1	Male	51	Frequent	Former	365	None	Adenocarcinoma	T1bN0M0	Poor-moderate	841 T > G	S281R	1135G > A	R379C
III4	Female	37	Never	Never		None	N/A	N/A	N/A	None	None	None	None
IV1	Male	10	Never	Never		None	N/A	N/A	N/A	None	None	None	None

Table 2 Clinicopathological features of 343 sporadic NSCLC cases and 280 healthy controls

	Lung cancer	Control
Characteristics	N(%)	N(%)
Age in years: median	62.04 ± 10.86	59.98 ± 9.65
Gender		
Female	125(36.44)	85(30.36)
Male	218(63.56)	195(69.64)
Smoking status		
Never	68(19.82)	N/A
Ever	107(31.20)	N/A
N/A	168(48.98)	N/A
Pack years (median, range)	432(20–732)	N/A
Histological types		
Adenocarcinoma	120(34.98)	
Squamous cell carcinoma	204(59.48)	
Adenosquamous carcinoma	8(2.33%)	
Others	11(3.21)[a]	
Clinical stage[b]		
I	76(22.89)	
II	58(17.47)	
III	91(27.41)	
IV	107(32.23)	
Tumor differentiation[b]		
Poor	204(61.44)	
Moderate and poor-moderate	121(36.45)	
Well and moderate-well	7(2.11)	
Lymph node metastasis		
Negative	154(44.90)	
Positive	189(55.10)	

[a] including: 11 small cell lung cancer; [b]for non-small cell lung cancer.

precise functions of CD133 protein are unknown although it is a member of five-transmembrane glycoproteins. The latter proteins specifically localize to plasma membrane protrusions. CD133 protein has two short N (extracellular)- and C (cytoplasmic)-terminal tails, and two large N-glycosylated extracellular loops (between TM2 and –3,and TM4 and –5). *PROM1* can translate into 7 isoforms of CD133 protein through the alternative splicing in human tissues and the alternatively spliced exons in the coding region only affect the short N- and the C-terminal domains [19,20]. CD133 is expressed in various normal and stem cells; thus, it was classified as a marker of primitive hematopoietic and neural stem cells. Recently, rapidly accumulated evidence indicates that CD133 had been described as the most important marker inherent to a number of types of cancer stem cells (CSCs) [21-24]. These CSCs can initiate the

tumors with performing unique functions, such as asymmetric division, self-renewal, drug resistance and quiescence [25]. A previous study reported that a rare population of tumorigenic cells in small cell and non-small cell lung cancer expressed CD133 protein [26]. Lung cancer contained a rare population of CD133[+] CSCs able to self-renew and generated an unlimited progeny of non-tumorigenic cells [27]. Cui *et al.* investigated several human lung cancer cell lines and found that CD133 was a marker for the small cell lung cancer and had stem cell-like features, such as self-renewal, differentiation, proliferation and tumorigenic capacity [25]. *PROM1* mutations, including p.R373C, p.Q576X, p.G614fsX626, and p.Y452fsX12, were associated with a heterogeneous group of inherited retinal disorders with autosomal recessive and dominant inheritance patterns [26,28]. In our current study, we identified somatic *PROM1* mutation (p.S281R) in 8/550 lung cancer patients and this novel *PROM1* mutation was a T to G change, resulting in an S281R amino acid change. All of the 8 patients had non-small cell lung cancer (i.e., 5 adenocarcinomas and 3 SCCs), five of which developed lung cancer at the age younger than 60 years old and two of which were classified as stage IV disease with poor differentiation. The mutation of *PROM1* is novel in lung cancer and further study will investigate the role of CD133 protein in lung cancer development and progression.

