Lung Cancer 130 (2019) 50–58

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Mutation patterns in a population-based non-small cell lung cancer cohort and prognostic impact of concomitant mutations in *KRAS* and *TP53* or *STK11*

Linnéa La Fleur[a], Elin Falk-Sörqvist[a], Patrik Smeds[a], Anders Berglund[b], Magnus Sundström[a], Johanna SM Mattsson[a], Eva Brandén[c], Hirsh Koyi[c], Johan Isaksson[a,c], Hans Brunnström[d], Mats Nilsson[e], Patrick Micke[a], Lotte Moens[a], Johan Botling[a,*]

[a] Dept. of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Sweden
[b] Epistat, Uppsala, Sweden
[c] Dept. of Respiratory Medicine, Gävle Hospital, Gävle, Centre for Research and Development, Uppsala university/County Council of Gävleborg, Gävle, Sweden
[d] Dept. of Clinical Sciences Lund, Division of Oncology and Pathology, Lund University, Lund, Sweden
[e] Dept. of Biochemistry and Biophysics, Stockholm University, Sweden

ARTICLE INFO

Keywords:
Non-small cell lung cancer
Targeted resequencing
Mutation patterns
KRAS
TP53
STK11

ABSTRACT

Objectives: Non-small cell lung cancer (NSCLC) is a heterogeneous disease with unique combinations of somatic molecular alterations in individual patients, as well as significant differences in populations across the world with regard to mutation spectra and mutation frequencies. Here we aim to describe mutational patterns and linked clinical parameters in a population-based NSCLC cohort.

Materials and methods: Using targeted resequencing the mutational status of 82 genes was evaluated in a consecutive Swedish surgical NSCLC cohort, consisting of 352 patient samples from either fresh frozen or formalin fixed paraffin embedded (FFPE) tissues. The panel covers all exons of the 82 genes and utilizes reduced target fragment length and two-strand capture making it compatible with degraded FFPE samples.

Results: We obtained a uniform sequencing coverage and mutation load across the fresh frozen and FFPE samples by adaption of sequencing depth and bioinformatic pipeline, thereby avoiding a technical bias between these two sample types. At large, the mutation frequencies resembled the frequencies seen in other western populations, except for a high frequency of *KRAS* hotspot mutations (43%) in adenocarcinoma patients. Worse overall survival was observed for adenocarcinoma patients with a mutation in either *TP53*, *STK11* or *SMARCA4*. In the adenocarcinoma *KRAS*-mutated group poor survival appeared to be linked to concomitant *TP53* or *STK11* mutations, and not to *KRAS* mutation as a single aberration. Similar results were seen in the analysis of publicly available data from the cBioPortal. In squamous cell carcinoma a worse prognosis could be observed for patients with *MLL2* mutations, while *CSMD3* mutations were linked to a better prognosis.

Conclusion: Here we have evaluated the mutational status of a NSCLC cohort. We could not confirm any survival impact of isolated driver mutations. Instead, concurrent mutations in *TP53* and *STK11* were shown to confer poor survival in the *KRAS*-positive adenocarcinoma subgroup.

1. Introduction

Lung cancer is the leading cause of cancer related deaths worldwide and has a poor prognosis with a five-year survival of only 16.8% [1,2]. Based on histology, lung cancer is separated into two main subsets, small cell carcinoma and non-small cell lung cancer (NSCLC) [3]. The two main histological NSCLC subsets, adenocarcinoma and squamous cell carcinoma belong to the top three cancer types with regard to the prevalence of somatic mutations [4]. The Cancer Genome Atlas (TCGA) has conducted comprehensive genome studies of NSCLC, displaying a great diversity of molecular variations [5,6]. Except for *TP53* and the *KEAP1/NFE2L2* axis, few genes are frequently altered in both histologic subtypes. In adenocarcinoma, molecular targeted therapies have been introduced for patients with activating mutations in *EGFR* or translocated *ALK*, *ROS1* and *RET* [7–10]. Mutated *BRAF*, *ERBB2* exon 20 insertions and mutations causing *MET* exon 14 skipping can also be

* Corresponding author at: Department of Immunology Genetics and Pathology, Rudbeck Laboratory, Uppsala University, SE, 751 85, Uppsala, Sweden.
E-mail address: johan.botling@igp.uu.se (J. Botling).

https://doi.org/10.1016/j.lungcan.2019.01.003
Received 21 August 2018; Received in revised form 3 January 2019; Accepted 8 January 2019

targeted successfully [11–15]. However, the majority of adenocarcinomas harbors an activating mutation in *KRAS* or lacks an identifiable oncogenic driver. Squamous cell carcinoma, although it is a heavily mutated malignancy, lacks common recurrent hotspot mutations in targetable driver genes [3]. Potential targets in squamous cell carcinoma that have been considered are mutated *DDR2* and amplified *FGFR1* [16,17]. Other frequently mutated genes include *MLL2*, *CDKN2A* and *PIK3CA* [6].

