

<https://helda.helsinki.fi>

Hot spot mutations in Finnish non-small cell lung cancers

Maki-Nevala, Satu

2016-09

Maki-Nevala , S , Sarhadi , V K , Ronty , M , Kettunen , E , Husgafvel-Pursiainen , K , Wolff , H , Knuutila , A & Knuutila , S 2016 , ' Hot spot mutations in Finnish non-small cell lung cancers ' , Lung Cancer , vol. 99 , pp. 102-110 . <https://doi.org/10.1016/j.lungcan.2016.06.024>

<http://hdl.handle.net/10138/228322>

<https://doi.org/10.1016/j.lungcan.2016.06.024>

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.



Hot spot mutations in Finnish non-small cell lung cancers



Satu Mäki-Nevala^a, Virinder Kaur Sarhadi^a, Mikko Rönty^{b,1}, Eeva Kettunen^c,
Kirsti Husgafvel-Pursiainen^c, Henrik Wolff^c, Aija Knuutila^d, Sakari Knuutila^{a,*}

^a University of Helsinki, Faculty of Medicine, Department of Pathology, P.O. Box 21, 00014 Helsinki, Finland

^b Helsinki University Central Hospital, HUSLAB, Department of Pathology, P.O. Box 400, 00029 Helsinki, Finland

^c Finnish Institute of Occupational Health, P.O. Box 40, 00251 Helsinki, Finland

^d University of Helsinki and Helsinki University Hospital, Heart and Lung Center, Department of Pulmonary Medicine, P.O. Box 340, 00029 Helsinki, Finland

ARTICLE INFO

Article history:

Received 21 April 2016

Received in revised form 7 June 2016

Accepted 27 June 2016

Keywords:

Lung cancer

Non-small cell lung cancer

Mutation

ABSTRACT

Objectives: Non-small cell lung cancer (NSCLC) is a common cancer with a poor prognosis. The aim of this study was to screen Finnish NSCLC tumor samples for common cancer-related mutations by targeted next generation sequencing and to determine their concurrences and associations with clinical features. **Materials and methods:** Sequencing libraries were prepared from DNA isolated from formalin-fixed, paraffin-embedded tumor material of 425 patients using the AmpliSeq Colon and Lung panel covering mutational hot spot regions of 22 cancer genes. Sequencing was performed with the Ion Torrent Personal Genome Machine (PGM).

Results: Data analysis of the hot spot mutations revealed mutations in 77% of the patients, with 7% having 3 or more mutations reported in the Catalogue of Somatic Mutations in Cancer (COSMIC) database. Two of the most frequently mutated genes were *TP53* (46%) and *KRAS* (25%). *KRAS* codon 12 mutations were the most recurrently occurring mutations. *EGFR* mutations were significantly associated with adenocarcinoma, female gender and never/light-smoking history; *CTNNB1* mutations with light ex-smokers, *PIK3CA* and *TP53* mutations with squamous cell carcinoma, and *KRAS* with adenocarcinoma. *TP53* mutations were most prevalent in current smokers and *ERBB2*, *ERBB4*, *PIK3CA*, *NRAS*, *NOTCH1*, *FBWX7*, *PTEN* and *STK11* mutations occurred exclusively in a group of ever-smokers, however the association was not statistically significant. No mutation was found that associated with asbestos exposure.

Conclusion: Finnish NSCLC patients have a similar mutation profile as other Western patients, however with a higher frequency of *BRAF* mutations but a lower frequency of *STK11* and *ERBB2* mutations. Moreover, *TP53* mutations occurred frequently with other gene mutations, most commonly with *KRAS*, *MET*, *EGFR* and *PIK3CA* mutations.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Non-small cell lung cancer (NSCLC) is the most common group of lung cancers with high mortality rates worldwide. In Finland, with a population of 5.5 million, approximately 2500 new cases of lung cancer are diagnosed annually [1]. Lung tumors, even of same histological subtype, can differ widely at the molecular level in their gene expression patterns, epigenetic changes, copy number alterations

and mutations. Several of these genetic markers can be displayed simultaneously and influence the prognosis and treatment outcome [2]. Patients with activated *EGFR* or *ALK* oncogenes benefit from the targeted therapies, such as afatinib, erlotinib and gefitinib for *EGFR*, and crizotinib for *ALK* fusion [3]. Many NSCLC patients also harbor other molecular alterations, such as *MET* amplification or *KRAS* mutations, which might influence the efficacy of a targeted treatment, leading to primary or secondary resistance [4,5]. Thus, it is important to investigate these concurrences of genetic markers in tumors, since they could reveal clinically important variants and/or serve as predictive and prognostic markers.

Ethnicity plays a role in the prevalence of some genetic markers [6]. Due to the genetic isolation of the Finnish population, we have examined these markers in a Finnish patient cohort. We used targeted next generation sequencing to determine mutations in hot spot regions, i.e. those reported in the Catalogue of Somatic Muta-

* Corresponding author.

E-mail addresses: satu.maki-nevala@helsinki.fi (S. Mäki-Nevala), virinder.sarhadi@helsinki.fi (V.K. Sarhadi), mikko.ronty@helsinki.fi (M. Rönty), eeva.kettunen@ttl.fi (E. Kettunen), kirsti.husgafvel-pursiainen@ttl.fi (K. Husgafvel-Pursiainen), henrik.wolff@ttl.fi (H. Wolff), aija.knuutila@hus.fi (A. Knuutila), sakari.knuutila@helsinki.fi (S. Knuutila).

¹ Present address: Fimlab Laboratories, P.O. Box 66, 33101 Tampere, Finland.

tions in Cancer (COSMIC) database [7], 22 lung cancer associated genes in 442 tumor samples of Finnish NSCLC patients.

2. Materials and methods

2.1. Tumor samples

We collected 442 formalin-fixed, paraffin-embedded (FFPE) tumor tissue specimens from NSCLC patients, of these, 425 patient samples were sequenced successfully. Tumor samples were obtained during radical lung cancer surgery in the Hospital District of Helsinki and Uusimaa (HUS), Finland, during 2006–2014. The majority of the samples were material removed during the surgery of the cancer (411 cases) with the remaining 14 obtained by core or fine needle biopsies. All samples were collected before any treatments. The ALK fusion status was tested by immunohistochemistry (IHC) and confirmed by fluorescence in situ hybridization (FISH) in 362 cases [8]. EGFR status, as tested by a real-time PCR kit, was known for 368 samples, due to partly overlapping material with our previous study [9].