Furthermore, *CRTC2* is localized at chromosome 1q21.3 and codes a 693 amino acid protein, namely CRTC2 [29]. CRTC2 protein contains an N-terminal GREB-binding domain (CBD), a central regulatory (REG) domain, a splicing domain (SD), and a C-terminal domain (TAD). CRTCs have shown to induce expression of cyclic AMP-responsive genes [30]. CRTCs are sequestered in the cytoplasm through phosphorylation-dependent interactions with 14-3-3 proteins. In the absence of AMPK activity, CRTC2 is dephosphorylated and translocated into the nucleus where it associated with CREB and induces expression of the target genes [31]. It has reported that CRTC2 was able to regulate gluconeogenesis by integrating calcium and metabolic signaling through calcineurin and AMPK/SIK kinase families [32]. Another study identified LKB1 as an essential activator for the AMPK gene family and a key regulator for CRTC2 transcriptional activity [33], while Shaw *et al.* demonstrated that SIKI may participate in mediating LKB1 tumor suppressor activity [34]. Ji *et al.* found that LKB1 mutations are detected in human tumor samples, including lung cancer [35]. Komiya *et al.* found enhanced activity of CRTC1 in LKB1-null lung cancer [36]. Brown *et al.* reported CRTC2 play an important role in breast cancer of postmenopausal women [37]. In our current study, we found somatic *CRTC2* mutation in 7/550 lung cancer patients. This novel mutation site could change G to A and result
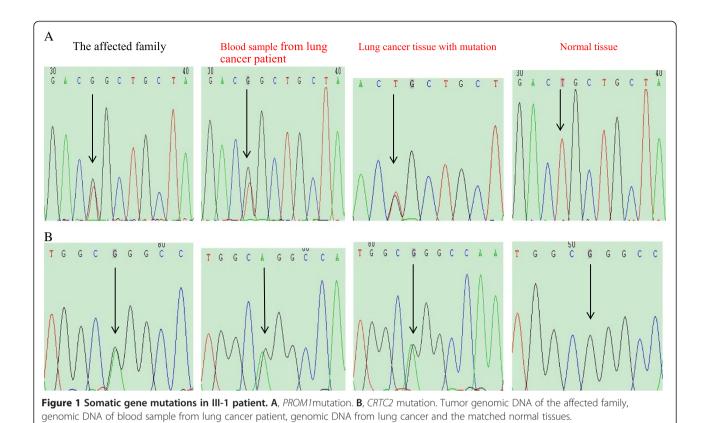
Figure 1 Somatic gene mutations in III-1 patient. A, *PROM1* mutation. **B**, *CRTC2* mutation. Tumor genomic DNA of the affected family, genomic DNA of blood sample from lung cancer patient, genomic DNA from lung cancer and the matched normal tissues.

Table 3 Clinicopathological features of patients with *PROM1* and *CRTC2* mutations

Patient	Gender	Age (yrs.)	Type	T	N	M	Stage	Differentiation	Tumor size (cm)	Family history	Smoking status	Pack years
								PROM1				
LC[a]-15	M	58	ADC	1a	0	0	IA	Poor	1.2 × 1 × 0.8	No	Never	
LC-116	M	70	SCC	4	0	0	IIIA	Poor-moderate	13.5 × 12 × 6	Yes	Current	182
LC-122	F	59	ADC	2a	1	1b	IV	Poor	4.3 × 2.3 × 3	No	Never	
LC-215	M	68	SCC	2b	1	0	IIB	Poor	6.1 × 4.9 × 5	No	Current	546
LC-275	F	54	ADC	4	2	1a	IV	Poor	17 × 11 × 4.1	No	Never	
LCT[b]-17	M	58	ADC	2a	1	0	IIA	Moderate	4 × 3.5 × 3.5	No	Current	182
LCT-51	M	61	SCC	1a	0	0	IA	Poor	2 × 1.5 × 1.4	No	Current	364
LCT-181	F	56	ADC	2a	1	0	IIA	Moderate-well	4 × 3 × 3	No	Never	
								CRTC2				
LC-207	M	72	ADC	2a	0	0	IB	Moderate	2.3 × 3.4 × 2	No	Current	365
LC-250	F	65	ADC	2a	1	0	IIA	Poor-moderate	4 × 4 × 3.5	No	Never	
LCT-81	F	57	ADC	2a	0	0	IB	Poor-moderate	3 × 2.5 × 2.5	No	Never	
LCT-94	M	69	SCC	1b	0	0	IA	Poor	2.5 × 1.5 × 1.0	No	Current	500
LCT-111	M	52	SCC	2a	1	0	IIA	Poor-moderate	4 × 3.5 × 2.5	No	Current	280
LCT-160	M	41	ADC	2a	1	0	IIA	Moderate	4 × 3.5 × 4	Yes	Never	
LCT-171	F	69	ADC	2a	0	0	IB	Poor-moderate	4 × 4 × 3	No	Never	

[a]blood sample of lung cancer; [b]tissue sample of lung cancer and matched normal lung tissue.
ADC, adenocarcinoma; SCC, squamous cell carcinoma.

