

Hotspot Mutations Detectable by Next-generation Sequencing in Exhaled Breath Condensates from Patients with Lung Cancer

OMAR YOUSSEF¹, AIJA KNUUTTILA^{2,3,4}, PÄIVI PIIRILÄ⁵,
TOM BÖHLING^{1,6}, VIRINDER SARHADI¹ and SAKARI KNUUTTILA¹

¹Department of Pathology, Faculty of Medicine, University of Helsinki, Helsinki, Finland;
²Departments of ²Pulmonary Medicine, ³Heart and Lung Center, ⁴Cancer Center,
Helsinki University Hospital and University of Helsinki, Helsinki, Finland;
⁵Unit of Clinical Physiology, HUS-Medical Imaging Center,
Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland;
⁶HUSLAB, Helsinki University Hospital, Helsinki, Finland

Abstract. *Background:* Genetic alterations occurring in lung cancer are the basis for defining molecular subtypes and essential for targeted therapies. Exhaled breath condensate (EBC) is a form of non-invasive sample that, amongst components, contains DNA from pulmonary tissue. Next-generation sequencing (NGS) was herein used to analyze mutations in EBC from patients with lung cancer. *Materials and Methods:* EBC was collected from 26 patients with cancer and 20 healthy controls. Amplicon-based sequencing using Ion Ampliseq Colon and Lung Cancer gene panel v2 was applied. *Results:* The sequencing was successful in 17 patients and 20 controls. EBC from patients revealed 39 hotspot mutations occurring in: adenomatous polyposis coli (APC), v-raf murine sarcoma viral oncogene homolog B (BRAF), discoidin domain receptor tyrosine kinase 2 (DDR2), epidermal growth factor receptor (EGFR), erb-b2 receptor tyrosine kinase 4 (ERBB4), F-box and WD repeat domain containing 7 (FBXW7), fibroblast growth factor receptor 1 (FGFR1), FGFR3 (fibroblast growth factor receptor 3), Kirsten rat sarcoma viral oncogene homolog (KRAS), mitogen-activated protein kinase kinase 1 (MAP2K1), met proto-oncogene (MET), neuroblastoma RAS viral (v-ras) oncogene homolog (NRAS), phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA), phosphatase and tensin homolog (PTEN), ret proto-oncogene (RET), SMAD family member 4

(SMAD4), serine/threonine kinase II (STK11), and tumor protein p53 (TP53) genes. EBC from controls revealed 35 hotspot mutations. The average mutant allele fraction was higher in patients than controls. *Conclusion:* NGS can identify mutations in EBCs from patients with lung cancer. This could provide a promising non-invasive method for the assessment of gene mutations in lung cancer.

Lung cancer is one of the most common malignancies worldwide. Lung cancer has a poor prognosis and a long-term survival rate that has not improved in the past 20 years (1). One reason for the poor tumor course and outcome is the very late phase at which the malignancy is diagnosed. At this phase, tumoral molecular changes lead to resistance to multiple forms of treatment (2-5). Sufficient tumor tissue biopsies are essential for making appropriate diagnostic and therapeutic decisions during the course of the disease. However, such biopsies are not always easily obtained. Recently, free circulating DNA has been investigated as a non-invasive method for the detection of numerous genetic alterations associated with malignancy and, thus, offers an alternative source of material for early cancer diagnosis (6-9).

Amongst other components, exhaled breath condensate (EBC) contains DNA that seemingly originates from the tumorous tissue in patients with lung carcinoma (10). Gene mutations occurring at tumor protein p53 (TP53), Kirsten rat sarcoma viral oncogene homolog (KRAS), epidermal growth factor receptor (EGFR), and cyclin-dependent kinase inhibitor 2A (CDKN2A) have been studied in EBC from patients with different lung malignancies (11-14). We recently demonstrated that next-generation sequencing (NGS) can be used to study mutations in EBC from healthy individuals (15). In the present study, we investigated the applicability of EBC in patients with cancer.

This article is freely accessible online.

Correspondence to: Sakari Knuutila, Haartmaninkatu 3, P.O. Box 21, FI-00014, Finland. Tel: +358 504601554, e-mail: sakari.knuutila@helsinki.fi

Key Words: Hotspot mutations, exhaled breath condensate, lung cancer, next-generation sequencing.

Materials and Methods

Patients and controls. EBC samples were collected from 26 patients with different lung neoplasms. We included any type of patient with lung neoplasms who were physically fit to give 15 minutes breathing and gave their consent. Seventeen patients had non-small cell lung carcinoma (NSCLC) (eleven adenocarcinoma and six squamous cell carcinoma), six had small cell lung carcinoma, two had pleural mesothelioma, and one patient had a suspected malignancy but did not have histopathological diagnosis (Table I). EBC samples from eight patients were collected during or after treatment (either chemotherapy, radiotherapy, or surgery). EBC from the remaining 18 patients was collected before any specific cancer treatment (Table II). Patients were diagnosed and treated at Helsinki University Hospital. For controls, EBC was collected from 20 healthy individuals including healthy smokers, ex-smokers, and non-smokers, and excluding those with pulmonary infections *e.g.* common flu. Detailed information and results from controls were reported previously (15). In brief, 10 healthy females (aged 25-56 years) and 10 healthy males (aged 25-68 years) were studied.