Histological subtypes and tumor cell contents of the samples were confirmed by a pulmonary pathologist. Tumor cell content was a minimum of 20% in 91% of the samples, and only eight samples had tumor content less than 10% (overall range 5–90%). The majority of the tumors were adenocarcinoma (67%) or squamous cell carcinoma (21%). The detailed histological subtyping and other clinical characteristics of the patients are presented in Table 1. Patients were grouped according to their smoking status as: never-smokers, ex-light smokers (<20 years and ceased >10 years ago), ex-medium smokers (>20 years and ceased), and current smokers (>20 years and ongoing). Asbestos exposure was evaluated by interview about the occupational history and/or fiber count by SEM on normal lung tissue [10]. Samples with more than 2.0 million asbestos fibers per gram of dried lung tissue, were grouped as asbestos-exposed, and those with less than 1.0 million fibers per gram of dried lung tissue were considered as non-exposed. Samples not fulfilling these criteria were considered as uncertain with regard to asbestos exposure. Survival outcome data were cut-off on March 5, 2016. The mean time period of follow-up was 43.7 months (range 1–141 months). The study was approved by the institutional ethics committee.

2.2. DNA extraction

DNA was extracted by the QIAamp DNA Mini Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's protocol with some changes, as described [11]. DNA samples were quantified with the Qubit® 2.0 fluorometer (Thermo Fisher Scientific) using the Qubit® dsDNA BR assay kit (Thermo Fisher Scientific).

2.3. Next generation sequencing

Sequencing libraries were prepared from 10 ng of DNA samples using the Ion AmpliSeq™ Colon and Lung Cancer Panel together with the Ion AmpliSeq™ Library Preparation Kit (Thermo Fisher Scientific) according to the manufacturer's guidelines. The Colon and Lung Cancer panel contains 504 mutational hot spots (totaling 14.6 kb) of 22 genes: *AKT1*, *ALK*, *BRAF*, *CTNNB1*, *DDR2*, *EGFR*, *ERBB2*, *ERBB4*, *FBXW7*, *FGFR1*, *FGFR2*, *FGFR3*, *KRAS*, *MAP2K1*, *MET*, *NOTCH1*, *NRAS*, *PIK3CA*, *PTEN*, *SMAD4*, *STK11*, and *TP53*. The libraries were quantified with the Qubit® 2.0 Fluorometer, using the Qubit® dsDNA HS assay kit. Templates were prepared and enriched using the Ion OneTouch™ 2 system and the Ion PGM™ Template OT2 200 Kit. Sequencing was performed with the Ion PGM™ Sequencing 200 Kit v2, the Ion 316v2 chip and the Ion Torrent PGM instrument.

Table 1
Characteristics of the patients.

	N (%) N total = 425
Tumor histology	
ADC	283 (66.6)
SCC	89 (20.9)
LCC	34 (8.0)
ADSQ	13 (3.1)
NSCLC, other	6 (1.4)
Tumor cell content, %	
Median	45
Range	5–90
Gender	
Male	239 (56.2)
Age, years	
Median	66.0
Range	40–84
Clinical stage	
Ia	100
Ib	15
IIa	60
IIb	30
IIIa	39
IIIb	7
Ia–IIIa	3
IIIa–IIIb	1
IV	11
Unknown	159
Smoking history	
Never-smokers	38 (8.9)
Smokers	383 (89.4)
Ex-light smokers	32 (7.5)
Ex-medium smokers	151 (35.5)
Current smokers	197 (46.4)
Not available	7 (1.6)
Asbestos exposure	
Exposed	29 (6.8)
Interview + fiber count	0
Fiber count only	3
Interview only	26
Non-exposed	234 (55.1)
Interview + fiber count	6
Fiber count only	22
Interview only	206
Not known/uncertain	162 (38.1)

Abbreviations: ADC, adenocarcinoma; ADSQ, adenosquamous carcinoma; LCC, large cell carcinoma; NSCLC, non-small cell lung cancer; NOS, not otherwise specified; SCC, squamous cell carcinoma.

2.4. Data analysis

The raw sequencing data were aligned to the hg19 and analyzed with TorrentSuite™ Software (v.4.0.2) (Thermo Fisher Scientific). Mutations were detected using the Variant Caller plug-in (v4.0-r76860) (Thermo Fisher Scientific) using default settings, i.e. a quality score of minimum 6, relative read quality of minimum 6.5, coverage of minimum 6 for SNP/COSMIC variant and 15 for indel, and strand bias maximum of 95% for SNP/COSMIC variant and 90% for indel. Further data filtering was conducted to include only those variants that had a frequency of variant reads with a minimum of 3%, and an average phred quality score minimum of 15 for previously known COSMIC mutations and a minimum of 20 for novel variants.

The coverage analysis was performed using the Coverage Analysis plug-in (v4.0-r77897) (Thermo Fisher Scientific). Statistical analyses were performed with IBM SPSS Statistics version 22 (IBM Corp., Armonk, NY). Fisher's exact test was used to compare mutation positive and negative groups with categorical features. Independent *t*-test was used for comparison of mutation statuses