Table 4 Clinicopathological features of these 200 sporadic lung cancer tissues

	Lung cancer
Characteristics	N(%)
Age in years: median	57.93 ± 10.480
Gender	
Female	56(28.00)
Male	144(72.00)
Smoking status	
Never	92(46.00)
Ever	108(54.00)
Pack years (median, range)	741.94(10–3000)
Histological type	
adenocarcinoma	95(47.50)
squamous cell carcinoma	93(46.50)
adenosquamous carcinoma	10(5.00)
others	2(1.00)[a]
Clinical stage[b]	
I	54(27.27)
II	56(28.28)
III	70(35.35)
IV	18(9.10)
Tumor differentiation[b]	
Poor	92(46.47)
Moderate and poor-moderate	100(50.51)
Well and moderate-well	6(3.02)
Lymph node metastasis	
Negative	98(49.00)
Positive	102(51.00)

[a]including: 2 neuroendocrine carcinoma.
[b]for non-small cell lung cancer.

in an R379C amino acid change. All of these 7 patients had stage I and II non-small cell lung cancer (i.e., 5 adenocarcinomas and 2 SCCs) without metastasis. However, it remains unknown what function of CRTC2 protein plays in lung cancer and further studies are warranted.

Conclusions

In summary, our current study identified and confirmed novel mutations of *PROM1* and *CRTC2* inNSCLC patients. Since the numbers affected patients are too small, we don't know whether these mutations can affect survival of patients. Further studies will investigate the role of these two proteins in lung cancer development and progression.

Methods

Study population

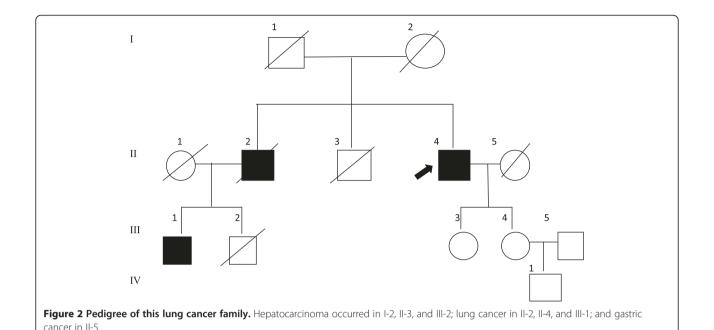This study included a four-generation Chinese Han family with lung cancer cases that had 13(6 living) family members (Figure 2; Table 1). The proband (II-4) was a 65-year-old man who had been diagnosed with adenocarcinoma of the lung in 2012 with a 10-year history of chronic obstructive pulmonary disease (COPD) and bronchiectasis. He had smoked for the past 38 years with approximately 1095 pack per year. Biochemical examination showed an increased CEA (8.97 ng/mL, normal < 5 ng/mL), while CT scanning showed a single nodule with $1.5 \times 1.5 \times 1.0$ cm in the apicoposterior segment of the left upper lobe. This proband's 51-year-old nephew (III-1) presented with cough, hemoptysis and dyspnea. CT scanning confirmed unilateral mass in the superior segment of the right lower lobe. In 2009, he underwent an open lung lobectomy and the result of histopathological examination confirmed as lung adenocarcinoma. In this family, other members also died of lung cancer (II-2), hepatocarcinoma (I-2, II-3, III-2), or gastric cancer (II-5).

Furthermore, to confirm gene mutations in lung cancer of this family members, we recruited 343sporadic lung cancer patients vs. 280 healthy controls as well as 200 pairs of lung cancer and the corresponding normal tissues from The West China Hospital of Sichuan University, The Second People's Hospital of Sichuan, and The Seventh People's Hospital of Chengdu, Chengdu, China between March 2010 and June 2012. All of the sporadic lung cancer patients and the healthy controls were non-related and were of Chinese Han ethnicity (Table 2). Those with secondary lung cancer or other serious disease were intentionally excluded. The diagnosis was confirmed by histopathological examination of the resected or biopsy tissue specimens in all cases. The demographic and clinical information were collected, including sex, age at admission, tobacco smoking, histological diagnosis, tumor-node-metastasis clinical stage according to the American joint committee on cancer 2010 guidelines [38,39]. The study was approved by the Ethics Committee of West China Hospital of Sichuan University and all participants provided informed consent.