For most patients with solid tumors only formalin fixed paraffin embedded (FFPE) samples are available in routine healthcare. This sampling bias is a problem for research on mutation spectra and associated biological and clinical impact in population-based clinical cohorts. Formalin fixation leads to fragmentation of the DNA and can induce artificial mutations during PCR amplification [18,19]. Hence, in order to include all samples of a patient cohort for sequencing the techniques needs to be adapted to samples with poor DNA quality.

The first aim of this study was to evaluate the mutational status in a representative NSCLC cohort of patients treated with surgery at a single center. The goal was to avoid sample selection bias and to achieve population-based coverage by combined use of fresh frozen and FFPE samples in an optimized and robust targeted resequencing approach. The second aim was to assess the prognostic impact of mutations in individual genes as well as impact of co-mutation patterns, in relation to known clinical patient parameters.

2. Material and methods

2.1. Patient cohort

The patient material used for this study consisted of either fresh frozen or FFPE tissue from primary tumors from 352 NSCLC patients who underwent surgical resection at Uppsala University Hospital between 2006 and 2010 [20]. All fresh specimens were routinely processed for fresh frozen biobanking at the department of Pathology at Uppsala University Hospital, whereas the FFPE tissue blocks were obtained after fixation and further processing for routine diagnostics. In addition, non-neoplastic lung tissue samples from the specimens was included in the study for 35 of the patients. Lung pathologists performed re-evaluation of the original cases in accordance to the fourth edition of the WHO Classification of Lung Tumors [21]. Clinical parameters were obtained from the records of the population-based Uppsala-Örebro Regional Lung Cancer Registry - age at diagnosis, sex, smoking history, tumor stage, performance status (according to WHO) and overall survival (defined as time from surgery until date of death or last follow up). Information regarding adjuvant therapy was retrieved from patient records. The study was conducted according to the Swedish Biobank Legislation and Ethical Review Act (reference 2012/532, Ethical Review Board in Uppsala).

2.2. DNA preparation

DNA was extracted from either sections of fresh frozen tissue (200 samples) using the QIAamp DNA Mini Kit (Qiagen, Hamburg, Germany) or from 1 mm FFPE tissue cores (152 samples), using the QIAamp DNA FFPE Tissue Kit (Qiagen). DNA quantification was conducted with Qubit dsDNA BR assay kit (Life Technologies, Thermo Fisher Scientific Inc., Sunnyvale, CA, USA).

2.3. Target enrichment and sequencing

Using the Haloplex target enrichment system (Agilent Technologies, Santa Clara, CA, USA) target enrichment was performed with 200 ng input DNA. An assay targeting the coding exons of 82 genes was designed for this study. The design covers a total region of approximately 0.47Mb and uses 39,328 probes to capture the targeted genes. The design coverage and mean read depth for the individual genes is provided in Supplementary table S1. The panel design utilized reduced target fragment length and two-strand capture making it compatible with degraded FFPE samples, as previously described [22]. The enriched and amplified samples concentrations were determined using D1000 Screen Tapes (Agilent). Subsequently the samples were pooled in equimolar amounts and thereafter purified using the Agencourt AMPure XP system (Beckman Coulter, Indianapolis, IN, USA). The pooled samples were then sequenced on the Illumina HiSeq 2500 platform.

2.4. Data analysis

2.4.1. Alignment

CutAdapt version 1.8.0 was used to remove incorporated parts of the Illumina sequencing adaptors, which happens on target fragments shorter than the read length (150bp). Reads which had a read length of shorter than 1 bp after trimming were excluded from the analysis. The remaining trimmed reads were aligned to the reference genome (hg19, October 2010 assembly) using BWA version 0.7.12 (BWA settings; algorithm BWA-MEM, number of threads 3).