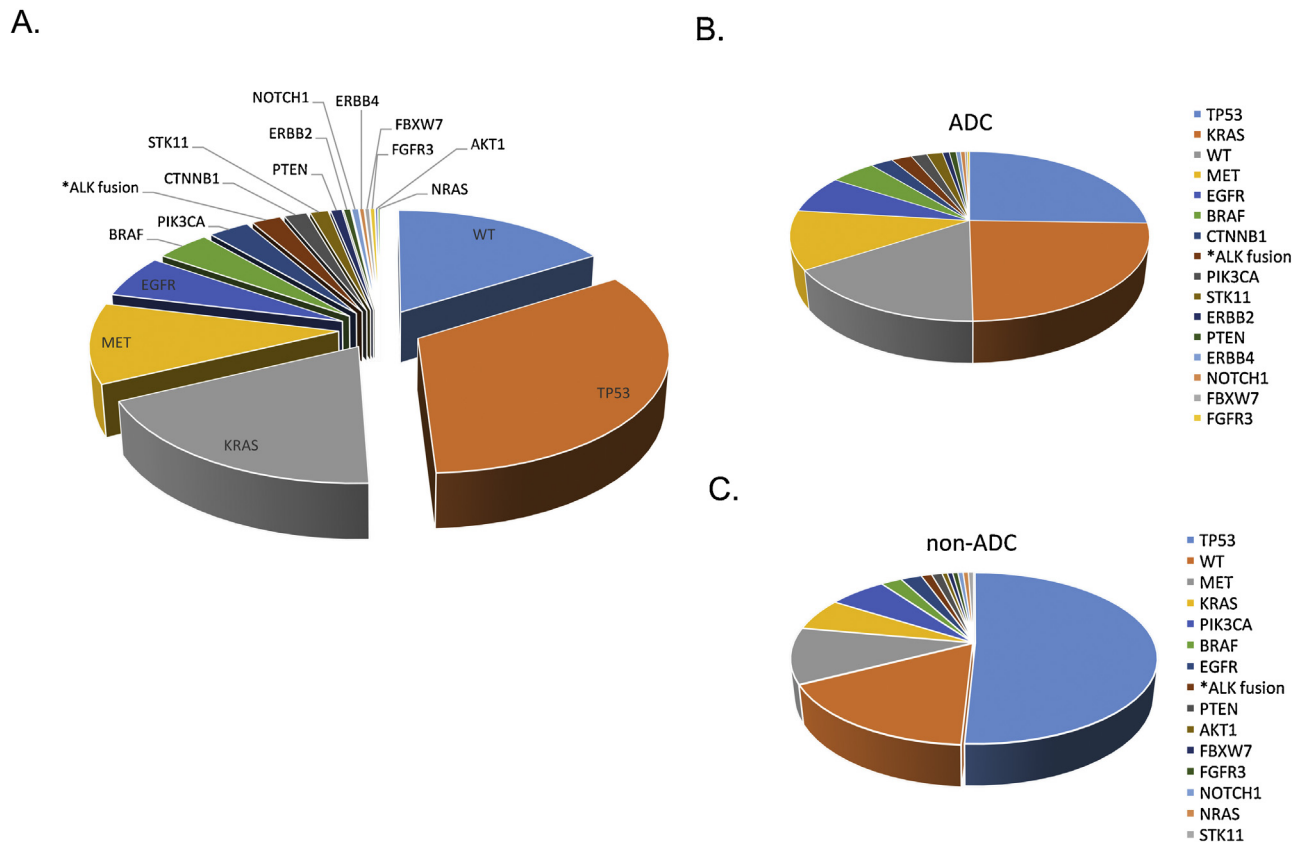


Fig. 1. A pie chart representing hot spot mutations in different genes and ALK fusions proportions in Finnish NSCLC patients. (A.) Mutations in all NSCLC patients ($n = 425$). Twenty three percent of the patients harbored no hot spot mutations (WT, blue section separated). The ALK fusion proportion is calculated from 362 patients. (B.) Mutations in ADC histology ($n = 283$). (C.) Mutations in non-ADC histology ($n = 142$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and age. Two-sided p values < 0.05 were considered statistically significant.

3. Results

Ion Torrent PGM sequencing was performed on 442 samples, with the results from 425 NSCLC samples being successfully completed and presented here. The samples not included were mainly those with low sequencing coverage/libraries. The samples included in the final sequencing analysis had an average coverage of 1623, (range 146–4137) and a uniformity of coverage of 98.7% (range 87.7–100%). Only non-synonymous variants were included.

The majority (77%) of the tumor samples had one or more mutations listed in the COSMIC database (Fig. 1); 46% carried one mutation, 25% two mutations, 4% three mutations, 2% four mutations, two cases had five mutations, one case displayed seven mutations and one case had as many as 15 mutations. Mutation frequencies differed between ADC and non-ADC histologies, as represented in Fig. 1 (b and c). *TP53* was the most frequently mutated gene, followed by *KRAS* (Fig. 2). No COSMIC mutations were seen in the following genes included in the Colon and Lung Cancer panel: *ALK*, *DDR2*, *FGFR1*, *FGFR2*, *MAP2K1* and *SMAD4*. However, non-synonymous novel mutations predicted as deleterious were detected in those genes (Supplemental File S1 and Fig. 2a). The top most recurrent mutations seen in tumors are presented in Fig. 2b.

3.1. COSMIC mutations

The COSMIC mutations detected in the tumors are shown in Table 2, and *TP53* mutations in Supplemental File S2. *TP53* was

the most frequently mutated gene with 46% of the patients harboring one or more *TP53* mutations. *KRAS* mutations were the second most common, with mutations in codons 12 and 13 being the most recurrent (25%). Two patients carried double Gly12Cys and Gly12Val *KRAS* mutations. The concurrencies of the different gene mutations are presented in Table 3. *KRAS* mutations (codon 12/13) were seen together with many other gene mutations and also with *EGFR* mutations His870Tyr in one case and Leu858Arg + Ala871Thr + Gly735Ser in one other.

One sample was exceptional, having multiple mutations in many different genes that included *KRAS* Gly13Cys; *EGFR* Leu858Arg, Ala871Thr and Gly735Ser; *BRAF* Ser605Asn; *CTNNB1* Pro44Ser; *ERBB4* His618Tyr; *PIK3CA* Met1043Ile. This sample was from a 72 years old female, who was an ex-medium-smoker, non-exposed to asbestos and had a radically operated adenocarcinoma; she was still living 47 months after her operation.

EGFR mutations were detected in 8% of the samples. One of these was a triple mutant, as described above; there were three double *EGFR* mutants: Leu858Arg + Asp761Tyr, Leu858Arg + Arg776Cys, and Gly719Ser + Ser768Ile. Clinically relevant exon 19 deletions, exon 20 insertions, Leu858Arg or Gly719Ala were seen together with *TP53* mutations in 17 patients. Moreover, Leu858Arg was detected simultaneously with *MET* Asn375Ser in four cases, and Gly719Ala with *CTNNB1* Ser37Phe in one case.

BRAF mutations were found in 5.6%, but only four cases harbored the Val600Glu mutation (0.1%) and no concurrent mutations were detected in these cases. Other mutations were detected in residues 466, 469 (exon 11), 586, 594, 597, 601 and 605 (exon 15). Two patients with *BRAF* mutations (Ser605Asn or Gly466Val) also harbored activating *KRAS* (codon 12 or 13) mutations.

