研　究　報　告　書

研究課題：A（一般）

（平成２５年度）

平成 27 年　6 月　1 日

公益財団法人　がん研究振興財団

理事長　　髙　山　昭　三　　殿

研究施設　国立がん研究センター　研究所

住　　所　東京都中央区築地 5-1-1

研究者氏名　白石　航也　　印

（研究課題）

ドライバー変異である RET, ALK と RET 融合遺伝子を伴う肺腺がん発症に関わる遺伝要因の解明

平成２６年　2 月　14 日付助成金交付のあった標記指定課題について研究が終了致しましたのでご報告いたします。

研究題目：ドライバー変異である RET, ALK と RET 融合遺伝子を伴う肺腺がん発症に関わる遺伝要因の解明
研究代表者：白石　航也

目的：　本邦における肺がんはがん死因第一位を占める難治がんであり、罹患数は男女ともに年々増加傾向にある。近年、次世代シークエンサーによるゲノム網羅的な解析が行われ、治療標的となる新規のドライバー変異が同定されてきている。その中で、遺伝子融合を伴うドライバー変異（ALK, RET, ROS1）陽性肺腺がんで認められる遺伝子異常には共通したパターンが認められることが、当研究室で行われた全エキソームシークエンスによる解析から明らかとなってきている（Saito M et al., Cancer Research 2015)。また融合遺伝子を伴う肺腺がん症例の多くが非喫煙者もしくは軽喫煙者であり、タバコによらない発症要因があると考えら、その一つの可能性に遺伝要因、例えば遺伝子多型（SNP）が関わっている可能性がある。

そこで、本研究では既にゲノム解析が終了している遺伝子融合を伴うドライバー変異（ALK, RET, ROS1）陽性肺腺がん症例 50 例（男性 14、女性 30）と健常群 2,823 名(男性 1,512, 女性 1,311)に対して、全ゲノム関連解析を行い、融合遺伝子を伴う肺腺がんの発症に関わる遺伝要因を探索する。本研究を通して、個別化診療のためのバイオマーカー構築に資する基盤情報を提供することを目的とする。

本研究に用いた解析対象：

国立がん研究センター中央病院にて、外科的手術を受けた肺腺がん症例で Omni-1 SNP chip によるゲノム網羅的解析が終了している 1,670 例のうち、国立がん研究センターバイオバンクにて保存されていた凍結がん組織検体より RNA を抽出し、RT-PCR 法にて融合遺伝子の検出を行った。その結果、遺伝子融合を伴うドライバー変異陽性肺腺がん症例 50 例（ALK 融合 27 例, RET 融合 12 例, ROS1 融合 11 例）を同定した。健常群は、日本の製薬企業６社によって構築された日本 PGx データサイエンスコンソーシアム（JPDSC）で収集された集団で、既に Illumina Omni2.5 SNP chip によりゲノム網羅的解析が終了している。

研究方法・結果：

全ゲノム関連解析を行う前にSample及びSNPに対するquality controlを実施した。まず、症例内に同一症例並びに性別の不一致等を調べるため、EIGENSOFTを使用した。関連解析等は、RおよびPLINKといったソフトウェアを使用した。関連解析によって得られたデータに偏りがないかどうかをQ-Q plotで確認したところ、λ_{GC}<1.10であったことから（図1）、このまま解析を続けた。全ゲノムインピュテーションに関しては、ソフトウェアとして，SHAPEIT (pre-phasing 用) およびIMPUTE2(imputation 用)

を使用した。また，reference として1000 Genomes ASN (サンプルサイズはJPT (Japanese in Tokyo, Japan) 89, CHB (Han Chinese in Beijing, China) 97, CHS (Han Chinese South) 100, 計286)のデータを使用した。肺腺がん症例についてはOmni1 SNP chipを用いて解析し、健常群はOmni2.5 SNP chipを用いているため、搭載されたSNPやimputedされたSNP数も異なる（表1）。imputation の対象となるSNP (reference に存在するSNP) の一部はタイピングされているため、タイピングされたSNP のデータと、当該SNP を除外してimpute した当該SNP のデータとを比較することによって、imputation の精度評価をおこなった。その結果精度は、99%以上一致したため、このまま継続して解析を行うこととした。またアレル頻度が1%のSNPについては、肺腺がんの症例数が少ないため、本研究では解析対象外とした。

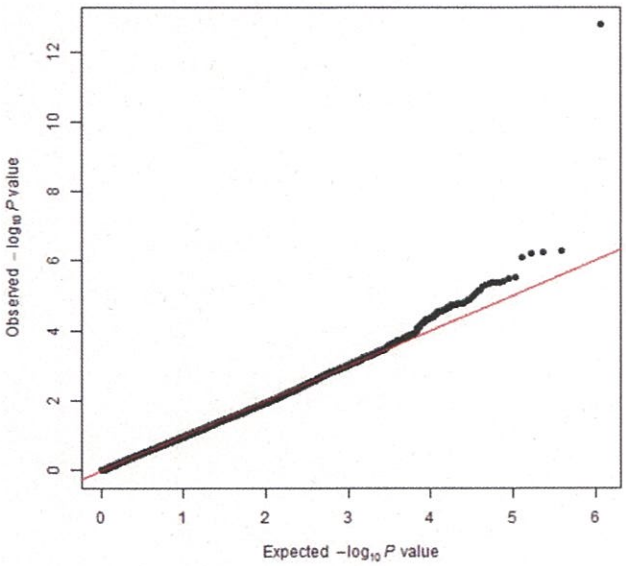

図1.Q-Q plot。横軸は予測値（-log10P)を示し、縦軸は実測値（-log10P)をしめす。
もしこのプトットが直線上に位置しない場合、集団の構造化が疑われる。