The study was approved by the Hospital and Uusimaa (HUS) Ethical Review Board (ethical permission number 253/13/03/01/2015). Written informed consent was obtained from all patients and controls.

EBC sample collection. Each patient provided an EBC sample by breathing through an ECoScreen® (Eric Jaeger, Würzburg, Germany) for 15 minutes. Mean breath volume and breathing frequency were checked every 5 minutes until the end of the collection procedure. Collected EBC samples were transported immediately on ice to the laboratory. Samples were transferred into 2-ml tubes. The volume of EBC was measured for each sample and then samples were stored at -70°C .

DNA extraction. DNA was extracted from the whole EBC sample using QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The kit uses RNA carrier and a vacuum for fast and efficient processing of DNA from the whole sample (1-3 ml). DNA was measured using a Qubit® 2.0 Fluorometer and Qubit® dsDNA HS Assay kit (Thermo Fisher Scientific, Waltham, MA, USA).

NGS: Preparation of libraries. Ten nanograms of EBC DNA were used for preparing amplicon libraries using Ion AmpliSeq™ Library kit 2.0 (Thermo Fisher Scientific) according to the manufacturer's instructions. In order to amplify the template, Ion AmpliSeq Colon and Lung Cancer panel v2 (Thermo Fisher Scientific) was used. The panel consisted of a primer pool of 92 amplicons covering 504 hotspot mutations in 22 genes frequently mutated in lung cancer. The genes included in this panel are AKT serine/threonine kinase 1 (*AKT1*), ALK receptor tyrosine kinase (*ALK*), v-raf murine sarcoma viral oncogene homolog B (*BRAF*), catenin beta 1 (*CTNNB1*), discoidin domain receptor tyrosine kinase 2 (*DDR2*), *EGFR*, erb-b2 receptor tyrosine kinase 2 (*ERBB2*), erb-b2 receptor tyrosine kinase 4 (*ERBB4*), F-box and WD repeat domain containing 7 (*FBXW7*), fibroblast growth factor receptor 1 (*FGFR1*), *FGFR2*, *FGFR3*, *KRAS*, mitogen-activated protein kinase kinase 1 (*MAP2K1*), met proto-oncogene (*MET*), Notch homolog 1, translocation-associated (Drosophila) (*NOTCH1*), neuroblastoma RAS viral (v-ras) oncogene homolog (*NRAS*), phosphatidylinositol-4,5-bisphosphate 3-kinase

catalytic subunit alpha (*PIK3CA*), phosphatase and tensin homolog (*PTEN*), SMAD family member 4 (*SMAD4*), serine/threonine kinase 11 (*STK11*), and *TP53*. Only two samples from mesothelioma patients were amplified by using the Ion AmpliSeq Cancer Hotspot Panel v2 (Life Technologies, California, USA). This panel consisted of primer pool for 207 amplicons covering 2800 mutational hotspot regions in 50 genes, including ret proto-oncogene (*RET*) mutations. The amplified libraries were purified using Agencourt AMPure XP beads (Beckman Coulter Genomics, High Wycombe, UK). Library concentrations were measured using Qubit® dsDNA HS Assay kit and Qubit® 2.0 Fluorometer (Thermo Fisher Scientific). The libraries were stored at -20°C until the sequencing step was performed.

Template preparation and sequencing. The purified libraries were diluted to 100 pM and subsequently amplified on Ion Sphere™ particles (Life Technologies). The templates were prepared and enriched using the Ion OneTouch™ 2 System (Life Technologies), an automated emulsion polymerase chain reaction system. Sequencing was performed on the Ion Personal Genome Machine System (PGM™; Life Technologies) using Ion 316™ chips and Ion PGM™ Sequencing Hi-Q view kit v2.

NGS data analysis. Torrent Suite Software v.5.2.2 (Life Technologies) was used to assess run performance and data analysis. The Integrative Genomics Viewer (IGV v 2.4; Broad Institute, Cambridge, MA, USA) was used for visual inspection of the aligned reads. The Variant Caller plug-in (v5.2) was used for variant calling. The same cutoff score as that used in our healthy control study was applied, with a threshold for quality score ≥ 20 and for mutant allele fraction $\geq 3\%$ (15). This NGS 3% allele fraction cutoff was tested previously for *EGFR* mutations in patients with lung cancer when compared with a clinically approved real-time polymerase chain reaction assay (16). Only single nucleotide variants resulting in a non-synonymous amino acid change, a premature stop codon, and indels resulting in either a frameshift or insertion or deletion of amino acids were selected. All single nucleotide variants were analyzed for previously reported hotspot mutations [somatic mutations reported in the Catalogue of Somatic Mutations in Cancer (COSMIC) database] and novel variations such as new mutations detected by NGS, but not reported in either COSMIC (<https://cancer.sanger.ac.uk/cosmic>) or dbSNP (build 151) databases (<https://www.ncbi.nlm.nih.gov/projects/SNP/>).