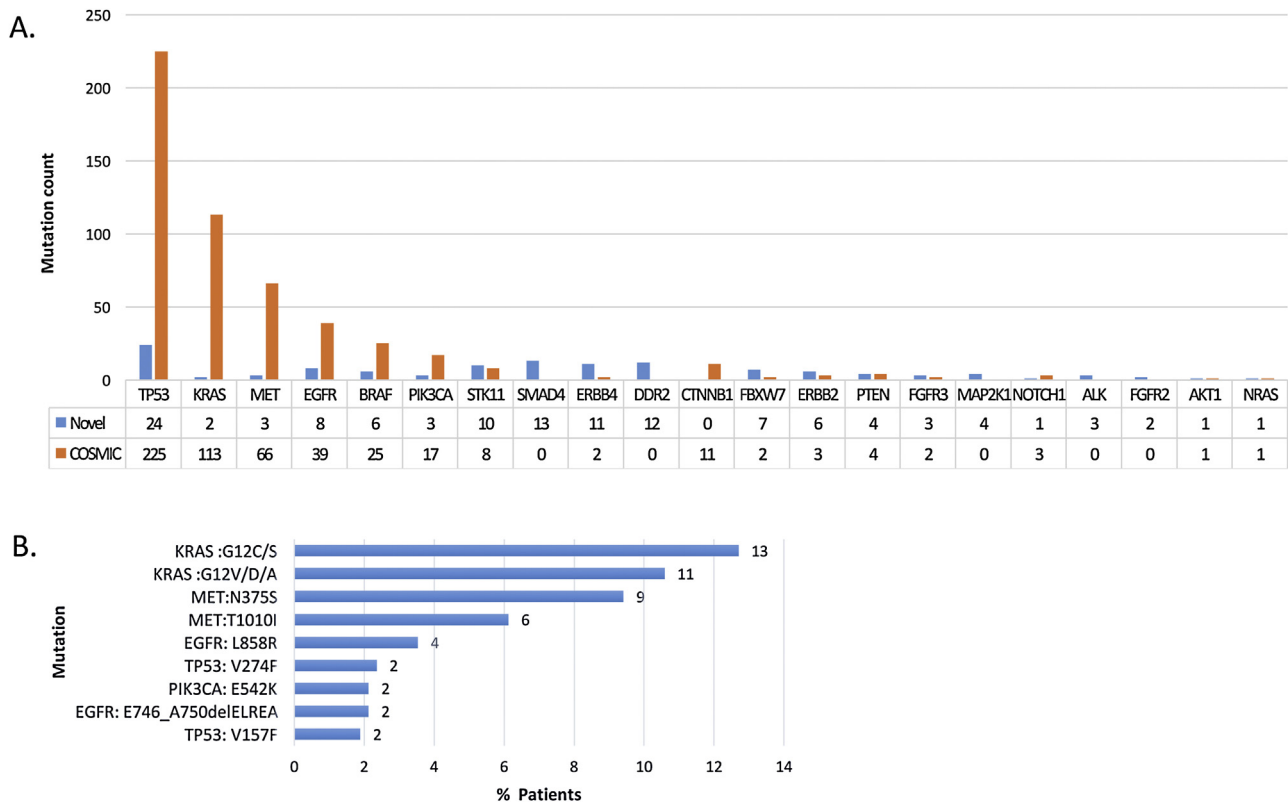


Fig. 2. (A.) A histogram of all non-synonymous hot spot COSMIC (orange, right bar) mutations and novel (blue, left bar) mutations predicted as being damaging by both SIFT and PROVEAN analysis (for novel mutations) in NSCLC patients. (B.) Recurrent hot spot mutations in Finnish NSCLC patients (N = 425). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

TP53 mutations were the most frequent concurrent gene mutations, occurring together with nearly all mutated genes (except for *NRAS* or *FBXW7*). Approximately 50% of the *EGFR* and *BRAF* and 35% of *KRAS* mutation positive samples also harbored a *TP53* mutation. *MET* mutations were detected in 33% of *PIK3CA* and 17% of *RAS* mutants.

3.2. Novel variants

There were 195 novel mutations that led to an amino acid change (including missense, nonsense and indels) which have not been reported in dbSNP (Build 146). The novel variants were analyzed in-silico with SIFT [12] and PROVEAN [13] tools to check the effect of the variants on protein function. Almost half (93) of the 195 novel missense variants were predicted to be deleterious/damaging by both SIFT and PROVEAN. Eleven variants were nonsense variants and 20 were in-frame/frameshift variations, thus also potentially damaging. All of the novel variants resulting in amino acid change are listed in Supplemental File S1 and all 124 deleterious missense, nonsense and in-frame/frameshift mutations in different genes are shown in Fig. 2a. Two novel exon 19 deletions of *EGFR*, one 11 bp insertion/frameshift variation of *ERBB2* and one 21 bp in-frame deletion in *TP53* were among the novel variants.

3.3. COSMIC mutations and their association with clinical characteristics

Statistically significant associations were seen between *EGFR* mutations and adenocarcinoma histology ($p=0.037$). Similarly, never and light ex-smokers ($p<0.001$) and female patients ($p<0.001$) harbored the majority of the *EGFR* mutations. Moreover, *PIK3CA* ($p=0.016$) and *TP53* ($p<0.001$) mutations were most

prevalent in squamous cell carcinoma and *KRAS* in adenocarcinoma ($p<0.001$). *CTNNB1* mutations tended to be associated with a group of light ex-smokers ($p=0.011$) (Table 4).

3.4. Patient survival

Survival data were available for 420 patients with a median follow-up of 43.7 months. Due to a short follow-up time, Kaplan-Meier analyses were not applicable. As an observational result, 53% (216/406) of the operated patients were alive; a total of 165 lung cancer-related death events occurred. Two out of 14 non-operated patients (14%) were alive at the end of follow-up. In adeno-, adenosquamous, squamous cell and large cell carcinomas, there were cancer related deaths observed in 34%, 46%, 28% and 62%, respectively. Lung cancer related death occurred in 53% (16/30) of the patients with more than three COSMIC mutations, whereas the proportions in other patients were: two mutations –20% (21/104), one mutation –37% (70/191) and no mutations –36% (34/95).

4. Discussion

In present study, the most frequently mutated genes were, in descending order: *TP53*, *KRAS*, *MET*, *EGFR*, *BRAF*, *PIK3CA* and *CTNNB1*. In other genes; *STK11*, *PTEN*, *ERBB2*, *NOTCH1*, *ERBB4*, *FBXW7*, *FGFR3*, *AKT1* and *NRAS*, COSMIC mutations were present in less than 2% of the patients. A majority of *EGFR* mutation statuses ($n=368$) were validated by golden standard PCR, results of which were compared in more detail in our previous study [14]. Briefly, a great concordance of 99.7% was obtained. Thus, we believe the results of Ion Torrent PGM NGS are trustful. Moreover, we were able to detect the known *EGFR* and *KRAS* mutations with good quality valued even in those samples with lower tumor content (<20%). However,

Table 2
Hot spot mutations in NSCLC patients.