2.4.2. Variant detection

Detection of single-nucleotide variants was performed with SNPmania version 0.0.7. Only reads with a minimum base quality of 20 and alignment quality of 5 were considered during variant detection. The variants were then annotated using Annovar version 2015Mar22. To lower the number of possible causative variants a filtering against normal samples was done. Variants listed in the European 1000 genome database were excluded from further analysis, as well as all variants detected in thirty-five sequenced non-neoplastic lung samples from the study cohort. Only variants present on both the plus and minus strand, with at least a read depth of 30, predicted by Annovar to be non-synonymous and having a minimum variant allele frequency (VAF) of 5% were kept for further analysis, except for common hotspots in the genes *EGFR*, *KRAS*, *BRAF*, *PIK3CA, ERBB2*, *MET* and *NRAS* and variants located within the genes TP53 and STK11. For these genes variants with a VAF down to 1% were included.

2.4.3. Large indels

Large insertions and deletions (larger than 5 base pairs) were detected with Pindel version 0.2.5a8. The complete bam-file from Bwa was used for a non-targeted analysis with Pindel using the following parameters: max_range_index 2 and balance_cutoff 100. Indels detected on both the plus and minus strand with at least a read depth of 30 and a VAF of 5% were kept and a filtering was applied where all indels found in the non-neoplastic samples were excluded. The remaining indels were annotated using Annovar.

2.5. Statistics

Overall survival was analyzed using the Kaplan-Meier method and comparison of survival curves was done using the log-rank test. Multivariate Cox regression models were used to evaluate the relative risk of death of all causes, expressed as hazard ratios (HR) with 95% confidence intervals (CI). The multivariate models were adjusted for gender, age at surgery, pathological stage, smoking status, performance status and adjuvant therapy. Clinical characteristics for the patients at diagnosis are presented using standard descriptive statistics. Fisher´s exact test was used to analyze observed differences in frequencies, comparing the groups with mutated *KRAS*, *EGFR*, *TP53, STK11, SMARCA4, APC, MLL2* and *CSMD3* to cases with a wild-type phenotype in the respective genes. Multiple testing adjustments of significance levels were not performed. Statistical significance was set to $p < 0.05$. All analyses were performed using R version 3.2.0.

Table 1
Clinical characteristics of the NSCLC cases included in this study.

	n (%)
All patients	352 (100)
Gender	
Men	175 (49.7)
Women	177 (50.3)
Age	
≤70	231 (65.6)
> 70	121 (34.4)
Stage	
IA	142 (40.3)
IB	75 (21.3)
IIA	41 (11.6)
IIB	34 (9.7)
IIIA	50 (14.2)
IV	10 (2.8)
Smoking	
Current	179 (50.9)
Ex > 1year	134 (38.1)
Never	39 (11.1)
Histology	
Adenocarcinoma	225 (63.9)
Squamous cell carcinoma	102 (29.0)
Large cell carcinoma	8 (2.3)
Large-cell neuroendocrine carcinoma	9 (2.6)
Adenosquamous	5 (1.4)
Sarcomatoid	3 (0.9)
Performance status	
0	211 (59.9)
1	136 (38.6)
2	5 (1.4)

3. Results

3.1. Clinical samples and histopathologic data

In this study 352 NSCLC patients were included. Of these, 225 were classified as adenocarcinoma and 102 as squamous cell carcinomas. For thirty-five of the tumor samples a matched sample from adjacent, histologically non-neoplastic lung tissue was analyzed. A summary of clinical parameters is displayed in Table 1.

3.2. Technical evaluation of sequencing strategy

The 82 genes selected for our study (Fig. 1A) have been shown to harbor genetic alterations that are believed to be relevant for the disease progression in NSCLC and present at a frequency of at least 1% in genome-wide studies of lung cancer cases [5,6]. The gene panel covered 99.87% of the targeted coding regions and the sequencing yielded an average read depth of 4409 reads/base within the region of interest. In total, 2469 non-synonymous genomic alterations could be detected across the 352 lung cancer samples (Supplementary table S2).

DNA from either fresh frozen (n = 200) or FFPE (n = 152) tumor samples was extracted for library preparation and targeted resequencing. A higher fraction of short sequencing reads (< 100bp) was obtained from FFPE samples, while longer reads (> 400bp) were almost completely absent (Fig.1B). Still, by sequencing the FFPE samples at a greater depth and adaption of the bioinformatics pipeline for variant calling we obtained a uniform sequencing coverage (Supplementary figure S1) and mutational load (Fig. 1C). The mean mutational load was 20.1 mutations/Mb in the fresh frozen samples and 18.8 mutations/Mb in the FFPE samples. Thus, technical bias was avoided with regard to the two sample types. Notably, the mean mutational load in our study is higher compared to other studies using whole exome sequencing. This is due to enrichment of commonly mutated genes included in the design of the targeted gene panel.