Statistical analysis. Data were analyzed using IBM SPSS advanced statistics version 24 (IBM, Armonk, NY, USA). Comparison between the median EBC DNA concentration between two groups was performed using the Mann-Whitney test (non-parametric *t*-test). Independent samples *t*-test was used to compare significant differences in the number of detected mutations between 'before treatment' and the 'during/after treatment' groups. All tests were two-sided and *p*-values of 0.05 or less were considered statistically significant.

Results

Success rate. DNA was successfully isolated from all 26 patients; NGS was successful in 17 patients. Nine EBC samples failed to amplify during sequencing (Table I). In the previously studied healthy controls, DNA was successfully isolated and sequenced in all 20 participants (15).

Table I. Summary of exhaled breath condensate samples collected and sequenced.

Category	Total samples collected	Successful DNA extraction	Successful sequencing	Hotspot mutation only	All mutations (hotspot + novel)
Patients	26	26	17	14	17
NSCLC (AC)	11	11	6	5	6
NSCLC (SCC)	6	6	5	3	5
SCLC	6	6	5	5	5
Mesothelioma	2	2	1	1	1
Suspected lung cancer [†]	1	1	0	0	0
Healthy controls [‡]	20	20	20	15	18

NSCLC: Non-small cell lung carcinoma; AC: adenocarcinoma; SCC: squamous cell carcinoma; SCLC: small cell lung carcinoma. [†]No biopsy confirmation. [‡]See reference (15).

Table II. Sample collection, sequencing, and detected mutations according to treatment status for all patients.

Category	Total samples collected		Successfully sequenced (number of detected mutations)	
	Before treatment	During/after treatment	Before treatment	During/after treatment
NSCLC (AC)	7	4	4 (23)	2 (7)
NSCLC (SCC)	5	1	5 (55)	0 (0)
SCLC	4	2	3 (27)	2 (14)
Mesothelioma	1	1	0 (0)	1 (10)
Suspected malignancy	1	0	0 (0)	0 (0)
Total	18	8	12 (105)	5 (31)

NSCLC: Non-small cell lung carcinoma; AC: adenocarcinoma; SCC: squamous cell carcinoma; SCLC: small cell lung carcinoma.

Hotspot mutations in EBC from patients and controls. For both controls and patients, the same criteria for cutoff were set and mutations with quality ≥ 20 and allelic frequency $\geq 3\%$ were considered to be true mutations. By applying these parameters to patient EBC samples, a total of 39 hotspot mutations that were previously reported as somatic mutations in the COSMIC database and 98 novel mutations (*i.e.* not previously reported) were identified. The most frequent hotspot mutations occurred in the following genes: *TP53* (n=6), *SMAD4* (n=6), *PIK3CA* (n=4), *KRAS* (n=3), *BRAF* (n=2), *ERBB4* (n=2), *FBXW7* (n=2), *MAP2K1* (n=2), *PTEN* (n=2), *STK11* (n=2), *APC* (n=1), *DDR2* (n=1), *EGFR* (n=1), *FGFR1* (n=1), *FGFR3* (n=1), *MET* (n=1), *NRAS* (n=1), and *RET* (n=1).

A total of 35 hotspot variants and 106 novel mutations were detected at an allelic frequency $\geq 3\%$ in EBC from healthy controls, with hotspot mutations occurring at the following genes: *TP53* (11), *KRAS* (4), *NRAS* (3), *SMAD4* (3), *CTNNB1* (2), *MET* (2), *PTEN* (2), *BRAF* (1), *DDR2* (1), *EGFR* (1), *ERBB2* (1), *FBXW7* (1), *FGFR3* (1), *NOTCH1* (1), and *PIK3CA* (1) (15). Detailed information on all hotspot mutations detected in both patients and controls,

along with corresponding mutant allele fraction and quality score, are shown in Table III.

Mutations in corresponding tumor tissue. Out of the 17 successfully sequenced patient EBC samples, NGS reports for tumor tissue from routine clinical diagnostics were available for only four patients with adenocarcinomas (Table IV). Tissue NGS in clinical diagnostic laboratory was performed on Ion Torrent PGM using in-house cancer gene panel 1, which included the following genes: *EGFR*, *KIT*, *KRAS*, *MET*, *NRAS*, platelet-derived growth factor receptor A (*PDGFRA*), *PIK3CA* (all exons), and *BRAF* (exons 11-15). For patient EP-8, tumor DNA revealed mutations in *KRAS* exon 3 (c.183A>C) and *PDGFRA* exon 4 (c.423G>T). *PDGFRA* was not included in our patient panel; the *KRAS* mutation was however not identified in EBC sequencing results. In the tumor DNA of patient EP-16, a mutation in *MET* exon 14 (c.3028+ 3del) was observed, which in our analysis was outside of the amplicon region and was not studied. In the other two patients with *KRAS* exon 2 (c.35 G>A) mutations in their tumor specimens, no corresponding mutations were seen in their EBC samples (Table IV).

Table III. Mutant allele fraction (MAF) and quality score of hotspot mutations detected in patients and controls.