Gene	Total no. of patients ^a	Mutation CDS	Mutation AA	Sample count		
EGFR	34	c.2155G>A	p.G719S	1		
		c.2156G>C	p.G719A	3		
		c.2203G>A	p.G735S	1		
		c.2235_2249del15	p.E746_A750delELREA	9		
		c.2240_2257del18	p.L747_P753>S	1		
		c.2281G>T	p.D761Y	1		
		c.2300C>T	p.A767V	1		
		c.2303G>T	p.S768I	1		
		c.2311_2312insGCGTGGACA	p.D770_N771insSVD	2		
		c.2320G>A	p.V774M	1		
		c.2326C>T	p.R776C	1		
		c.2573T>G	p.L858R	15		
		c.2608C>T	p.H870Y	1		
		c.2611G>A	p.A871T	1		
		ERBB2	3	c.2305G>C	p.D769H	1
				c.2326_2327insTGT	p.G776>VC	1
				c.2524G>A	p.V842I	1
		ERBB4	2	c.1825G>A	p.D609N	1
				c.1852C>T	p.H618Y	1
		FGFR3	2	c.1150T>C	p.F384L	1
MET	65	c.1138G>A	p.G380R	1		
		c.1124A>G	p.N375S	40		
		c.3029C>T	p.T1010I	26		
AKT1	1	c.49G>A	p.E17K	1		
PIK3CA	18	c.1624G>A	p.E542K	9		
		c.1635G>C	p.E545D	1		
		c.3059C>T	p.A1020V	1		
		c.3129G>A	p.M1043I	1		
		c.3140A>G/T	p.H1047R/L	6		
		c.34G>T/A	p.G12C/S	54		
KRAS	111	c.35G>T/A/C	p.G12V/D/A	45		
		c.37G>T	p.G13C	5		
		c.38G>A/T	p.G13D/V	4		
		c.57G>T	p.L19F	1		
		c.176C>G	p.A59G	1		
		c.183A>C	p.Q61H	2		
		c.351A>T	p.K117N	1		
		c.182A>T	p.Q61L	1		
		NRAS	1	c.1397G>T	p.G466V	2
				c.1405G>A	p.G469R	5
				c.1406G>C	p.G469A	4
				c.1756G>A	p.E586K	2
				c.1780G>A	p.D594N	2
				c.1781A>G	p.D594G	1
c.1790T>G	p.L597R			1		
c.1799T>A	p.V600E			4		
c.1801A>G	p.K601E			2		
c.1814G>A	p.S605N			2		
CTNNB1	10	c.110C>G/T	p.S37C/F	4		
		c.115G>A	p.A39T	1		
		c.130C>T	p.P44S	2		
		c.133T>C	p.S45P	1		
		c.134C>T	p.S45F	1		
		c.98C>G/T	p.S33C/F	2		
NOTCH1	3	c.4732_4734delGTG	p.V1578delV	3		
FBWX7	2	c.1393C>T	p.R465C	1		
		c.1513C>G	p.R505G	1		
PTEN	5	c.1003C>T	p.R335*	1		
		c.209+1delGT	p.?	1		
		c.209+1G>T	p.?	1		
STK11	8	c.517C>T	p.R173C	1		
		c.109C>T	p.Q37*	2		
		c.169delG	p.E57fs*7	1		
		c.580G>A	p.D194N	1		
TP53	196	c.837delC	p.P281fs*6	4		
		Supplemental file S2	Supplemental file S2			

^a Some samples carry multiple mutations.

we might have missed some mutations in few (n = 8) samples with tumor content less than 10%. Unfortunately, normal material was not available for testing the somatic nature of mutations, thus, some of the mutations reported might be of germ line origin, although they have been reported in COSMIC.

The frequency of activating *EGFR* mutations (7%; COSMIC 8%) was similar or slightly higher than that in other studies examining Western or Central European NSCLC patients (4.9–13.8%) [15–19]. *KRAS* mutations were the second most common mutations after *TP53*. ADCs were mutated at a frequency of 35%, while 26% of all

Table 3
Concurrences of the COSMIC hotspot mutations.^a

Gene	Total% (N)	EGFR	ERBB2	ERBB4	FGFR3	MET ^b	AKT1	PIK3CA	KRAS	NRAS	BRAF	CTNNB1	NOTCH1	FBWX7	PTEN	STK11
<i>EGFR</i>	8.0 (34/425)															
<i>ERBB2</i>	0.7 (3/425)	0														
<i>ERBB4</i>	0.5 (2/425)	1	0													
<i>FGFR3</i>	0.5 (2/425)	0	0	0												
<i>MET^b</i>	15.3 (65/425)	4	1	0	0											
<i>AKT1</i>	0.2 (1/425)	0	0	0	0	0										
<i>PIK3CA</i>	4.2 (18/425)	1	0	1	0	6	0									
<i>KRAS</i>	26.1 (111/425)	2	1	2	0	18	0	3								
<i>NRAS</i>	0.2 (1/425)	0	0	0	0	1	0	0	0							
<i>BRAF</i>	5.6 (24/425)	1	0	1	0	3	0	1	2	0						
<i>CTNNB1</i>	2.4 (10/425)	2	0	2	0	2	0	1	5	0	2					
<i>NOTCH1</i>	0.7 (3/425)	0	0	0	0	0	0	1	1	0	1	0				
<i>FBWX7</i>	0.5 (2/425)	0	0	0	0	0	0	1	1	0	0	0	0			
<i>PTEN</i>	1.2 (5/425)	0	0	0	0	0	0	0	0	0	0	1	0	0		
<i>STK11</i>	1.9 (8/425)	0	0	0	0	0	0	0	2	0	0	0	0	0	0	
<i>TP53</i>	46.4 (197/425)	18	1	2	2	25	1	11	39	0	12	6	2	0	5	3

^a Numbers represent the number of patients having concurrent mutations in the corresponding gene pairs. Some patients harbored mutations in multiple genes. Total percentages are calculated from the number of cases harboring the mutation in a particular gene out of the total number of cases (N = 425).

^b MET mutations detected in this study were Asn375Ser and Thr1010Ile, which are described in NCBI dbSNP as germline, and in COSMIC database being confirmed as somatic.

Table 4The frequency (% (N)) of the COSMIC mutations in different subgroups. Statistical significance ($p < 0.05$) is marked with an asterisk (*) and p value is given.

	<i>EGFR</i>	<i>ERBB2</i>	<i>ERBB4</i>	<i>FGFR3</i>	<i>MET</i>	<i>AKT1</i>	<i>PIK3CA</i>	<i>KRAS</i>	<i>NRAS</i>	<i>BRAF</i>	<i>CTNNB1</i>	<i>NOTCH1</i>	<i>FBWX7</i>	<i>PTEN</i>	<i>STK11</i>	<i>TP53</i>
Histology	*0.037	1.000	1.000	0.557	0.861	0.334	*0.016	*<0.001	0.125	0.080	0.341	0.098	0.557	0.794	0.489	*<0.001
ADC	11 (30)	1 (3)	1 (2)	0.4 (1)	16 (46)	00	3 (7)	35 (100)	0	7 (20)	4 (10)	1 (2)	0.4 (1)	1 (3)	2 (7)	37 (105)
ADSQ	0	0	0	0	7 (1)	0	0	23 (3)	0	8 (1)	0	0	0	0	0	54 (7)
LCC	3 (1)	0	0	0	18 (6)	0	3 (1)	12 (4)	3 (1)	3 (1)	0	0	0	0	0	59 (20)
SCC	2 (2)	0	0	1 (1)	14 (12)	1 (1)	10 (9)	2 (2)	0	1 (3)	0	0	1 (1)	2 (2)	1 (1)	69 (61)
other	17 (1)	0	0	0	0	0	17 (1)	33 (2)	0	17 (1)	0	17 (1)	0	0	0	67 (4)
Smoking 1	*<0.001	0.377	0.221	0.307	0.510	^a	0.542	0.720	1.000	0.237	*0.011	1.000	0.437	0.863	0.097	0.135
NS	16 (6)	0	0	3 (1)	18 (7)	0	0	18 (7)	0	11 (4)	3 (1)	0	0	0	0	32 (12)
LS	25 (8)	3 (1)	3 (1)	0	22 (7)	0	6 (2)	25 (8)	0	9 (3)	13 (4)	0	0	0	3 (1)	38 (12)
MS	8 (12)	1 (1)	0	0	13 (20)	0	4 (6)	28 (42)	0	4 (6)	1 (2)	1 (1)	1 (2)	2 (3)	4 (6)	47 (71)
CS	4 (8)	1 (1)	1 (1)	1 (1)	14 (28)	0	5 (10)	27 (53)	1 (1)	5 (10)	2 (3)	1 (2)	0	1 (2)	1 (1)	50 (99)
Smoking 2	0.109	1.000	1.000	0.174	0.478	^a	0.392	0.334	1.000	0.146	1.000	1.000	1.000	0.619	0.463	0.061
Never	16 (6)	0	0	3 (1)	18 (7)	0	0	18 (7)	0	11 (4)	3 (1)	0	0	0	0	32 (38)
Ever	8 (28)	1 (3)	1 (2)	0.3 (1)	15 (55)	0	5 (18)	27 (103)	0.3 (1)	5 (19)	2 (9)	1 (3)	1 (2)	1 (5)	2 (8)	48 (182)
Asbestos	0.146	1.000	1.000	0.110	0.436	1.000	1.000	0.819	1.000	1.000	1.000	0.297	1.000	1.000	1.000	0.118
Asb-	9 (22)	1 (2)	0.4 (1)	0	17 (40)	0.4 (1)	5 (11)	25 (58)	0.4 (1)	6 (15)	3 (7)	1 (2)	1 (2)	1 (2)	1 (3)	46 (108)
Asb+	0	0	0	3 (1)	10 (3)	0	3 (1)	21 (6)	0	3 (1)	0	3 (1)	0	0	0	62 (18)
Gender	*<0.001	0.584	0.191	0.507	0.345	1.000	0.647	1.000	1.000	0.145	0.754	1.000	0.507	0.071	0.085	0.624
Female	14 (26)	1 (2)	1 (2)	0	17 (32)	0	4 (8)	26 (49)	0	8 (14)	3 (5)	1 (1)	0	0	1 (1)	48 (89)
Male	3 (8)	0.4 (1)	0	1 (2)	14 (33)	0.4 (1)	4 (10)	26 (62)	0.4 (1)	4 (10)	2 (5)	1 (2)	1 (2)	2 (5)	3 (7)	45 (108)