DNA samples

Peripheral blood collected from patients and controls into an anticoagulation tube and stored at −80°C until use. Genomic DNA was then extracted using a commercial DNA isolation kit from Bioteke (Beijing, China) according to the manufacturer's instructions. Tumor tissues were fixed in formalin and embedded in paraffin and then sectioned for hematoxylin and eosin staining. Pathologists then identified areas of adequate tumor lesion (>70%) for DNA extraction. The normal lung tissue of the patients obtained during surgical resection represented lung tissue located more than 5 cm from tumor lesions. Genomic DNA was extracted using the Tissue Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol.

Figure 2 Pedigree of this lung cancer family. Hepatocarcinoma occurred in I-2, II-3, and III-2; lung cancer in II-2, II-4, and III-1; and gastric cancer in II-5.

Whole genome exome capture and massively parallel DNA sequencing

We then performed a whole genome exome sequencing to identify the disease-causing genetic variant for this lung cancer family. Briefly, 15 μg of genomic DNA samples from each of these four individuals (II-4, III-1, III-4, and IV-1) were separately sheared into approximately 200 to 300 bp DNA fragments by Covaris, and adapters were then ligated to both ends of the resulting DNA fragments and amplified by ligation-mediated PCR (LM-PCR), purified, and hybridized to the NimleGen 2.1 M human exome array for enrichment, non-hybridized fragments were then washed out. Both non-capture and capture LM-PCR products were subjected to quantitative PCR to estimate the magnitude of enrichment. Each capture library was then loaded on Hiseq2000 platform, and high-throughput sequencing was performed for each captured library independently to ensure that each sample at least 50-fold coverage. Raw image files were processed by Illumina base calling Software 1.7 for base calling with default parameters and the sequences of each individual were generated as 90 bp paired-end reads.

Sequencing reads were aligned to the human genome (hg19) as the reference genome sequence, together with its gene annotation that was downloaded from the UCSC database (http://genome.ucsc.edu/). Single nucleotide polymorphisms (SNPs) were identified by *SOAPsnp* [40] and small insertion/deletions (InDels) were detected by *SAM tools* [41]. SNPs and indels were defined based on satisfaction of the following: i) The consensus quality score was bigger than Q20 (the quality score is a Phred score, generated by the program SOAPsnp12, quality score 20 represents 99% accuracy of a base call); ii) The depth of supported reads at that locus was no less than 4X and no more than 1000X; and iii)The distance between two SNPs no less than 5. The changes that were shared in the two affected individuals, but absent in the two unaffected individuals were obtained by further comparison of the variants of each of the four individuals. All changes were filtered against the Han Chinese Beijing SNPs in the dbSNP132, and the 1000 Genome Project (February 28, 2011 releases for SNPs and February 16, 2011 releases for indel http://www.1000 genome.org). All variants were then confirmed by direct DNA sequencing of polymerase chain reaction (PCR) product (BigDye® Terminator v3.1 Cycle Sequencing Kits; Applied Biosystems, Foster City, CA).

Validation of gene mutations

To further verify the genetic variants identified in these two lung cancer patients sporadic lung cancer patients and the healthy controls, we performed PCR according to the manufacturer's instructions. Primers were synthesized by using the on-line software (http://frodo.wi.mit.edu/primer3/) for PCR amplification of variants identified via exome sequencing. The primers of *PROM1* mutation were 5′-GACCGCAGGCTAGTTTTCAC-3′ and 5′-CTTGCAGTGTGTCCCTCTCA-3′ and the primers of *CRTC2* mutation were 5′-GAGGAGGAAGAGGAG GAGGA-3′ and 5′-CTAAGCAATCCCAACCTCCA-3′. PCR products were then digested overnight with specific restriction enzyme MspI and separated by a 6% polyacrylamide gel electrophoresis and stained with 1.5 g/l of argent nitrate to visualize both *PROM1* T/G and *CRTC2*

G/A mutations. The genotypes and all mutations were confirmed by the DNA sequencing analysis (BigDye® Terminator v3.1 Cycle Sequencing Kits; Applied Biosystems, Foster City, CA). Approximately 10% of the cases were randomly selected for the repeated assays and the results were 100% concordant.