3.3. Mutation frequencies in lung tumors

The mutation frequencies in adenocarcinoma and squamous cell carcinoma for the 82 genes included in the study are listed in Supplementary table S3. The most frequently altered gene in both adenocarcinoma and squamous cell carcinoma was *TP53*, mutated in 47% and 85% respectively. Out of the total 245 *TP53* variants 67% was considered pathogenic, 26% likely pathogenic and 7% variant of uncertain significance according to the TP53 database Seshat [23]. Either *KEAP1* or *NFE2L2* was mutated in 19% of adenocarcinoma cases and 23% of squamous cell carcinoma cases, in a mutually exclusive fashion [24]. Other frequently mutated genes in adenocarcinoma were *KRAS* (44%), *EGFR* (19%), *STK11* (21%; 34% missense, 30% frameshift, 17% nonsense, 15% splice site mutation, 4% non-frameshift indel) and *SMARCA4* (9%) (Fig. 2A). Nine out of twenty adenocarcinoma patients with a *SMARCA4* mutation also exhibited an activating *KRAS* mutation. An analysis of known hotspot mutations in *KRAS* and *NRAS* (codon 12, 13, 61, 146), *EGFR* (exon 18–21), *ERBB2* (exon 20 insertions), *MET* (exon 14 skipping), *BRAF* (codon 466, 469, 581, 594, 596, 600, 601) and *PIK3CA* (exon 9 and 20) is summarized in Fig. 2B. Point mutations in hotspot regions in *KRAS* could be observed in 96 patients (43%). The majority of alterations occurred in codon 12 (n = 77, 80% of the *KRAS* mutated cases), most frequently p.G12C, p.G12 V and p.G12D. Eleven patients exhibited mutations in codon 13 (11%) whereas mutations in codon 61 and 146 where present in seven (7%) and one (1%) patients respectively. In *EGFR*, hotspot alterations were seen in 30 cases (13%), most frequently exon 19 deletions (n = 15, 50%) and p.L858R (n = 11, 37%). Exon 20 insertions in *ERBB2* were seen in six patients (3%), predominantly p.E770_A771InsAYVM (n = 5). In other driver genes activating mutations were detected at low frequencies; *BRAF* (n = 5, 2%, p.V600E, p.K601E, p.G469 A, p.D594 G, p.G596R), *NRAS* (n = 1, < 1%, p.Q61 L), *MET* (n = 2, 1%, exon 14 skipping) and *PIK3CA* (n = 1, < 1%, p.E545 A). The occurrence of *ALK* rearrangements (n = 5, 2%) was assessed in a previous study [25]. In squamous cell carcinoma, frequently mutated genes were *MLL2* (26%), *PIK3CA* (20%) and *CDKN2 A* (15%), whereas mutations in *DDR2* was observed in 4% of patients (Fig. 2C).

3.4. Prognosis linked to mutation status

Univariate analysis of overall survival was performed to evaluate the prognostic effect of genes with mutations occurring in ten or more cases in the adenocarcinoma or squamous cell carcinoma subgroups. Genes with a significant impact of mutation status on overall survival are shown in Fig. 3, including *KRAS* and *EGFR* for reference. Descriptive statistics for these genes with regard to clinical baseline parameters are displayed in Supplementary table S4 (adenocarcinoma) and S5 (squamous cell carcinoma). In adenocarcinoma, no prognostic effect of mutation status in *KRAS* and *EGFR* could be demonstrated. Worse overall survival, or a trend for worse overall survival, were seen in patients with mutations in *TP53* (p = 0.003), *STK11* (p = 0.086) and *SMARCA4* (p = 0.003), whereas patients with mutant *APC* had a better overall survival (p = 0.027) (Fig. 3). In squamous cell carcinoma a worse overall survival was observed for patients with *MLL2* mutations (p = 0.015) whereas patients with mutations in *CSMD3* had a better prognosis (p = 0.034). An association to smoking was observed in adenocarcinoma patients with mutations in *KRAS* (p = 0.0002), *TP53* (p = 0.018), *STK11* (p = 0.021) and *SMARCA4* (p = 0.05). Conversely, *EGFR* mutations were more commonly found in never-smokers (p = 1.53e-06) (Supplementary table S4). In addition, female sex was associated with wild-type TP53 status (p = 0.043). No additional associations to clinical parameters reached significance, but this should be interpreted with caution due to small size of the subgroups in this analysis.