Abbreviations: ADC, adenocarcinoma; ADSQ, adenosquamous carcinoma; CS, current smoker; LCC, large cell carcinoma; LS, light ex-smoker; MS, medium ex-smoker; NS, never-smoker; SCC, squamous cell carcinoma.

^a No statistical assessment, because mutation occurred in a case with a missing value.

NSCLC cases carried a *KRAS* mutation, which is in concordance with the previous studies on European NSCLC patients (15–33%) [15,17]. Activating *KRAS*, *EGFR*, *BRAF* mutations and *ALK* fusions have been postulated to be mutually exclusive [6], and this proposal is confirmed in our study. Only one case, carrying 15 different mutations in seven genes, harbored both activating *KRAS* and *EGFR* mutations. Two cases harbored concurrent *KRAS* and *BRAF* (Gly466Val and Ser605Asn) mutations, that have been previously reported [20,21].

Even though the frequency of most mutations was similar as in other studies on Caucasian patients, there were two genes, *BRAF* and *MET* that did differ. The proportion of *BRAF* mutations was 6%, which is higher than reported in a recent meta-analysis (3%, range 0.9–8.9%) [21], or in the studies examining European patients (2.1–3.5%) [21,23,24]. The highest frequency of 8.9% in the meta-analysis can be explained mainly due to selection of only *EGFR*, *KRAS*, *PIK3CA*, *HER2* and *ALK* wild-type patients in that study [25]. Although Val600Glu mutations have been reported to account for approximately half of the *BRAF* mutations in NSCLC patients [21], here the proportion of Val600Glu mutations was much lower (4/24, 17%), and the Val600Glu mutation frequency (0.9%) was lower than that reported previously, such as in Norwegian patients (1.7%) [26]. *BRAF* mutations other than Val600Glu seem to be more prevalent in Finnish NSCLC patients. The clinical significance of these mutations still remains partly unresolved, although some of them seem to be kinase activating and thus might play an important role in driving tumorigenesis [27–29].

MET mutations were seen in 16% of ADCs, which is more than double than reported elsewhere (7%) [30]. Both *MET* mutations detected in this study, Asn375Ser and Thr1010Ile, have been reported as germline in NCBI dbSNP and as somatic in COSMIC. The frequency of Thr1010Ile was three time higher in our NSCLC patients than reported for the general Finnish population in 1000 Human Genomes (NCBI 1000 Genomes Browser): 6.1% (26/425) vs. 2.0%, but for Asn375Ser very similar: 9.4% (40/425) vs. 11%. Unfortunately, because of the lack of normal tissue material, we were unable to confirm their somatic/germline origin.

In this study, the frequency of hot spot mutations in *HER2* (0.7% vs. 1.6–6%) [6,30–33] and *STK11* (1.7% vs. 16–19%) [6,34,35] was lower than described elsewhere. *STK11* mutations have been reported to be associated with Caucasian ethnicity [6,33]. The lower frequency in our study could be partly explained by the fact that our NGS panel covered only the mutation hotspot regions of the genes [34,35]. However, two out of eight *STK11* mutations were seen together with *KRAS*, and five with *TP53*. A similar trend of concurrent *STK11* + *TP53* and *KRAS* + *STK11* was reported in a recent meta-analysis on Western ADCs [6].

HER2 mutations are believed to be more prevalent in adenocarcinoma, and never-smokers [6,32,36]. All three hot spot mutations detected in this study occurred in adenocarcinoma, but in ever-smokers (two medium ex-smokers and one current smoker). In addition to hot spot mutations, six novel mutations were detected, which increases *HER2* mutation frequency to 2%, which is more similar to previous reports [6,36].

In our patients, the frequency of *TP53* mutations was as high as 46.4%. The frequency of *TP53* mutations has been reported to be around 40% (17–68%) among smokers and 21% (0–40%) among non-smokers in a meta-analysis of *TP53* mutations in lung cancer patients [37]. We found clear evidence of concurrence of mutated *TP53* with many other mutant genes, especially with *EGFR* and *KRAS*, in our patients. Nearly half of the *EGFR* and *BRAF* mutants and 35% of the *KRAS*/*NRAS* mutants harbored one or more concomitant *TP53* mutations. Earlier reports have described similar concurrent *TP53* mutations with *EGFR*, *KRAS* and *STK11* mutations although at much lower frequencies (less than 10%) [6]. The high frequency of mutated *TP53* found here might be one reason for its higher concurrence with other gene mutations.

We also evaluated the association of smoking history, asbestos exposure, histological subtype, age, gender and survival with mutation status. *EGFR* mutations were associated with female gender, never-smokers and adenocarcinoma subtype as reported previously [6,31,38]. *CTNNB1* mutations were significantly associated with ex-light-smokers, and as far as we are aware, this association has not been reported previously. COSMIC mutations in *ERBB2*, *ERBB4*, *PIK3CA*, *NRAS*, *NOTCH1*, *FBWX7*, *PTEN* and *STK11* occurred exclusively in ever-smokers, but without statistical significance. *TP53* mutations were most prevalent in the current smokers, but similarly, statistical significance was not reached ($p=0.135$). In a recent meta-analysis, it was concluded that smokers have higher risk for *p53* mutations [37].