表1.肺腺がん症例並びに健常群での関連解析に用いたSNP数の内訳

染色体	肺腺がん症例			健常群		
	Genotyped	Imputed	計	Genotyped	Imputed	計
Chr1	59,346	403,901	463,247	99,066	405,253	504,319
Chr2	56,282	433,967	490,249	103,974	434,072	538,046
Chr3	46,429	383,143	429,572	89,935	385,448	475,383
Chr4	42,450	385,216	427,666	82,716	389,618	472,334
Chr5	41,978	342,500	384,478	78,119	343,723	421,842
Chr6	53,352	362,248	415,600	84,903	362,509	447,412
Chr7	38,363	300,114	338,477	70,334	304,763	375,097
Chr8	37,761	284,777	322,538	68,200	286,104	354,304
Chr9	34,213	225,517	259,730	58,429	224,999	283,428
Chr10	38,415	268,063	306,478	65,738	265,387	331,125
Chr11	36,123	259,769	295,892	62,240	255,710	317,950
Chr12	35,315	252,877	288,192	61,689	252,021	313,710
Chr13	25,607	197,286	222,893	46,001	195,057	241,058
Chr14	22,311	172,259	194,570	42,516	172,037	214,553
Chr15	21,758	144,740	166,498	40,663	146,848	187,511
Chr16	22,194	140,804	162,998	43,058	147,288	190,346
Chr17	19,588	120,479	140,067	36,826	125,361	162,187
Chr18	20,197	145,166	165,363	39,067	147,400	186,467
Chr19	15,553	99,303	114,856	27,071	107,077	134,148
Chr20	19,280	104,390	123,670	31,498	105,199	136,697
Chr21	10,558	69,438	79,996	18,562	70,614	89,176
Chr22	10,350	61,195	71,545	20,116	61,102	81,218
ChrX	15,868	148,624	164,492	28,212	154,070	182,282
計	723,291	5,305,776	6,029,067	1,298,933	5,341,660	6,640,593

肺腺がん症例並びに健常群で共にgenotyped SNPのみを抽出して、Manhattan plotを作成した（図2）。最終的に、imputationされた結果は、$P<10^{-8}$は76 SNP, $P<10^{-7}$は8SNP, $P<10^{-6}$は35 SNP, $P<10^{-5}$は363 SNPとなった。肺腺がん感受性遺伝子座として報告されているSNPは、ORは同じ方向を示していたが、統計学的には強い関連は認められなかった。（rs2736100 in TERT: OR=1.21, P=0.35, rs10937405 in TP63: OR=0.93, P=0.73, rs2395185 in HLA-class II: OR=1.62, P=0.018, rs3817963 in BTNL2: OR=1.66, P=0.014, rs9387478 near ROS1: OR=1.05, P=0.82, rs7086803 in VTI1A: OR=1.08, P=0.74, rs7216064 in BPTF: OR=0.62, P=0.055）。

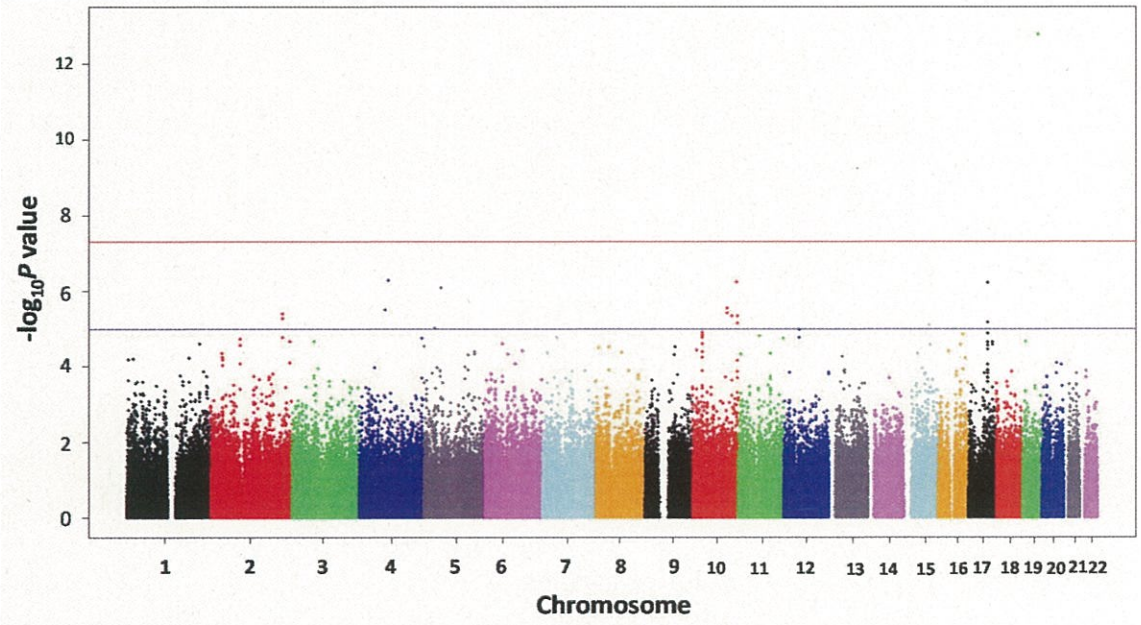

図2.本研究の解析で得られた各SNPのP値(-log10)を縦軸に、染色体上の位置を横軸にとった図（Manhattan plot）。関連が強いほど、大きな数値を示す。