Gene symbol	Gene name	Patients			Controls		
		Mutation	MAF (%)	Quality	Mutation	MAF (%)	Quality
<i>APC</i>	Adenomatous polyposis coli	p.A1582P	99	1825	ND	NA	NA
<i>BRAF</i>	V-raf murine sarcoma viral oncogene homolog B	p.Q461*	19	671	p.K601E	12	134
		p.I457T	7	117			
<i>CTNNB1</i>	Catenin beta 1	ND	NA	NA	p.T41A	4	36
					p.D32Y	11	99
<i>DDR2</i>	Discoidin domain receptor tyrosine kinase 2	p.M629V	6	148	p.R105C	17	296
<i>EGFR</i>	Epidermal growth factor receptor	p.K708R	3	22	p.D761N	12	320
<i>ERBB2</i>	Erb-B2 receptor tyrosine kinase 2	ND	NA	NA	p.V773M	49	5897
<i>ERBB4</i>	Erb-B2 receptor tyrosine kinase 4	p.G237S	6	28	ND	NA	NA
		p.R232*	100	1914			
<i>FBXW7</i>	F-Box and WD repeat domain containing 7	p.G391D	18	57	p.R278*	19	2123
		p.R479*	13	246			
<i>FGFR1</i>	Fibroblast growth factor receptor 1	p.P150S	38	496	ND	NA	NA
<i>FGFR3</i>	Fibroblast growth factor receptor 3	p.A634T	10	278	p.H274Y	7	298
<i>KRAS</i>	Kirsten rat sarcoma viral oncogene homolog	p.G13D	17	393	p.G12V	4	27
		p.D57N	12	374	p.D30Y	5	44
		p.K117R	6	133	p.T20M	4	24
					p.A11V	6	33
<i>MAP2K1</i>	Mitogen-activated protein kinase kinase 1	p.R49C	18	588	ND	NA	NA
		p.F53L	8	39			
<i>MET</i>	Met proto-oncogene	p.N375S	41	439	p.R359L	6	192
					p.S349G	7	50
<i>NOTCH1</i>	Notch homolog 1, translocation-associated (Drosophila)	ND	NA	NA	p.V1578delV	4	76
<i>NRAS</i>	Neuroblastoma RAS viral (v-ras) oncogene homolog	p.V45A	12	210	p.D54G	5	23
					p.A11T	5	24
					p.G60E	25	682
<i>PIK3CA</i>	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha	p.H1047Y	92	3492	p.M1043V	18	513
		p.P421L	100	239			
		p.R537Q	6	59			
		p.G1049A	45	7095			
<i>PTEN</i>	Phosphatase and tensin homolog	p.321 fs	12	180	p.F258L	9	496
		p.Y315H	3	42	p.L57S	100	1677
<i>RET</i>	RET proto-oncogene	p.C609V	96	711	ND	NA	NA
<i>SMAD4</i>	SMAD Family Member 4	p.R497H	26	2003	p.S171L	4	21
		p.509 fs	7	173	p.F408S	10	144
		p.H528R	6	63	p.A452T	56	2322
		p.S191L	20	813			
		p.G352E	16	779			
		p.W509R	5	80			
<i>STK11</i>	Serine/threonine kinase 11	p.W332*	13	951	ND	NA	NA
		p.E33K	21	649			
<i>TP53</i>	Tumor protein p53	p.M246I	39	2106	p.G112S	27	1191
		p.A353T	34	280	p.I251T	4	20
		p.H214Y	15	448	p.P72R	5	2639
		p.R196P	9	249	p.S166P	11	763
		p.G154S	27	1161	p.Q104*	20	354
		p.P219L	13	1066	p.Y163*	11	202
					p.F134L	6	23
					p.R337C	27	1407
					p.N345D	24	1199
					p.M169I	9	496
					p.S269N	7	111

ND: Not detected; NA: not applicable.

Mutation in relation to treatment status. EBC samples from eight patients were collected during or after treatment (chemotherapy, radiotherapy, or surgery); five out of these eight samples were successfully sequenced. The number of

mutations detected in EBC samples was lower in these treated patients compared to the non-treated patients. However, this difference was not statistically significant ($p=0.83$) (Table II).

Table IV. Clinical characteristics of patients with successfully sequenced samples.