When evaluating the association of mutation status and asbestos-exposure, *EGFR* mutations were seen more frequently in the non-exposed group (9% vs. 0%, $p=0.146$) and *TP53* mutations in the asbestos-exposed group (62% vs. 46%, $p=0.118$). However, in our previous study, *EGFR* mutations were seen in both asbestos-exposed (interview only) and non-exposed groups (11% vs. 10%, $p=0.498$) [9]. The frequency of *MET* and *KRAS* mutations was similar in both non-exposed and exposed groups, 17% vs. 10% and 25% vs. 21%, respectively. This is concordant with our previous findings that *KRAS* mutations are not associated with asbestos exposure [39]. However, we were unable to find any mutations significantly associated with asbestos exposure due to the small number of asbestos-exposed patients and the high proportion of uncertain/only partial information regarding asbestos-exposure.

PIK3CA and *TP53* mutations were associated with SCCs, which is concordant with previous studies [30,40,41]. In agreement with a meta-analysis [22], *BRAF* mutations showed no significant association with histology, gender or smoking status. However, similarly to the meta-analysis, *BRAF* mutations were more likely to occur in females (8% vs. 4% in male) and never-smokers (11% vs. 5% in ever-smokers).

The majority of our patients were in the early operable stage of disease and their follow-up time was short, therefore, no clear conclusions can be drawn regarding the association between mutations and survival. *TP53* [42–44] and *KRAS* mutations [45], and multiple mutations [46], have been reported to predict poorer cancer-related overall survival. However, a similar trend was observed, i.e. 53% of the patients with 3 or more hot spot mutations, died of lung cancer in comparison to 32% of the patients with less than 3 mutations.

5. Conclusions

This hot spot mutation screening of a large Finnish NSCLC cohort suggests that although mutation profiles do not differ essentially from other Western patients, mutations in *BRAF* are seen more frequently, and those in *STK11* and *ERBB2* less often in Finnish NSCLC patients. Mutations in *TP53* were the most frequent (46%) and they tended to occur along with *EGFR* and *KRAS* mutations. As a novel finding, we observed *CTNNB1* mutations to be significantly associated with ex-light smokers.

Conflict of interest

None declared.

Acknowledgements

We thank Milja Tikkanen for the technical and clinical data collection support, and Tiina Wirtanen, Päivi Tuominen, Jaana Kierikki and Helinä Hämäläinen for the technical support. We are also grateful to Lauri Lehtimäki for helping to handle the data, Leo Lahti for

his help with statistical matters and Ewen Mac Donald for language editing. The study was funded by the Finnish Work Environment Fund (no. 115112 for SK, no. 111100 for KHP, no. 112269 for HW), the Sigrid Jusélius Foundation, the Cancer Society of Finland, which covered the sequencing costs of this project; and the Finnish Cultural Foundation for Mäki-Nevala's personal grant.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.lungcan.2016.06.024>.