In multivariate analysis (Supplementary table S6), following adjustment for gender, age, stage, smoking history and performance

A

AKT1	CDKN2A	EPHA4	ERBB4	KRAS	NOTCH2	RUNX1T1
ALK	CREBBP	EPHA5	FAM123B	LRP1B	NRAS	SETD2
APC	CSMD3	EPHA6	FBXW7	LTK	PDGFRA	SMAD4
ARID1A	CTNNB1	EPHA7	FGFR1	MAP2K1	PDYN	SMARCA4
ASCL1	CUL3	EPHA8	FGFR2	MET	PIK3CA	SOX2
ASCL4	DDR2	EPHB1	FGFR3	MGA	PTEN	STK11
ATM	DOK2	EPHB2	FGFR4	MLL2	PTPRD	TP53
ATRX	EGFR	EPHB3	FOXP1	MLL3	RB1	TRRAP
BAI3	EPHA1	EPHB4	GRM8	MUC16	RBM10	TSC1
BRAF	EPHA10	EPHB6	HRAS	NF1	RET	U2AF1
CBL	EPHA2	ERBB2	KDR	NFE2L2	RIT1	
CCND1	EPHA3	ERBB3	KEAP1	NOTCH1	ROS1	

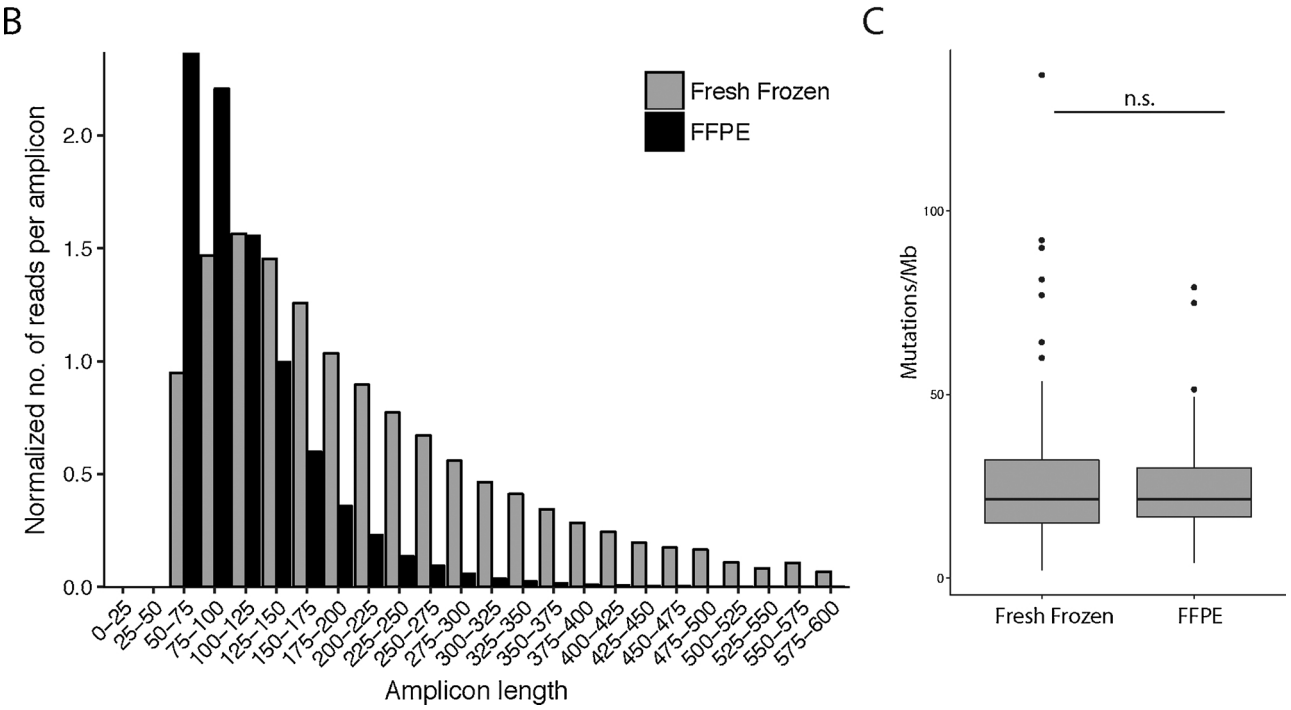

Fig. 1. (A) The 82 genes covered in the Haloplex gene panel used for this study. (B) The number of reads obtained per amplicon length bin was normalized to the number of expected reads per bin and is shown as the mean for the fresh frozen and formalin fixed paraffin embedded (FFPE) samples respectively. (C) The mutation load in fresh frozen and FFPE samples.