得られた解析結果の内、$P<10^{-5}$でかつSNPinfoもしくはHaploRegといった機能推定ソフトウェアを用いて、遺伝子発現制御やアミノ酸置換を伴うSNPを選択した。その結果、選択的スプライシングに関わると推定されるアミノ酸置換を伴う2つのSNPに注目した（表2）。

表2.本研究によって同定された候補SNP

SNP	Gene	Amino acid change	Chr.	Position	Major allele	Major allele frequency		P value	OR
						Case	Control		
rs1345658	ZNF30	R380K	19	35435006	G	0.49	0.18	1.78E-14	5.3
rs35750729	EPX	K441T	17	56276940	A	0.08	0.01	4.07E-06	6.1

ZNF30は、zinc finger protein 30で転写因子として働くと考えられているが、機能については殆どわかっていない。一方、EPX (Eosinophil Peroxidase)はperoxidase遺伝子ファミリーの一つで、好酸球の中で発現し、アレルギーや寄生虫感染によって誘導されることがわかっている。しかし、肺がんリスクとの関連については不明な点が多い。

結論

今回、2つの候補となるSNPを同定するに至った。しかしながら、症例数が少ないことや検証研究に用いるための症例数が現状では確保できなかったため、未だ遺伝子融合を伴うドライバー変異陽性肺腺がん症例特異的な感受性遺伝子座の同定には至っていない。現在、さらに症例数を確保し、真の感受性遺伝子座の同定や機能的意義についても検討して行く予定である。

本研究成果論文：

Shiraishi K and Kohno T. Genetic susceptibility to lung adenocarcinoma. Genes and Environment. 2014; 36 (3): 160-166.

以上

Genes and Environment, Vol. 36, No. 3 pp. 160–166 (2014)

Review

Genetic Susceptibility to Lung Adenocarcinoma

Kouya Shiraishi[1,2] and Takashi Kohno[1]

[1]Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan

Received June 19, 2014; Revised July 1, 2014; Accepted July 4, 2014
J-STAGE Advance published date: July 10, 2014

Lung adenocarcinoma (LADC) is the most common histological type of lung cancer and its incidence is increasing worldwide. Genetic polymorphisms, including single nucleotide polymorphisms (SNPs), underlie inter-individual differences in cancer susceptibility, and genetic loci for LADC risk have been identified by genome-wide association studies (GWAS) and candidate gene association studies. Recently, three GWAS of LADC and subsequent pooled GWAS analyses identified genetic susceptibility variants on chromosome 15q25 (*CHRNA*), 5p15 (*TERT*), 3q28 (*TP63*), 6p21 (*BAT3-MSH2* and *BTNL2*), and 17q.24 (*BPTF*). SNPs in *TERT* and *TP63* are associated with increased risk for LADC in both never-smokers and smokers, whereas those in *CHRNA* are associated with increased risk of lung cancer irrespective of histological type. However, the risk alleles for *CHRNA* SNPs are rare in Asian populations, including Japanese. The association of 5p15 and 3q28 variants with increased risk of LADC was validated in both European and Asian populations; however, strength of association with LADC risk seems different by ethnicity. The association of SNPs in *BTNL2* and *BPTF* with LADC risk was replicated in one study. On the other hand, significant associations of functional variants in DNA repair and metabolic genes have not been reported in lung cancer GWAS. Here, we review previously reported GWAS and candidate gene analyses and discuss identified genetic factors for LADC risk, which may be useful for early detection or prevention of LADC.

Key words: lung adenocarcinoma (LADC), genome-wide association studies (GWAS), single nucleotide polymorphisms (SNPs)

Introduction

Lung cancer is the leading cause of cancer mortality worldwide (1). According to vital statistics collected in Japan between 1950 and 2010, high rates of lung cancer mortality are seen in males in age groups 60–64 or older, although age-standardized lung cancer mortality rates have increased in women. Lung cancer mortality and occurrence rates remain high. Despite advancements in recent molecular targeted drugs, there are few efficient therapies for advanced stage lung cancer patients. In fact, the 5-year survival rate for stage IV lung cancer is under 20%, in contrast to 71.4% for stage IA (2). These results suggest that earlier diagnosis and treatment of lung cancer would significantly improve clinical outcomes and reduce mortality.

Tobacco smoking and exposure to tobacco smoke are a major cause of lung cancer. Tobacco smoking is associated with a substantially increased risk of mortality, accounting for approximately 2 million deaths in adults aged ≥45 years old throughout Asia (3). Tobacco cessation is a very useful tool for prevention of lung cancer. The different histological types of lung cancer are typically divided into small cell lung cancer (SCC) and non-small cell lung cancer, comprising adenocarcinoma (LADC) and squamous cell carcinoma (SQC) (1). Development of LADC is less associated with smoking than SQC and SCC (4), indicating that the mechanisms of carcinogenesis differ among these types. Although lung cancer is predominantly caused by tobacco smoking, several genome-wide association studies (GWAS) have reported that inherited genetic factors (i.e., genetic polymorphisms) increase the risk of lung cancer (5–13). Risk variants may result in different magnitudes of lung cancer risk depending on the population, smoking behavior, and histological type. In this review, we summarize previously reported GWASs for lung cancer, particularly LADC. Further studies of genetic factors will help clarify disease etiology and identify high risk individuals for targeted screening and/or prevention.