References

- [1] G. Engholm, J. Ferlay, N. Christensen, A.M.T. Kejs, T.B. Johannesen, S. Khan, M. Leinonen, et al., NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries. Version 7.1 (2015), 2015.
- [2] E. Thunnissen, K. van der Oord, M. den Bakker, Prognostic and predictive biomarkers in lung cancer. A review, *Virchows Arch.* 464 (2014) 347–358.
- [3] J. Minguet, K.H. Smith, P. Bramlage, Targeted therapies for treatment of non-small cell lung cancer—recent advances and future perspectives, *Int. J. Cancer* 138 (2016) 2549–2561.
- [4] R.C. Doebele, A.B. Pilling, D.L. Aisner, et al., Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer, *Clin. Cancer Res.* 18 (2012) 1472–1482.
- [5] T.E. Stinchcombe, Novel agents in development for advanced non-small cell lung cancer, *Ther. Adv. Med. Oncol.* 6 (2014) 240–253.
- [6] S. Dearden, J. Stevens, Y.L. Wu, D. Blowers, Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap), *Ann. Oncol.* 24 (2013) 2371–2376.
- [7] S.A. Forbes, D. Beare, P. Gunasekaran, et al., COSMIC: exploring the world's knowledge of somatic mutations in human cancer, *Nucleic Acids Res.* 43 (2015) D805–11.
- [8] K. Tuononen, M. Kero, S. Mäki-Nevala, et al., ALK fusion and its association with other driver gene mutations in Finnish non-small cell lung cancer patients, *Genes Chromosomes Cancer* 53 (2014) 895–901.
- [9] S. Mäki-Nevala, M. Ronty, M. Morel, et al., Epidermal growth factor receptor mutations in 510 Finnish non-small-cell lung cancer patients, *J. Thorac. Oncol.* 9 (2014) 886–891.
- [10] T. Tuomi, Fibrous minerals in the lungs of mesothelioma patients: comparison between data on SEM, TEM, and personal interview information, *Am. J. Ind. Med.* 21 (1992) 155–162.
- [11] K. Tuononen, S. Mäki-Nevala, V.K. Sarhadi, et al., Comparison of targeted next-generation sequencing (NGS) and real-time PCR in the detection of EGFR, KRAS, and BRAF mutations on formalin-fixed, paraffin-embedded tumor material of non-small cell lung carcinoma—superiority of NGS, *Genes Chromosomes Cancer* 52 (2013) 503–511.
- [12] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, *Nat. Protoc.* 4 (2009) 1073–1081.
- [13] Y. Choi, G.E. Sims, S. Murphy, J.R. Miller, A.P. Chan, Predicting the functional effect of amino acid substitutions and indels, *PLoS One* 7 (2012) e46688.
- [14] S. Mäki-Nevala, A. Knuutila, S. Knuutila, et al., Concordant results of epidermal growth factor receptor mutation detection by real-time polymerase chain reaction and Ion Torrent technology in non-small cell lung cancer, *J. Clin. Respir. Dis. Care* 2 (2016) 107.
- [15] C. Boch, J. Kollmeier, A. Roth, et al., The frequency of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC): routine screening data for central Europe from a cohort study, *BMJ Open* 3 (2013) Print 2013, <http://dx.doi.org/10.1136/bmjopen.2013-002560>.
- [16] S. Gahr, R. Stoehr, E. Geissinger, et al., EGFR mutational status in a large series of Caucasian European NSCLC patients: data from daily practice, *Br. J. Cancer* 109 (2013) 1821–1828.
- [17] A.J. Smits, J.A. Kummer, J.W. Hinrichs, et al., EGFR and KRAS mutations in lung carcinomas in the Dutch population: increased EGFR mutation frequency in malignant pleural effusion of lung adenocarcinoma, *Cell Oncol.* 35 (2012) 189–196.
- [18] B.G. Skov, E. Hogdall, P. Clementsen, et al., The prevalence of EGFR mutations in non-small cell lung cancer in an unselected Caucasian population, *APMIS* 123 (2015) 108–115.
- [19] R. Ramlau, T. Cufer, P. Berzinec, et al., Epidermal growth factor receptor mutation-positive non-small-cell lung cancer in the real-world setting in central Europe: the INSIGHT study, *J. Thorac. Oncol.* 10 (2015) 1370–1374.
- [20] T. Kinno, K. Tsuta, K. Shiraishi, et al., Clinicopathological features of non-small cell lung carcinomas with BRAF mutations, *Ann. Oncol.* 25 (2014) 138–142.
- [21] C. Tissot, S. Couraud, R. Tanguy, P.P. Bringuier, N. Girard, P.J. Souquet, Clinical characteristics and outcome of patients with lung cancer harboring BRAF mutations, *Lung Cancer* 91 (2016) 23–28.
- [22] D. Chen, J.F. Huang, K. Liu, et al., BRAFV600E mutation and its association with clinicopathological features of colorectal cancer: a systematic review and meta-analysis, *PLoS One* 9 (2014) e90607.
- [23] K. Schmid, N. Oehl, F. Wrba, R. Pirker, C. Pirker, M. Filipits, EGFR/KRAS/BRAF mutations in primary lung adenocarcinomas and corresponding locoregional lymph node metastases, *Clin. Cancer Res.* 15 (2009) 4554–4560.
- [24] A. Marchetti, L. Felicioni, S. Malatesta, et al., Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations, *J. Clin. Oncol.* 29 (2011) 3574–3579.
- [25] M. Ilie, E. Long, V. Hofman, et al., Diagnostic value of immunohistochemistry for the detection of the BRAFV600E mutation in primary lung adenocarcinoma Caucasian patients, *Ann. Oncol.* 24 (2013) 742–748.
- [26] O.T. Brustugun, A.M. Khattak, A.K. Tromborg, et al., BRAF-mutations in non-small cell lung cancer, *Lung Cancer* 84 (2014) 36–38.
- [27] H. Davies, G.R. Bignell, C. Cox, et al., Mutations of the BRAF gene in human cancer, *Nature* 417 (2002) 949–954.
- [28] K.B. Dahlman, J. Xia, K. Hutchinson, et al., BRAF(L597) mutations in melanoma are associated with sensitivity to MEK inhibitors, *Cancer Discov.* 2 (2012) 791–797.
- [29] V. De Falco, R. Giannini, A. Tamburrino, et al., Functional characterization of the novel T599I-VKSRdel BRAF mutation in a follicular variant papillary thyroid carcinoma, *J. Clin. Endocrinol. Metab.* 93 (2008) 4398–4402.
- [30] Cancer genome atlas research network comprehensive molecular profiling of lung adenocarcinoma, *Nature* 511 (2014) 543–550.
- [31] H. Shigematsu, L. Lin, T. Takahashi, et al., Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers, *J. Natl. Cancer Inst.* 97 (2005) 339–346.
- [32] M.E. Arcila, J.E. Chaft, K. Nafa, et al., Prevalence, clinicopathologic associations, and molecular spectrum of ERBB2 (HER2) tyrosine kinase mutations in lung adenocarcinomas, *Clin. Cancer Res.* 18 (2012) 4910–4918.
- [33] J. Mazieres, S. Peters, B. Lepage, et al., Lung cancer that harbors an HER2 mutation: epidemiologic characteristics and therapeutic perspectives, *J. Clin. Oncol.* 31 (2013) 1997–2003.
- [34] J.P. Koivunen, J. Kim, J. Lee, et al., Mutations in the LKB1 tumour suppressor are frequently detected in tumours from Caucasian but not Asian lung cancer patients, *Br. J. Cancer* 99 (2008) 245–252.
- [35] N. Pecuchet, P. Laurent-Puig, A. Mansuet-Lupo, et al., Different prognostic impact of STK11 mutations in non-squamous non-small-cell lung cancer, *Oncotarget* (2015) (in press).
- [36] H. Shigematsu, T. Takahashi, M. Nomura, et al., Somatic mutations of the HER2 kinase domain in lung adenocarcinomas, *Cancer Res.* 65 (2005) 1642–1646.
- [37] X. Liu, X.J. Lin, C.P. Wang, et al., Association between smoking and p53 mutation in lung cancer: a meta-analysis, *Clin. Oncol. (R. Coll. Radiol.)* 26 (2014) 18–24.
- [38] W. Pao, V. Miller, M. Zakowski, et al., EGF receptor gene mutations are common in lung cancers from never smokers and are associated with sensitivity of tumors to gefitinib and erlotinib, *Proc. Natl. Acad. Sci. U. S. A.* 101 (2004) 13306–13311.
- [39] S. Mäki-Nevala, V.K. Sarhadi, A. Knuutila, et al., Driver gene and novel mutations in asbestos-exposed lung adenocarcinoma and malignant mesothelioma detected by exome sequencing, *Lung* 194 (2015) 125–135.
- [40] Cancer genome atlas research network comprehensive genomic characterization of squamous cell lung cancers, *Nature* 489 (2012) 519–525.
- [41] S. Li, L. Li, Y. Zhu, et al., Coexistence of EGFR with KRAS, or BRAF, or PIK3CA somatic mutations in lung cancer: a comprehensive mutation profiling from 5125 Chinese cohorts, *Br. J. Cancer* 110 (2014) 2812–2820.
- [42] T. Mitsudomi, N. Hamajima, M. Ogawa, T. Takahashi, Prognostic significance of p53 alterations in patients with non-small cell lung cancer: a meta-analysis, *Clin. Cancer Res.* 6 (2000) 4055–4063.
- [43] T. Sioris, K. Hუსgafvel-Pursiainen, A. Karjalainen, et al., Survival in operable non-small-cell lung cancer: role of p53 mutations, tobacco smoking and asbestos exposure, *Int. J. Cancer* 86 (2000) 590–594.
- [44] E. Steels, M. Paesmans, T. Berghmans, et al., Role of p53 as a prognostic factor for survival in lung cancer: a systematic review of the literature with a meta-analysis, *Eur. Respir. J.* 18 (2001) 705–719.
- [45] W. Pan, Y. Yang, H. Zhu, Y. Zhang, R. Zhou, X. Sun, KRAS mutation is a weak, but valid predictor for poor prognosis and treatment outcomes in NSCLC: A meta-analysis of 41 studies, *Oncotarget* 7 (2016) 8373–8388.
- [46] E. Bria, S. Pilotto, E. Amato, et al., Molecular heterogeneity assessment by next-generation sequencing and response to gefitinib of EGFR mutant advanced lung adenocarcinoma, *Oncotarget* 6 (2015) 12783–12795.