status, the differences in all-cause mortality remained significant between the mutated and wild-type group in adenocarcinomas for *TP53* (HR = 1.47, 95% CI 1.02–2.13), *STK11* (HR = 1.66, 95% CI 1.05–2.61) and *SMARCA4* (HR = 2.1, 95% CI 1.22–3.61) and for *CSMD3* (HR = 0.3, 95% CI 0.14-0.62) in squamous cell carcinoma. However, independent prognostic impact of *APC* mutations (HR = 0.5, 95% CI 0.18–1.37) in adenocarcinoma and *MLL2* (HR = 1.81, 95% CI 0.96–3.41) in squamous cell carcinoma could not be confirmed. For completion, even though this study was not powered to assess the effect of adjuvant therapy in small genetic subgroups, and aware of risks with statistical "overfitting", a second multivariate Cox regression analysis (Supplementary table S7) was conducted with addition of this parameter. The difference in all-cause mortality dependent on mutation status remained significant for *STK11* (HR = 1.8, 95% CI 1.11–2.93) and *SMARCA4* (HR = 2.06, 95% CI 1.14–3.72) in adenocarcinoma. However, for *TP53* (HR = 1.25, 95% CI 0.85–1,84) the survival impact of mutations as an independent variable did not remain significant. In squamous cell carcinoma, *CSMD3* mutations (HR = 0.17, 95% CI 0.07-0.4) retained a significant independent positive prognostic impact.

3.5. Survival impact of co-occurring mutations in KRAS, TP53 and STK11

KRAS mutations in lung adenocarcinoma patients are frequently occurring together with mutations in *TP53* and/or *STK11* (Fig. 4A). Of the 195 adenocarcinoma cases without an activating *EGFR* mutation, 81 were mutated in one of the three genes (*KRAS* n = 33, *TP53* n = 45, *STK11* n = 3), while 38 had a wild-type genotype in all three genes. Co-occurring mutations in two of the three genes were present in 73 patients ($KRAS^{mut}$/$TP53^{mut}$ n = 34, $KRAS^{mut}$/$STK11^{mut}$ n = 26, $TP53^{mut}$/$STK11^{mut}$ n = 13), whereas three patients had mutations in all three genes. The best overall survival was observed in the wild-type and $KRAS^{mut}$ groups (Fig.4B). However, there was no significant survival difference between these two groups (p = 0.91). Patients with an isolated *TP53* mutation displayed a worse prognosis in comparison to both