GWASs of LADC Risk

Previously reported GWAS are summarized in Table 1. Several GWASs report that three chromosomal loci, 15q24-25.1, 5p15, and 6p21, are associated with increased risk of lung cancer in European/American populations (5–7), while loci 3q28, 6p21, and 17q24 are associated with increased risk of LADC in Japanese and/or Korean populations (10,11). In addition, loci 13q12 and 22q12 are associated with lung cancer risk in

[2]Correspondense to: Kouya Shiraishi, Division of Genome Biology, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Tel: +81-3-3542-2511, Fax: +81-3-3542-0807, E-mail: kshirais@ncc.go.jp.
doi: org/10.3123/jemsge.2014.021

Table 1. Succeptibility genes for lung cancer risk identified by GWAS

Susceptibility locus	Reported gene(s)	Strongest SNP-Risk allele	Risk allele frequency			Odd ratio	*P*-value	Reference
			CEU	JPT	HCB			
3q28	TP63	rs4488809-C/ rs10937405-C	0.508/ 0.636	0.589/ 0.478	0.478/ 0.722	1.19–1.31	10^{-9}~10^{-26}	10, 11, 12, 14
5p15.33	TERT	rs2736100-G/ rs2853677-C	0.527/ 0.420	0.376/ 0.262	0.407/ 0.221	1.12–1.41	10^{-10}~10^{-40}	9, 10, 11, 12, 14
5p15.33	CLPTM1L	rs31489-C/ rs401681-G	0.607/ 0.566	0.866/ 0.657	0.814/ 0.721	1.12–1.15	10^{-9}~10^{-10}	7, 8, 9
6q22.1	ROS1, DCBLD1	rs9387478-C	0.496	0.576	0.5		10^{-10}	14
	TRNAA-UGC	rs4324798-A	0.088	0	0		10^{-8}	9
6p21.32	HLA class II	rs2395185-T	0.433	0.372	0.407	1.17	10^{-8}	14
6p21.32	BTNL2	rs3817963-G	0.345	0.32	0.244	1.18	10^{-10}	11
6p21.33	BAT3-MSH5-APOM	rs3117582-C	0.08	0	0	1.22–1.24	10^{-10}~10^{-12}	7, 8, 9
10q25.2	VTI1A	rs7086803-A	0.023	0.268	0.279	1.28	10^{-18}	14
13q12.12	MIPEP	rs753955-G	0.659	0.384	0.43	1.18	10^{-12}	12
15q25.1	CHRNA3,CHRNA5, CHRNB4, PSMA4, LOC123688	rs8034191-C/ rs1051730-T	0.419/ 0.385	0.014/ 0.012	0.047/ 0.035	1.29–1.35	10^{-12}~10^{-51}	5, 6, 7, 8, 9
17q24.2	BPTF	rs7216064-A	0.221	0.733	0.64	1.16–1.20	10^{-6}~10^{-11}	11, 14
22q12.2	MTMR3	rs36600-A	0.252	0.076	0.105	1.29	10^{-13}	12

Chinese populations (12), and loci 10q25 and 6p21 are associated with susceptibility to lung cancer in Asian female never-smokers (14).

The chromosomal 15q24–25.1 region contains nicotinic acetylcholine receptor subunit genes, i.e., *CHRNA3* (cholinergic receptor, nicotinic, alpha 3) and *CHRNA5*. Their products are expressed in pulmonary epithelial cells and bind to nicotine, an addictive compound found in tobacco smoke, and nitrosamines, which are potential lung carcinogens found in tobacco smoke and certain foods (15,16). In Asians, SNPs in these genes are associated with lung cancer risk (17,18), although conflicting results have been reported by different investigators (19). Since the frequency of these risk alleles is much lower in Asian populations than in populations of European descent, these results probably reflect low statistical power of the studies. At least, the significance of CHRNA risk alleles on lung cancer risk varies between Asian and European populations. Thus, GWAS conducted with sample sets of Asian populations are necessary to fully identify genetic factors underlying lung adenocarcinoma.

The 5p15.33 region contains *TERT* (telomerase reverse transcriptase) and *CLPTM1L* (cleft lip and palate transmembrane protein 1). TERT functions in telomere replication and maintenance, and promotes epithelial cell proliferation (20). The risk allele frequency of the *TERT* SNP, rs2736100, is similar among various ethnicities, and associations have been detected in European, American, and Asian populations (Fig. 1A) (8,10–12,19,21,22). A recent meta-analysis investigating the association of this SNP with susceptibility to various types of lung cancer, including LADC, SQC, and SCC, reported that the association with LADC is stronger than that of SQC or SCC (Fig. 1A) (23). Genetic modifiers and/or environmental factors in the different histological types might have contributed to the difference. The rs2736100 SNP is also associated with susceptibility to other types of cancer, including cancers of the brain, bladder, prostate, uterine cervix, skin, testes, and chronic lymphocytic leukemia (24,25). The rs2736100 SNP is located in intron 2 of *TERT*, which is a putative region regulating telomere length. It was previously reported that risk alleles in the *TERT* gene may be associated with shorter telomere length in leukocytes (24). However, rs2736100 SNP, which is associated with increased risk of glioma, was strongly correlated with longer telomere length in the same cell type (26). A recent GWAS of breast and ovarian cancers indicated that multiple independent variants at the *TERT* locus were associated with shorter telomere length and cancer risk (26). Therefore, multiple variants at the *TERT* locus may independently affect leukocyte telomere length. To elucidate the functional significance of *TERT* SNPs, further studies are warranted.

Interestingly, the *CLPTM1L* gene located near the *TERT* gene. *CLPTM1L* was identified as a cisplatin (CDDP) resistance genes. A recent meta-analysis suggested that the association of rs31489, a SNP in the promoter region of *CLPTM1L*, with lung cancer risk was stronger in Caucasians than in Asians (25). Michael *et al.* showed that *CLPTM1L* is required for lung tumorigenesis in a conditional K-Ras^{G12D} transgenic mouse model (27). The frequency of *KRAS*-mutated LADC was different among Caucasians and Asians. KRAS is activated by a single amino acid substitution (codons 12 and 13) in 15–25% of LADC cases in Americans and 5–10% of LADC cases in East Asians (28). The *CLPTM1L* SNP may be preferentially associated with increased risk of *KRAS*-mutated LADC.