Author details
[1]Department of Respiratory Medicine, West China Hospital of Sichuan University, 37 Guoxue Xiang, Chengdu, Sichuan 610041, China. [2]Laboratory of Molecular Translational Medicine, West China Institute of Women and Children's Health, The Key Laboratory of Obstetric & Gynecologic and Pediatric Diseases and Birth Defects of Ministry of Education, West China Second University Hospital, Sichuan University, Chengdu 610041, China. [3]Department of Immunology, West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, China. [4]Department of Forensic Biology, West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, China. [5]Department of Radiation Oncology, The Second People's Hospital of Sichuan, Chengdu 610041, China. [6]Department of Clinic Laboratory, Chengdu Tumor Hospital, Chengdu 610041, China.

References
1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: **Global cancer statistics.** *CA: Cancer J Clin* 2011, **61**(2):69–90.
2. Chen WQ: **Estimation of cancer incidence and mortality in China in 2004–2005.** *Zhonghua zhong liu za zhi [Chin J Oncol]* 2009, **31**(9):664–668.
3. Yun YH, Lim MK, Jung KW, Bae JM, Park SM, Shin SA, Lee JS, Park JG: **Relative and absolute risks of cigarette smoking on major histologic types of lung cancer in Korean men.** *Cancer Epidemiol Biomark Prevent* 2005, **14**(9):2125–2130.
4. Amos CI, Caporaso NE, Weston A: **Host factors in lung cancer risk: a review of interdisciplinary studies.** *Cancer Epidemiol Biomark Prevent* 1992, **1**(6):505–513.
5. Lissowska J, Foretova L, Dabek J, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Fabianova E, Cassidy A, Mates D, Bencko V, Janout V, Hung RJ, Brennan P, Boffetta P: **Family history and lung cancer risk: international multicentre case–control study in Eastern and Central Europe and meta-analyses.** *Cancer Causes Control* 2010, **21**(7):1091–1104.
6. Cote ML, Liu M, Bonassi S, Neri M, Schwartz AG, Christiani DC, Spitz MR, Muscat JE, Rennert G, Aben KK, Andrew AS, Bencko V, Bickeboller H, Boffetta P, Brennan P, Brenner H, Duell EJ, Fabianova E, Field JK, Foretova L, Friis S, Harris CC, Holcatova I, Hong YC, Isla D, Janout V, Kiemeney LA, Kiyohara C, Lan Q, Lazarus P, *et al*: **Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium.** *Eur J Cancer* 2012, **48**(13):1957–1968.
7. D'Amelio AM Jr, Cassidy A, Asomaning K, Raji OY, Duffy SW, Field JK, Spitz MR, Christiani D, Etzel CJ: **Comparison of discriminatory power and accuracy of three lung cancer risk models.** *British J Cancer* 2010, **103**(3):423–429.
8. Wistuba II, Gazdar AF, Minna JD: **Molecular genetics of small cell lung carcinoma.** *Sem Oncol* 2001, **28**(2 Suppl 4):3–13.
9. Horowitz JM, Park SH, Bogenmann E, Cheng JC, Yandell DW, Kaye FJ, Minna JD, Dryja TP, Weinberg RA: **Frequent inactivation of the retinoblastoma anti-oncogene is restricted to a subset of human tumor cells.** *Proc Natl Acad Sci USA* 1990, **87**(7):2775–2779.
10. Mori N, Yokota J, Akiyama T, Sameshima Y, Okamoto A, Mizoguchi H, Toyoshima K, Sugimura T, Terada M: **Variable mutations of the RB gene in small-cell lung carcinoma.** *Oncogene* 1990, **5**(11):1713–1717.