ID	Histopathology	Treatment status	Stage at diagnosis	Mutated genes	
				Tissue*	EBC**
EP-1	Mesothelioma (epithelial subtype)	Chemotherapy	IV	NA	<i>ABL1, APC, EGFR, FBXW7, FGFR3, PIK3CA, RET, SMAD4, SMO</i>
EP-2	NSCLC (AC)	Before treatment	IVB	No specific findings	<i>ERBB2, ERBB4, FGFR1, FGFR3, KRAS, PIK3CA, PTEN</i>
EP-3	NSCLC (SCC)	Before treatment	IIIB	NA	<i>DDR2, ERBB4, FGFR3, KRAS, MAP2K1, NRAS, PIK3CA, SMAD4, STK11, TP53</i>
EP-4	SCLC	Before treatment	IIIB	NA	<i>DDR2, FGFR3, KRAS, MET, PIK3CA, SMAD4</i>
EP-6	NSCLC (SCC)	Before treatment	IIIA	NA	<i>ALK, DDR2, ERBB4, FBXW7, FGFR1, SMAD4</i>
EP-7	NSCLC (AC)	Before treatment	IVA	No specific findings	<i>ALK, BRAF, DDR2, EGFR, FGFR3, MAP2K1, NRAS, PIK3CA</i>
EP-8	NSCLC (AC)	Before treatment	IVB	<i>KRAS</i> exon 3 (c.183A>C), <i>PDGFRA</i> exon 4 (c.423G>T)	<i>ERBB4, FGFR3</i>
EP-9	NSCLC (SCC)	Before treatment	IIIA	NA	<i>EGFR, MET, AKT1</i>
EP-10	NSCLC (AC)	Chemotherapy	IVB	<i>KRAS</i> exon 2 (c.35G>A)	<i>ERBB2, FGFR1</i>
EP-16	NSCLC (AC)	Before treatment	IVA	<i>MET</i> exon14 c.3028+3del	<i>PIK3CA, ERBB2</i>
EP-17	NSCLC (AC)	Surgery, chemotherapy, radiotherapy	IIIA	<i>KRAS</i> exon2 (c.35G>A)	<i>NRAS, FBXW7, PTEN, MAP2K1, SMAD4</i>
EP-24	SCLC	Chemotherapy	IVA	NA	<i>MET, SMAD4, TP53</i>
EP-25	SCLC	Before treatment	IVB	NA	<i>AKT1, DDR2, ERBB4, FGFR3, NRAS, PTEN, SMAD4, TP53</i>
EP-26	NSCLC (SCC)	Before treatment	IIIB	NA	<i>ERBB4</i>
EP-27	SCLC	Chemotherapy	IIIA	NA	<i>FBXW7, FGFR3, MAP2K1, MET, NRAS, PTEN, STK11, TP53</i>
EP-28	SCLC	Before treatment	IVB	NA	<i>BRAF, FGFR1, FGFR3, KRAS, MET, SMAD4, TP53</i>
EP-29	NSCLC (SCC)	Before treatment	IIIA	NA	<i>ALK, BRAF, DDR2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, PIK3CA, PTEN, SMAD4, STK11, TP53</i>

NGS: Next-generation sequencing; NA: tissue sequencing not applicable; NSCLC: non-small cell lung carcinoma; AC: adenocarcinoma; SCC: squamous cell carcinoma; SCLC: small cell lung carcinoma; SMO: smoothened, frizzled class receptor. *Gene panel for tissue sequencing included only *EGFR, KIT, KRAS, MET, NRAS, PDGFRA, PIK3CA*, and *BRAF*. **Gene panels for EBC sequencing included 22 genes (Ion Ampliseq Colon and Lung Cancer panel v2) and 50 genes (Ion AmpliSeq Cancer Hotspot Panel v2). See Table III for other gene names.

Discussion

In this study, we showed, to our knowledge for the first time, that NGS can be adopted for analysis of EBC specimens from patients with different lung malignancies.

In our previous study, we optimized the application of NGS to EBC from healthy individuals with a median DNA concentration of 2,240 ng/ml (15). In the current study, we applied the same methodological techniques to EBC from patients with lung cancer to identify driver gene mutations. The success rate in patients was 65.4%; nine samples did not have sufficient sequencing reads even after repeating the sequencing step. The common element between these samples was a low DNA library concentration. Accordingly, we examined DNA concentration as a possible cause for

failed amplification. The difference in median DNA concentration between samples with successful sequencing (1,560 ng/ml) and those with failed sequencing (1,380 ng/ml) was statistically significant ($p=0.037$). When comparing genes frequently mutated in both patients and controls (such as *TP53* and *KRAS*) for hotspot mutations, the average mutant allele fraction was higher in patients than controls (22.9% and 13.6% in *TP53* and 11.4% and 4.3% in *KRAS* in patients and controls, respectively). The presence of mutations in healthy controls does not necessarily have a clinical significance, but instead reflects a physiological homeostatic process and removal of damaged DNA due to continuous cellular turnover in pulmonary tissue (15). Similar findings were highlighted in 2012 by Kordiak *et al.*, who compared the ratio of mutant and wild-type *KRAS* in