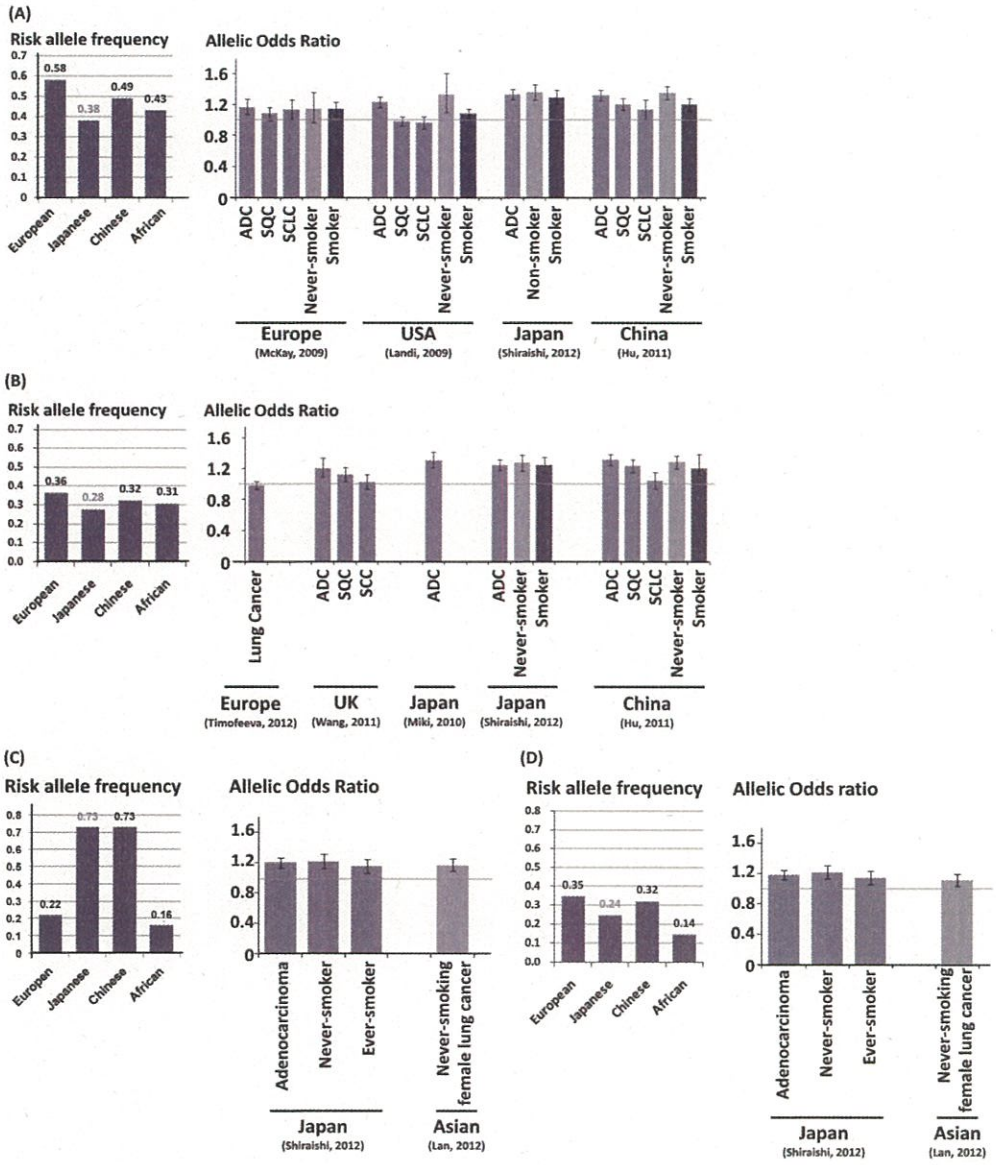

Fig. 1. Increased lung cancer risk with SNPs identified by GWASs according to population, smoking behavior, and histological type. (A) rs2736100 (*TERT*) at 5p15.33. (B) rs10937405 (*TP63*) at 3q28. (C) rs7216064 (*BPTF*) at 17q24.3. (D) rs3817963 (*BTNL2*) at 6p21. Frequencies of risk alleles in each population, as determined by the HapMap project or the International Histocompatibility Working Group, are shown on the left.

The 3q28 region contains the *TP63* gene that encodes a member of the tumor suppressor *TP53* (also known as p53) gene family, which is involved in development, differentiation, and the cellular stress (29). The risk allele frequency of the rs10937405 SNP was similar among ethnic groups (Fig. 1B). A recent meta-analysis showed that the rs10937405 SNP might be a risk conferring factor for development of non-small cell carcinoma, especially for LADC and of East Asian populations (30). SNPs associated with increased lung cancer risk are located in intron 1 of *TP63* and the region may have a functional role in the regulation of *TP63* gene expression (10). TP63 expression is induced by DNA damage. Therefore, inter-individual differences in DNA damage responses to genotoxic stress due to *TP63* SNPs may contribute to lung cancer susceptibility.