11. Yokomizo A, Tindall DJ, Drabkin H, Gemmill R, Franklin W, Yang P, Sugio K, Smith DI, Liu W: **PTEN/MMAC1 mutations identified in small cell, but not in non-small cell lung cancers.** *Oncogene* 1998, **17**(4):475–479.
12. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, Lin ML, Beare D, Lau KW, Greenman C, Varela I, Nik-Zainal S, Davies HR, Ordonez GR, Mudie LJ, Latimer C, Edkins S, Stebbings L, Chen L, Jia M, Leroy C, Marshall J, Menzies A, Butler A, Teague JW, Mangion J, Sun YA, McLaughlin SF, Peckham HE, Tsung EF, *et al*: **A small-cell lung cancer genome with complex signatures of tobacco exposure.** *Nature* 2010, **463**(7278):184–190.
13. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X, Vijayakrishnan J, Sullivan K, Matakidou A, Wang Y, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS: **Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1.** *Nat Genet* 2008, **40**(5):616–622.
14. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Chen C, Goodman G, Field JK, Liloglou T, Xinarianos G, Cassidy A, McLaughlin J, Liu G, Narod S, Krokan HE, Skorpen F, Elvestad MB, Hveem K, Vatten L, Linseisen J, *et al*: **A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25.** *Nature* 2008, **452**(7187):633–637.
15. Bailey-Wilson JE, Amos CI, Pinney SM, Petersen GM, de Andrade M, Wiest JS, Fain P, Schwartz AG, You M, Franklin W, Klein C, Gazdar A, Rothschild H, Mandal D, Coons T, Slusser J, Lee J, Gaba C, Kupert E, Perez A, Zhou X, Zeng D, Liu Q, Zhang Q, Seminara D, Minna J, Anderson MW: **A major lung cancer susceptibility locus maps to chromosome 6q23-25.** *Am J Hum Genet* 2004, **75**(3):460–474.
16. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsater A, Flex A, Aben KK, de Vegt F, *et al*: **A variant associated with nicotine dependence, lung cancer and peripheral arterial disease.** *Nature* 2008, **452**(7187):638–642.
17. Liu P, Vikis HG, Wang D, Lu Y, Wang Y, Schwartz AG, Pinney SM, Yang P, de Andrade M, Petersen GM, Wiest JS, Fain PR, Gazdar A, Gaba C, Rothschild H, Mandal D, Coons T, Lee J, Kupert E, Seminara D, Minna J, Bailey-Wilson JE, Wu X, Spitz MR, Eisen T, Houlston RS, Amos CI, Anderson MW, You M: **Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer.** *J Natl Cancer Inst* 2008, **100**(18):1326–1330.
18. Collins FS: **Positional cloning moves from perditional to traditional.** *Nature genetics* 1995, **9**(4):347–350.
19. Permanyer J, Navarro R, Friedman J, Pomares E, Castro-Navarro J, Marfany G, Swaroop A, Gonzalez-Duarte R: **Autosomal recessive retinitis pigmentosa with early macular affectation caused by premature truncation in PROM1.** *Invest Ophthalmol Visual Sci* 2010, **51**(5):2656–2663.
20. Fargeas CA, Huttner WB, Corbeil D: **Nomenclature of prominin-1 (CD133) splice variants - an update.** *Tissue antigens* 2007, **69**(6):602–606.
21. Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, Dirks PB: **Identification of a cancer stem cell in human brain tumors.** *Cancer Res* 2003, **63**(18):5821–5828.
22. Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ: **Prospective identification of tumorigenic prostate cancer stem cells.** *Cancer Res* 2005, **65**(23):10946–10951.