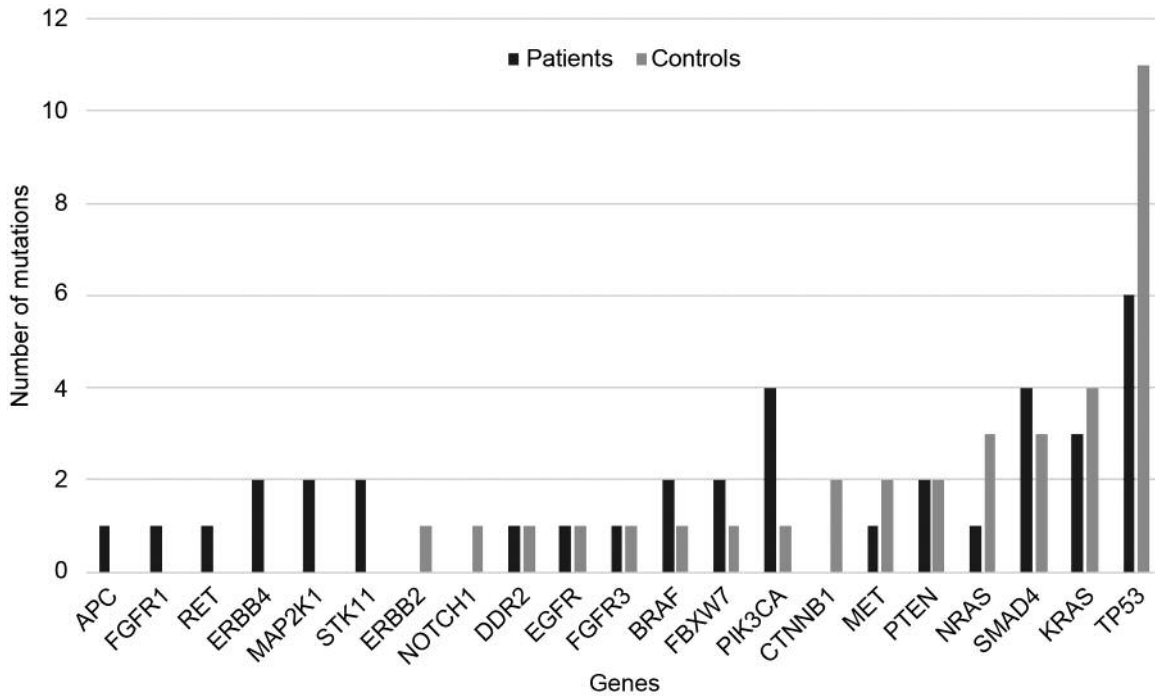


Figure 1. Pattern of hotspot mutations in exhaled breath condensate from patients and controls. See Table III for gene names.

EBC from patients with lung cancer and healthy controls. The authors revealed a significantly higher ratio of mutant:wild-type *KRAS* in patients (0.20 ± 0.03) than in controls (0.028 ± 0.032). They also compared the ratio of mutant:wild-type *KRAS* between tissue tumors and normal pulmonary parenchyma and found significantly higher values in cancerous tissues (12).

Testing for *KRAS* mutational status may have a predictive value in different treatment protocols, as this mutation confers resistance to a wide variety of therapeutic agents, including chemotherapy and inhibitors of angiogenesis and *EGFR* (17). Four patients with adenocarcinomas had NGS results from tumor tissue acquired during routine clinical diagnostics. We compared these results with the corresponding EBC sequencing data. In patient EP-8, *KRAS* exon 3 mutation (c.183A>C) was detected in the tissue specimen, but was not observed in the patient's EBC. Patients EP-10 and EP-17 had a *KRAS* exon 2 (c.35G>A) mutation in their tissue specimens collected at the time of diagnosis (6 months and 18 months prior to EBC sampling, respectively) that were not detected in EBC. However, the EBC samples from these patients were collected when they had already undergone cancer treatment. For patient EP-10, EBC was collected during chemotherapy and VEGFR-inhibitor treatment. For patient EP 17, EBC was collected after radical treatment (lobectomy with adjuvant

chemotherapy and radiotherapy). It not known whether the absence of mutations in EBC is due to treatment effect or other reasons. A previous reports showed a significant decrease in the mutant:wild-type allelic fraction of *KRAS* in EBC after tumor resection (12).

Comparing the pattern of hotspot mutations in EBC between patients and controls illustrated that some genes were mutated only in patient EBC samples (*APC*, *FGFR1*, *RET*, *MAP2K1*, and *STK11*). One of these mutations was *MAP2K1* F53L, which has been reported to be associated with a distinct subset of lung adenocarcinoma and is a possible target for mitogen-activated protein kinase inhibitors (18). Mutations in *BRAF*, *FBXW7*, *PIK3CA*, and *SMAD4* were found in both patients and controls but more frequently in patients (Figure 1). *PIK3CA* gene mutation was found in the patient group at codon 1047 (H1047Y), which is commonly mutated in NSCLC and may mediate resistance to anti-EGFR and other targeted therapies (19, 20).

Although the relevance of the mutational patterns observed in EBC is not yet clear, we feel that this should be reported due to the limited number of such studies on EBC. In addition, hotspot mutations in EBC categorized according to histological tumor types are given in Table V. The presence of hotspot mutations in apparently healthy individuals has also been reported in other studies (21, 22).

Table V. Hotspot mutations in exhaled breath condensate according to tumor type.

Tumor type	Gene	Mutation type	MAF (%)	Quality
Mesothelioma	<i>APC</i>	p.A1582P	99	1825
	<i>FBXW7</i>	p.G391D	18	57
	<i>PIK3CA</i>	p.P421L	100	239
	<i>RET</i>	p.C609V	96	711
NSCLC (AC)	<i>BRAF</i>	p.I457T	7	118
	<i>EGFR</i>	p.K708R	3	23
	<i>FGFR1</i>	p.P150S	38	497
	<i>KRAS</i>	p.G13D	17	393
	<i>MAP2K1</i>	p.F53L	8	40
	<i>NRAS</i>	p.V45A	12	210
	<i>PIK3CA</i>	p.R537Q	6	59
		p.G1049A	45	7095
	<i>PTEN</i>	p.321 fs	12	180
		p.Y315H	3	42
NSCLC (SCC)	<i>DDR2</i>	p.M629V	6	148
	<i>ERBB4</i>	p.R232*	100	1915
	<i>KRAS</i>	p.K117R	6	134
	<i>MAP2K1</i>	p.R49C	18	588
	<i>SMAD4</i>	p.W509R	5	81
		p.509fs	7	173
	<i>STK11</i>	p.W332*	13	952
	<i>TP53</i>	p.M246I	39	2106
SCLC		p.P219L	13	1066
	<i>BRAF</i>	p.Q461*	19	672
	<i>ERBB4</i>	p.G237S	6	29
	<i>FBXW7</i>	p.R479*	13	247
	<i>FGFR3</i>	p.A634T	10	278
	<i>KRAS</i>	p.D57N	12	375
	<i>MET</i>	p.N375S	41	440
	<i>PIK3CA</i>	p.H1047Y	92	3492
	<i>SMAD4</i>	p.S191L	20	813
		p.G352E	16	779
		p.R497H	26	2003
		p.H528R	6	63
	<i>STK11</i>	p.E33K	21	650
	<i>TP53</i>	p.A353T	34	281
		p.H214Y	15	449
		p.R196P	9	250
	p.G154S	27	1162	