The 17q24.3 region contains the *BPTF* (bromodomain PHD finger transcription factor) gene. *BPTF* encodes a chromatin remodeling factor that regulates transcription via specific recognition of methylated histone proteins (31). A GWAS performed by us showed that variants of *BPTF* were associated with

increased risk of LADC, and that the risk allele frequency of rs7216064, a SNP located in intron 4 of *BPTF* associated with LADC risk, was different in various ethnic populations (Fig. 1C) (11). One study validated this association using a population of Asian female never-smokers (Fig. 1C) (14). Recently, chromatin remodeling genes have been implicated as tumor suppressors in LADC (32) and in other cancers (33). Importantly, changes in copy number and somatic mutations in *BPTF* are present in many types of cancer (34) (http://www.cbioportal.org/public-portal/). Further studies are required to elucidate how *BPTF* contributes to LADC susceptibility.

Associations of SNPs in the 6p21 region were identified in GWAS of European and Asian populations (7,11,12,14,22). This region contains *BAT3* (HLA-B associated transcript 3) and *MSH5* (mutS homolog 5). The BAT3 protein complexes with a histone acetyltransferase, p300, which acetylates histone and p53 proteins in response to DNA damage, while MSH5 is involved in DNA mismatch repair. Variants of *BAT3-MSH5* were associated with increased lung cancer risk in European populations, irrespective of histological type; however, these *BAT3-MSH5* SNPs were not polymorphic in Asians. In our GWASs, variants of *BTNL2* (butyrophilin-like 2) and the *HLA-DQA1* locus, located in the *HLA* (human leukocyte antigen)-class II region, were significantly associated with increased LADC risk (11,35). Frequency of the *BTNL2* SNP is similar among ethnic groups (Fig. 1D). In addition, association of the *BTNL2* SNP with lung cancer risk was validated in Asian female never-smokers (14). It is possible that *HLA*-class II gene polymorphisms confer lung cancer susceptibility causing inter-individual differences in the immune response against tumor cells. Notably, the 6p21 region is an extremely highly polymorphic region which contains major histocompatibility complex genes and is in strong linkage disequilibrium; therefore, the observed associations might be affected by population structure (9). Further investigation is warranted to determine whether and how genotypes in this region contribute to risk for lung cancer.

Functional Polymorphisms in DNA Repair and Metabolic Genes

Studies of DNA adduction/damage, including that produced by tobacco carcinogens and associated repair processes, have identified various metabolic and DNA repair genes with functional polymorphisms (26,36–41). A recent meta-analysis showed that representative SNPs in *TP53* (tumor protein p53), *OGG1* (8-oxoguanine DNA glycosylase), *CYP1A1* (cytochrome P450, family 1, subfamily A, polypeptide 1), and *EPHX* (epoxide hydrolase 1, microsomal (xenobiotic)) were associated with increased lung cancer risk (42–46) (Table 2). For example, the risk (72Pro) allele of the *TP53*-Arg72Pro SNP encodes a protein with weaker apoptotic activity than that of the 72Arg allele, which enables increased survival of DNA-damaged cells (47), while the risk (326Cys) allele of the Ser326Cys SNP in *OGG1* encodes a DNA glycosylase with weaker activity in the repair of oxidative promutagenic base damage, 8-hydroxyguanine, produced by tobacco and other carcinogens than that of the 326Ser allele (48,49). The risk (462Val) allele of the Ile462Val SNP of *CYP1A1* encodes a metabolic protein with higher activity in bio-activating the major tobacco procarcinogens, polycyclic aromatic hydrocarbons (PAHs), than the 462Ile (50). EPHX catalyzes the hydrolysis of arene, alkene, and aliphatic epoxides from PAH and aromatic amines. The risk (113His) allele of

Table 2. Association of functional polymorphisms in metabolic and DNA repair genes with lung cancer risk

Gene	Gene product	SNP ID	Polymorphism	Odd ratio	Reference
Metabolic gene					
CYP1A1	Cytochrome P450	rs1048943	Ile462Val	2.36	36
mEH (EPHX1)	Epoxide hydrolase	rs1051740	His113Tyr	0.70	36
MPO	Myeloperoxidase	rs2333227	G-463A	0.71	36
GSTM1	Glutathione-S-transferase	—	Presence/Null	1.18	37
GSTT1	Glutathione-S-transferase	—	Presence/Null	1.28	36
DNA repair gene					37
XPA	Nucleotide excision repair protein	rs1800975	G-23A	0.73	37
XPC	Nucleotide excision repair protein	rs2228001	Lys939Gln	1.30	37
XRCC1	Base excision repair protein	rs25487	Arg399Gln	1.34	37
ERCC1	Nucleotide excision repair protein	rs11615 rs3212948	exon 4 T19007C 3′-UTR C8092A	1.24 0.79	39 39
ERCC2	Nucleotide excision repair protein	rs13181	Lys751Glu	1.30	37
ERCC4	Nucleotide excision repair protein	rs1800067	Arg415Gln	0.82	38
ERCC5	Nucleotide excision repair protein	rs17655	Asp1104His	1.23	38
OGG1	Base excision repair protein	rs1052133	Ser326Cys	1.24	40
TP53	Transcription factor	rs1042522	Arg72Pro	1.20	44
MDM2	Ubiquitine ligase	rs2279744	T309G	1.27	41

the His113Tyr SNP of *EPHX* leads to decreased enzyme activity. However, associations of these functional polymorphisms have never been reported in previous GWASs as a result of the lack of probes targeting these SNPs in commercially available DNA chips. To validate this association of these SNPs in this review, authors performed a case-control study using imputed genotypes consisted of Japanese LADC patients and non-cancerous control subjects. Notably, the rs1051740 SNP of *EPHX* was associated with increased risk of LADC (allelic odds ratio = 0.91, *P* value = 0.036); however, no other SNPs show associations with increased risk for LADC.