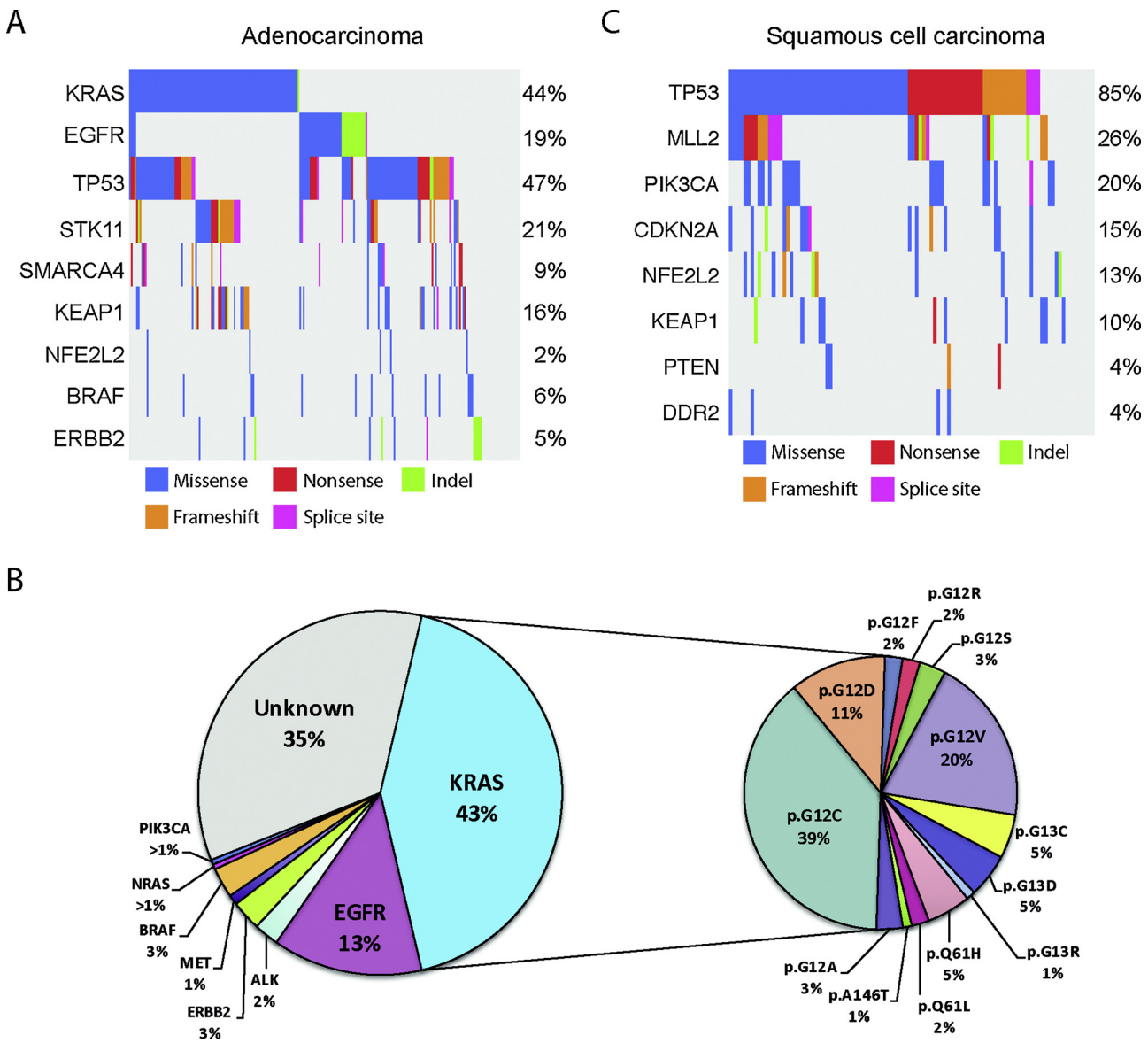

Fig. 2. (A) Mutation distribution and frequency in adenocarcinoma patients of selected genes. Color-coded after type of mutation: blue (missense), red (nonsense), green (non-frameshift indel), orange (frameshift), purple (splice site). (B) Frequency of hotspot mutations in commonly altered driver genes in adenocarcinoma patients. (C) Mutation distribution and frequency in squamous cell carcinoma patients of selected genes.

the wild-type and KRAS^{mut} group (p = 0.019; p = 0.014 respectively). For cases with co-occurring mutations a significantly worse outcome could be seen for the KRAS^{mut}/TP53^{mut} and KRAS^{mut}/STK11^{mut} groups in comparison to the wild-type group (p = 0.038; p = 0.046 respectively). Inferior survival was also observed in comparison to the KRAS^{mut} group lacking a co-mutation (p = 0.032; p = 0.04 respectively). To evaluate these findings in an independent cohort, mutation status and survival data from the MSK-IMPACT study [26] were retrieved from the cBioPortal [27,28] (Fig. 4C). A total of 576 lung adenocarcinoma patients (stage I–IV) were selected after removal of cases with an activating *EGFR* mutation. As in our cohort, the patients in the wild-type group and KRAS^{mut} group exhibited a better prognosis, without a significant difference between each other (p = 0.24). On the other hand, patients with an isolated mutation in either *TP53* or *STK11* had a worse outcome in comparison to the wild-type group (p = 0.0016; p = 0.0023 respectively). The groups KRAS^{mut}/TP53^{mut} and KRAS^{mut}/STK11^{mut} had a worse overall survival in comparison to patients with a wild-type phenotype (p = 0.006; p = 1.17e-07 respectively), whereas only patients in the KRAS^{mut}/STK11^{mut} group had a significant worse prognosis in comparison to the KRAS^{mut} group (p = 0.0001).

4. Discussion

We have performed targeted resequencing of a consecutive population-based Swedish NSCLC cohort, assessing the mutation patterns in 82 genes commonly altered in NSCLC and evaluated the prognostic impact in different mutation subgroups. In order to include all samples in the cohort we used FFPE tissue for the patients lacking fresh frozen biobank material. The sequencing approach was adapted previously by members of our group to function well on degraded FFPE material by use of two-strand capture a reduced target fragment lengh [22]. In addition, the FFPE samples were sequenced at a greater depth and the bioinformatics pipeline was optimized to avoid technical biases between the two sample types.