23. O'Brien CA, Pollett A, Gallinger S, Dick JE: **A human colon cancer cell capable of initiating tumour growth in immunodeficient mice.** *Nature* 2007, **445**(7123):106–110.
24. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, De Maria R: **Identification and expansion of human colon-cancer-initiating cells.** *Nature* 2007, **445**(7123):111–115.
25. Cui F, Wang J, Chen D, Chen YJ: **CD133 is a temporary marker of cancer stem cells in small cell lung cancer, but not in non-small cell lung cancer.** *Oncol Reports* 2011, **25**(3):701–708.
26. Arrigoni FI, Matarin M, Thompson PJ, Michaelides M, McClements ME, Redmond E, Clarke L, Ellins E, Mohamed S, Pavord I, Klein N, Hunt DM, Moore AT, Halcox J, Sisodiya SM: **Extended extraocular phenotype of PROM1 mutation in kindreds with known autosomal dominant macular dystrophy.** *Eur J Human Genet* 2011, **19**(2):131–137.
27. Eramo A, Lotti F, Sette G, Pilozzi E, Biffoni M, Di Virgilio A, Conticello C, Ruco L, Peschle C, De Maria R: **Identification and expansion of the tumorigenic lung cancer stem cell population.** *Cell Death Diff* 2008, **15**(3):504–514.
28. Michaelides M, Gaillard MC, Escher P, Tiab L, Bedell M, Borruat FX, Barthelmes D, Carmona R, Zhang K, White E, McClements M, Robson AG, Holder GE, Bradshaw K, Hunt DM, Webster AR, Moore AT, Schorderet DF, Munier FL: **The PROM1 mutation p.R373C causes an autosomal dominant bull's eye maculopathy associated with rod, rod-cone, and macular dystrophy.** *Invest Ophthalmol Vis Sci* 2010, **51**(9):4771–4780.
29. Iourgenko V, Zhang W, Mickanin C, Daly I, Jiang C, Hexham JM, Orth AP, Miraglia L, Meltzer J, Garza D, Chirn GW, McWhinnie E, Cohen D, Skelton J, Terry R, Yu Y, Bodian D, Buxton FP, Zhu J, Song C, Labow MA: **Identification of a family of cAMP response element-binding protein coactivators by genome-scale functional analysis in mammalian cells.** *Proc Natl Acad Sci USA* 2003, **100**(21):12147–12152.
30. Screaton RA, Conkright MD, Katoh Y, Best JL, Canettieri G, Jeffries S, Guzman E, Niessen S, Yates JR, 3rd, Takemori H, Okamoto M, Montminy M: **The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector.** *Cell* 2004, **119**(1):61–74.
31. Koo SH, Flechner L, Qi L, Zhang X, Screaton RA, Jeffries S, Hedrick S, Xu W, Boussouar F, Brindle P, Takemori H, Montminy M: **The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism.** *Nature* 2005, **437**(7062):1109–1111.
32. Altarejos JY, Montminy M: **CREB and the CRTC co-activators: sensors for hormonal and metabolic signals.** *Nat Rev Mol Cell Biol* 2011, **12**(3):141–151.
33. Shaw RJ, Bardeesy N, Manning BD, Lopez L, Kosmatka M, DePinho RA, Cantley LC: **The LKB1 tumor suppressor negatively regulates mTOR signaling.** *Cancer Cell* 2004, **6**(1):91–99.
34. Shaw RJ: **Tumor suppression by LKB1: SIK-ness prevents metastasis.** *Sci Signal* 2009, **2**(86):e55.
35. Ji H, Ramsey MR, Hayes DN, Fan C, McNamara K, Kozlowski P, Torrice C, Wu MC, Shimamura T, Perera SA, Liang MC, Cai D, Naumov GN, Bao L, Contreras CM, Li D, Chen L, Krishnamurthy J, Koivunen J, Chirieac LR, Padera RF, Bronson RT, Lindeman NI, Christiani DC, Lin X, Shapiro GI, Janne PA, Johnson BE, Meyerson M, Kwiatkowski DJ, et al: **LKB1 modulates lung cancer differentiation and metastasis.** *Nature* 2007, **448**(7155):807–810.
36. Komiya T, Coxon A, Park Y, Chen WD, Zajac-Kaye M, Meltzer P, Karpova T, Kaye FJ: **Enhanced activity of the CREB co-activator Crtc1 in LKB1 null lung cancer.** *Oncogene* 2010, **29**(11):1672–1680.
37. Brown KA, McInnes KJ, Hunger NI, Oakhill JS, Steinberg GR, Simpson ER: **Subcellular localization of cyclic AMP-responsive element binding protein-regulated transcription coactivator 2 provides a link between obesity and breast cancer in postmenopausal women.** *Cancer Res* 2009, **69**(13):5392–5399.
38. de Mello RA, Ferreira M, Costa S, Costa BM, Pires FS, Neves I, Almeida MI, Cunha J, Oliveira P, Hespanhol V, Reis RM: **Association between EGF +61 genetic polymorphisms and non-small cell lung cancer increased risk in a Portuguese population: a case–control study.** *Tumour Biol* 2012, **33**(5):1341–1348.
39. Sculier JP, Chansky K, Crowley JJ, Van Meerbeeck J, Goldstraw P: **The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th Edition of the TNM Classification of Malignant Tumors and the proposals for the 7th Edition.** *J Thoracic Oncol: Offic Publ Int Assoc Stud Lung Cancer* 2008, **3**(5):457–466.
40. Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K: **SNP detection for massively parallel whole-genome resequencing.** *Genome Res* 2009, **19**(6):1124–1132.
41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009, **25**(16):2078–2079.