Limitations and Applications of GWASs

Genetic factors affecting lung cancer risk have been identified using GWAS and other association studies. The results indicate that risk variants confer different magnitudes of risk in different populations for a variety of reasons, including differences in allele frequency and the genetic and environmental backgrounds that interact with each variant. In addition, it is important to find out rare variants with large effects on the risk of lung cancer. Wang et al. performed imputation using the 1000 Genomes resource and conducted a meta-GWAS that consisted of 21,594 European lung cancer cases and 54,156 controls (51). In this study, variants of *BRCA2* (p.Lys3326X: rs11571833, odds ratio = 2.47) and *CHEK2* (p.Ile157Thr: rs13314271, odds ratio = 0.38) were significantly associated with increased risk of SQC. These results indicate that the imputation strategy can identify rare disease-causing variants with large effects on lung cancer risk from existing GWAS data. Such an analysis should be done also in Asian population. Our studies and other collaborators using larger numbers of subjects in international and/or pan-Japan consortiums, such as ILCCO (International Lung Cancer Consortium) and FLCCA (Female Lung Cancer Consortium in Asia) and JLCS (Japanese Lung Cancer Collaborative Study), are promising.

GWASs focusing on specific types of lung cancer would also be worthwhile, because some genetic factors might be specifically associated with increased risk of a specific type of lung cancer, such as SCC, lung cancers in female never-smokers, or lung cancers with defined gene mutations. LADC develops via several carcinogenic pathways defined by oncogenic driver mutations in *EGFR*, *KRAS*, *HER2*, *ALK*, and *RET*, and etiological factors are thought to be different among these pathways (1,52,53). Understanding the remaining genetic factors will help clarify disease etiology and identify high risk individuals for targeted screening and/or prevention.

Acknowledgments: This work was supported in part by Grants-in-Aid from the Ministry of Health, Labor, and Welfare for Research for Promotion of Cancer Control Programs, from the Ministry of Education, Culture, Sports, Science and Technology of Japan for Scientific Research on Innovative Areas (22131006), Foundation for Promotion of Cancer Research in Japan and the National Cancer Center Research and Development Fund (25-A-1 and 26-A-8).

Conflicts of interest statement: The author declares that there are no conflicts of interest.

References

1 Herbst RS, Heymach JV, Lippman SM. Lung cancer. N Eng J Med. 2008; 359: 1367–80.
2 Matsuda A, Matsuda T, Shibata A, Katanoda K, Sobue T, Nishimoto H, Japan Cancer Surveillance Research G. Cancer incidence and incidence rates in Japan in 2008: a study of 25 population-based cancer registries for the Monitoring of Cancer Incidence in Japan (MCIJ) project. Jpn J Clin Oncol. 2014; 44: 388–96.
3 Zheng W, McLerran DF, Rolland BA, Fu Z, Boffetta P, He J, et al. Burden of total and cause-specific mortality related to tobacco smoking among adults aged >/=45 years in Asia: a pooled analysis of 21 cohorts. PLoS med. 2014; 11: e1001631.
4 Sobue T, Suzuki T, Fujimoto I, Matsuda M, Doi O, Mori T, et al. Case-control study for lung cancer and cigarette smoking in Osaka, Japan: comparison with the results from Western Europe. Japanese Journal of Cancer Research. Gann. 1994; 85: 464–73.
5 Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature. 2008; 452: 633–7.
6 McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet. 2008; 40: 1404–6.
7 Wanfg Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet. 2008; 40: 1407–9.
8 Broderick P, Wang Y, Vijayakrishnan J, Matakidou A, Spitz MR, Eisen T, et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. Cancer Res. 2009; 69: 6633–41.
9 Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85: 679–91.
10 Miki D, Kubo M, Takahashi A, Yoon KA, Kim J, Lee GK, et al. Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. Nat Genet. 2010; 42: 893–6.
11 Shiraishi K, Kunitoh H, Daigo Y, Takahashi A, Goto K, Sakamoto H, et al. A genome-wide association study identifies two new susceptibility loci for lung adenocarci-