MAF: Mutant allele fraction; NSCLC: non-small cell lung carcinoma; AC: adenocarcinoma; SCC: squamous cell carcinoma; SCLC: small cell lung carcinoma. See Table III for gene names.

Accumulation of mutations in somatic cells is a major step in the multistep process of tumorigenesis (23). There is also a correlation between increased mutation rate and increased lifetime incidence of lung cancer. It is believed that every ~30 mutations in the coding region are associated with a 0.2% increase of lifetime incidence of cancer in general (24). In our study, the average number of hotspot mutations in patient EBC samples (2.29) was higher than the average from control EBC samples (1.75). However, this was not

statistically significant ($p=0.292$). Additionally, tissue NGS was not applicable to 15 patients due to the widely accepted clinical practice that only adenocarcinomas, and never smokers/former light smokers with other lung cancer histology are routinely tested. In real-life situations, there is often insufficient tissue material for tumor molecular profiling, which highlights the need for easily accessible samples such as EBC for molecular analysis in such difficult cases. Importantly, EBC might be able to replace the need for tumor re-biopsies in following the disease course and could facilitate detection of new resistance mutations during treatment. EBC may also be an ideal sample for screening of mutations and other genetic alterations in individuals at substantial risk of lung cancer, such as smokers and persons exposed to asbestos.

Our study has limitations, as we only had NGS results from corresponding tumor tissue for four patients to compare mutations seen in EBC. Moreover, two of these patients had undergone treatment since their tissue analysis, which makes it difficult to draw any conclusion regarding the concordance of results between EBC and tumor tissue. Further studies on a larger number of patients and comparison of their tumor and EBC mutation profile is required to assess the accuracy of EBC mutation profiling for diagnostic applications.

Conclusion

In the present study, hotspot mutations were detectable in EBC from patients with different types of lung malignancies. The mutant allele fraction for hotspot mutations was higher in EBC samples from patients than from healthy controls. EBC may provide a helpful tool in cancer diagnostics and molecular profiling. Genomic studies on EBC in lung cancer are still at an early phase. Further investigations are required to test the applicability of NGS for clinical use.

Conflict of Interest

The Authors declare no conflict of interest in regard to this study.

Acknowledgements

The Authors would like to thank Minna Ratilainen and Julia Hämäläinen for their invaluable role and help in EBC sample collection. This study was funded by the Sigrid Jusélius Foundation and the Finnish Work Environment Fund. O.Y. is grateful for the financial support received from the Otto A. Malm Foundation.

References

- 1 Siegel RL, Miller KD and Jemal A: Cancer statistics, 2018. *CA Cancer J Clin* 68: 7-30, 2018.
- 2 Ellis PM and Vandermeer R: Delays in the diagnosis of lung cancer. *J Thorac Dis* 3: 183-188, 2011.