noma in the Japanese population. Nat Genet. 2012; 44: 900–3.
12 Hu Z, Wu C, Shi Y, Guo H, Zhao X, Yin Z, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet. 2011; 43: 792–6.
13 Kohno T, Shiraishi K. Genetic polymorphisms underlying lung cancer susceptibility and therapeutic response. Genes Environ. 2012; 34: 94–100.
14 Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nat Genet. 2012; 44: 1330–5.
15 Minna JD. Nicotine exposure and bronchial epithelial cell nicotinic acetylcholine receptor expression in the pathogenesis of lung cancer. J Clin Invest. 2003; 111: 31–3.
16 Schuller HM. Nitrosamines as nicotinic receptor ligands. Life Sci. 2007; 80: 2274–80.
17 Shiraishi K, Kohno T, Kunitoh H, Watanabe S, Goto K, Nishiwaki Y, et al. Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. Carcinogenesis. 2009; 30: 65–70.
18 Wu C, Hu Z, Yu D, Huang L, Jin G, Liang J, et al. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. Cancer Res. 2009; 69: 5065–72.
19 Truong T, Hung RJ, Amos CI, Wu X, Bickeboller H, Rosenberger A, et al. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. J Natl Cancer Inst. 2010; 102: 959–71.
20 Baird DM. Variation at the TERT locus and predisposition for cancer. Expert Rev Mol Med. 2010; 12: e16.
21 McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet. 2008; 40: 1404–6.
22 Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85: 679–91.
23 Nie W, Zang Y, Chen J, Xiu Q. TERT rs2736100 polymorphism contributes to lung cancer risk: a meta-analysis including 49,869 cases and 73,464 controls. Tumour Biol. 2014; 35: 5569–74.
24 Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet. 2009; 41: 221–7.
25 Speedy HE, Di Bernardo MC, Sava GP, Dyer MJ, Holroyd A, Wang Y, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2014; 46: 56–60.
26 Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet. 2013; 45: 371–84.
27 James MA, Vikis HG, Tate E, Rymaszewski AL, You M. CRR9/CLPTM1L regulates cell survival signaling and is required for Ras transformation and lung tumorigenesis. Cancer Res. 2014; 74: 1116–27.
28 Pao W, Girard N. New driver mutations in non-small-cell lung cancer. Lancet Oncol. 2011; 12: 175–80.
29 Petitjean A, Hainaut P, Caron de Fromentel C. TP63 gene in stress response and carcinogenesis: a broader role than expected. Bull Cancer 2006; 93: E126–135.
30 Zhang L, Wang XF, Ma YS, Xia Q, Zhang F, Fu D, et al. Quantitative assessment of the influence of TP63 gene polymorphisms and lung cancer risk: evidence based on 93,751 subjects. PLoS One. 2014; 9: e87004.
31 Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell. 2011; 145: 692–706.
32 Medina PP, Sanchez-Cespedes M. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. Epigenetics. 2008; 3: 64–8.
33 Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. Nat Rev Cancer. 2011; 11: 481–92.
34 Buganim Y, Goldstein I, Lipson D, Milyavsky M, Polak-Charcon S, Mardoukh C, et al. A novel translocation breakpoint within the BPTF gene is associated with a pre-malignant phenotype. PLoS One. 2010; 5: e9657.
35 Kohno T, Kunitoh H, Shimada Y, Shiraishi K, Ishii Y, Goto K, et al. Individuals susceptible to lung adenocarcinoma defined by combined HLA-DQA1 and TERT genotypes. Carcinogenesis. 2010; 31: 834–41.
36 Dong LM, Potter JD, White E, Ulrich CM, Cardon LR, Peters U. Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. JAMA. 2008; 299: 2423–36.
37 Ye Z, Song H, Higgins JP, Pharoah P, Danesh J. Five glutathione s-transferase gene variants in 23,452 cases of lung cancer and 30,397 controls: meta-analysis of 130 studies. PLoS Med. 2006; 3: e91.
38 He XF, Liu LR, Wei W, Liu Y, Su J, Wang SL, et al. Association between the XPG Asp1104His and XPF Arg415Gln Polymorphisms and Risk of Cancer: A Meta-Analysis. PLoS One. 2014; 9: e88490.
39 Zhu J, Hua RX, Jiang J, Zhao LQ, Sun X, Luan J, et al. Association Studies of ERCC1 Polymorphisms with Lung Cancer Susceptibility: A Systematic Review and Meta-Analysis. PLoS One. 2014; 9: e97616.
40 Yin ZB, Hua RX, Li JH, Sun C, Zhu JH, Su X, et al. Smoking and hOGG1 Ser326Cys polymorphism contribute to lung cancer risk: evidence from a meta-analysis. Tumour Biol. 2014; 35: 1609–18.
41 Wilkening S, Bermejo JL, Hemminki K. MDM2 SNP309 and cancer risk: a combined analysis. Carcinogenesis. 2007; 28: 2262–7.
42 Sugimura H, Wakai K, Genka K, Nagura K, Igarashi H, Nagayama K, et al. Association of Ile462Val (Exon 7) polymorphism of cytochrome P450 IA1 with lung cancer in the Asian population: further evidence from a case-control study in Okinawa. Cancer Epidemiol Biomarkers

Prev. 1998; 7: 413–7.
43 Sugimura H, Kohno T, Wakai K, Nagura K, Genka K, Igarashi H, et al. hOGG1 Ser326Cys polymorphism and lung cancer susceptibility. Cancer Epidemiol Biomarkers Prev. 1999; 8: 669–74.
44 Fan R, Wu MT, Miller D, Wain JC, Kelsey KT, Wiencke JK, et al. The p53 codon 72 polymorphism and lung cancer risk. Cancer Epidemiol Biomarkers Prev. 2000; 9: 1037–42.
45 Houlston RS. CYP1A1 polymorphisms and lung cancer risk: a meta-analysis. Pharmacogenetics. 2000, 10: 105–14.
46 Li H, Hao X, Zhang W, Wei Q, Chen K. The hOGG1 Ser326Cys polymorphism and lung cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2008; 17: 1739–45.
47 Dumont P, Leu JI, Della Pietra AC, 3rd, George DL, Murphy M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet 2003; 33: 357–65.
48 Kohno T, Shinmura K, Tosaka M, Tani M, Kim SR, Sugimura H, et al. Genetic polymorphisms and alternative splicing of the hOGG1 gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. Oncogene. 1998; 16: 3219–25.
49 Yamane A, Kohno T, Ito K, Sunaga N, Aoki K, Yoshimura K, et al. Differential ability of polymorphic OGG1 proteins to suppress mutagenesis induced by 8-hydroxyguanine in human cell in vivo. Carcinogenesis. 2004; 25: 1689–94.
50 Bartsch H, haugen A, Risch A, Shields P, Vineis P. Genetic susceptibility. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editons. World Health Organization Classification of Tumors: Pathology and Genetics, Tumours of Lung, Pleura, Thymus and Heart. edn. Lyon: IARC Press; 2004. p. 24–5.
51 Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet. 2014; 46: 736–41.
52 Lovly CM, Carbone DP. Lung cancer in 2010: One size does not fit all. Nat Rev Clin Oncol. 2011; 8: 68–70.
53 Ju YS, Lee WC, Shin JY, Lee S, Bleazard T, Won JK, et al. Fusion of KIF5B and RET transforming gene in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. Genome Res. 2011, 22: 436–45.