- 3 Kadara H, Scheet P, Wistuba II and Spira AE: Early events in the molecular pathogenesis of lung cancer. *Cancer Prev Res* 9: 518-527, 2016.
- 4 Massion PP and Carbone DP: The molecular basis of lung cancer: Molecular abnormalities and therapeutic implications. *Respir Res* 4: 12, 2003.
- 5 West H, Oxnard GR and Doebele RC: Acquired resistance to targeted therapies in advanced non-small cell lung cancer: New strategies and new agents. *Am Soc Clin Oncol Educ Book*, 2013. doi: 10.1200/EdBook_AM.2013.33.e272.
- 6 Ilie M, Hofman V, Long E, Bordone O, Selva E, Washetine K, Marquette CH and Hofman P: Current challenges for detection of circulating tumor cells and cell-free circulating nucleic acids, and their characterization in non-small cell lung carcinoma patients. What is the best blood substrate for personalized medicine? *Ann Transl Med* 2: 107, 2014.
- 7 Pérez-Callejo D, Romero A, Provencio M and Torrente M: Liquid biopsy-based biomarkers in non-small cell lung cancer for diagnosis and treatment monitoring. *Transl Lung Cancer Res* 5: 455-465, 2016.
- 8 Marmarelis M, Thompson JC, Aggarwal C, Evans TL, Carpenter E, Cohen RB, Langer CJ and Bauml J: Emerging uses of circulating tumor DNA in advanced stage non-small cell lung cancer. *Ann Transl Med* 5: 380, 2017.
- 9 Molina-Vila MÁ: Liquid biopsy in lung cancer: present and future. *Transl Lung Cancer Res* 5: 452-454, 2016.
- 10 Youssef O, Sarhadi VK, Armengol G, Piirilä P, Knuutila A and Knuutila S: Exhaled breath condensate as a source of biomarkers for lung carcinomas. A focus on genetic and epigenetic markers – A mini-review. *Genes Chromosomes Cancer* 55: 905-914, 2016.
- 11 Gessner C, Kuhn H, Toepfer K, Hammerschmidt S, Schauer J and Wirtz H: Detection of *p53* gene mutations in exhaled breath condensate of non-small cell lung cancer patients. *Lung Cancer* 43: 215-222, 2004.
- 12 Kordiak J, Szemraj J, Hamara K, Bialasiewicz P and Nowak D: Complete surgical resection of lung tumor decreases exhalation of mutated *KRAS* oncogene. *Respir Med* 106: 1293-1300, 2012.
- 13 Zhang D, Takigawa N, Ochi N, Tanimoto Y, Noujima D, Chen YY, Tanimoto M and Kiura K: Detection of the *EGFR* mutation in exhaled breath condensate from a heavy smoker with squamous cell carcinoma of the lung. *Lung Cancer* 73: 379-380, 2011.
- 14 Chen JL, Chen JR, Huang FF, Tao GH, Zhou F and Tao YJ: Analysis of p16 gene mutations and their expression using exhaled breath condensate in non-small-cell lung cancer. *Oncol Lett* 10: 1477-1480, 2015.
- 15 Youssef O, Knuutila A, Piirilä P, Böhling T, Sarhadi V and Knuutila S: Presence of cancer-associated mutations in exhaled breath condensates of healthy individuals by next generation sequencing. *Oncotarget* 8: 18166-18176, 2017.
- 16 Mäki-Nevala S, Knuutila A, Knuutila S and Sarhadi VK: 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.
- 17 Matikas A, Mistriotis D, Georgoulis V and Kotsakis A: Targeting *KRAS*-mutated non-small cell lung cancer: A history of failures and a future of hope for a diverse entity. *Crit Rev Oncol Hematol* 110: 1-12, 2017.
- 18 Arcila ME, Drlon A, Sylvester BE, Lovly CM, Borsu L, Reva B, Kris MG, Solit DB and Ladanyi M: *MAP2K1 (MEK1)* mutations define a distinct subset of lung adenocarcinoma associated with smoking. *Clin Cancer Res* 21: 1935-1943, 2015.
- 19 Vansteenkiste JF, Canon JL, De Braud F, Grossi F, De Pas T, Gray JE, Su WC, Felip E, Yoshioka H, Gridelli C, Dy GK, Thongprasert S, Reck M, Aimone P, Vidam GA, Roussou P, Wang YA, Tomaso ED and Soria JC: Safety and efficacy of buparlisib (BKM120) in patients with PI3K pathway-activated non-small cell lung cancer. *J Thorac Oncol* 10: 1319-1327, 2015.
- 20 Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, Bergethon K, Shaw AT, Gettinger S, Cosper AK, Akhavanfard S, Heist RS, Temel J, Christensen JG, Wain JC, Lynch TJ, Vernovsky K, Mark EJ, Lanuti M, Iafrate AJ, Mino-Kenudson M and Engelman JA: Genotypic and histological evolution of lung cancers acquiring resistance to *EGFR* inhibitors. *Sci Transl Med* 3: 75ra26, 2011.
- 21 Gormally E, Vineis P, Matullo G, Veglia F, Caboux E, Le Roux E, Peluso M, Garte S, Guarrera S, Munnia A, Airoidi L, Astrup H, Malaveille C, Dunning A, Overvad K, Tjønneland A, Lund E, Clavel-Chapelon F, Boeing H, Trichopoulou A, Palli D, Krogh V, Tumino R, Panico S, Bueno-de-Mesquita HB, Peeters PH, Pera G, Martinez C, Dorronsoro M, Barricarte A, Navarro C, Quirós JR, Hallmans G, Day NE, Key TJ, Saracci R, Kaaks R, Riboli E and Hainaut P: *TP53* and *KRAS2* mutations in plasma DNA of healthy subjects and subsequent cancer occurrence: a prospective study. *Cancer Res* 66: 6871-6876, 2006.
- 22 Yakubovskaya MS, Spiegelman V, Luo FC, Malaev S, Salnev A, Zborovskaya I, Gasparyan A, Polotsky B, Machaladze Z and Trachtenberg AC: High frequency of *K-RAS* mutations in normal appearing lung tissues and sputum of patients with lung cancer. *Int J Cancer* 63: 810-814, 1995.
- 23 Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA and Kinzler KW: Cancer genome landscapes. *Science* 339: 1546-1558, 2013.
- 24 Hao D, Wang L and Di L: Distinct mutation accumulation rates among tissues determine the variation in cancer risk. *Sci Rep* 6: 19458, 2016.

Received August 27, 2018
 Revised September 10, 2018
 Accepted September 13, 2018