At large, we observed mutation frequencies in our cohort comparable to other western populations, with an exception of a high frequency of activating *KRAS* mutations (43%) among lung

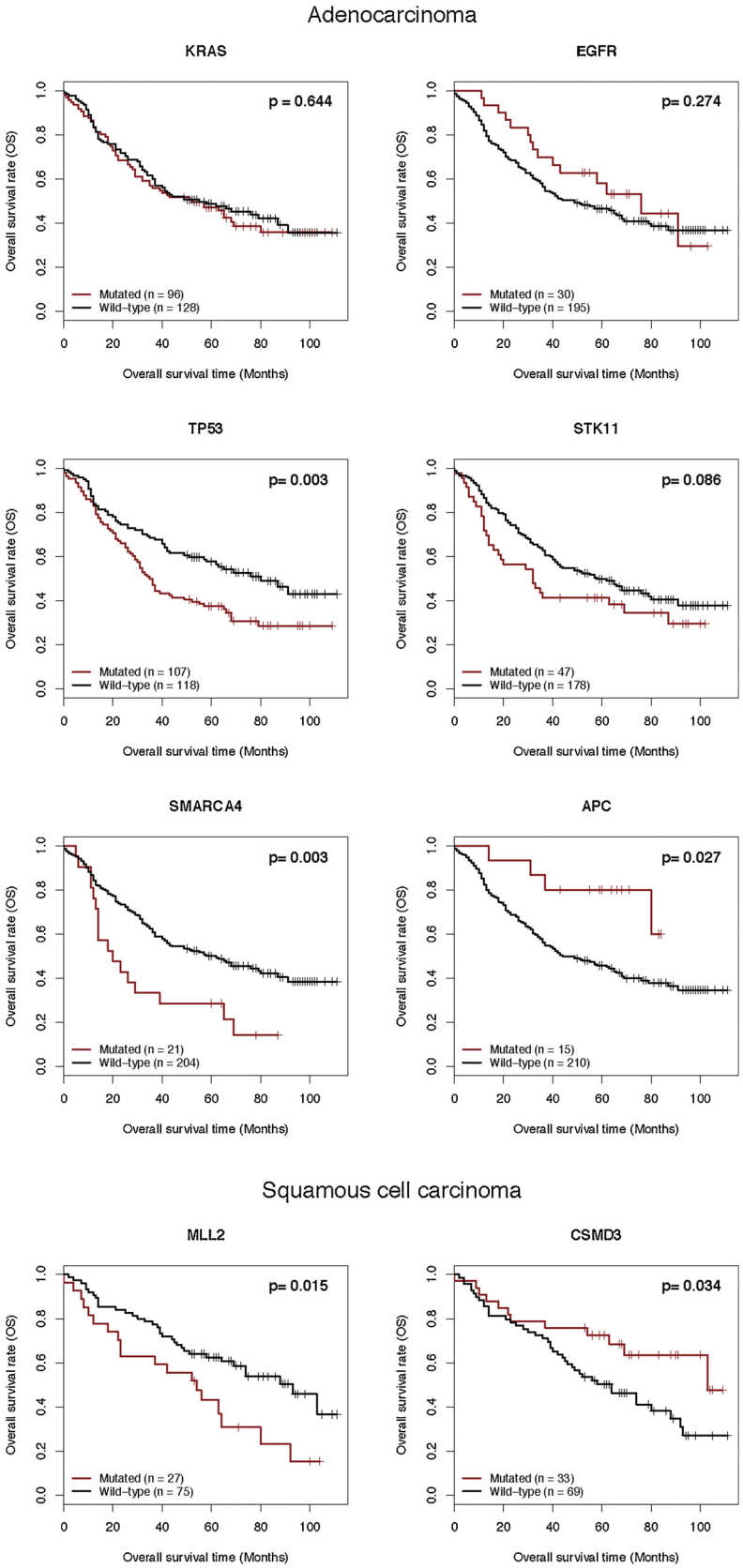

Fig. 3. Kaplan-Meier analysis of overall survival stratified by mutation status of the genes *KRAS*, *EGFR*, *TP53*, *STK11*, *SMARCA4* and *APC* for adenocarcinoma patients and *MLL2* and *CSMD3* for squamous cell carcinoma patients; comparing mutated to wild-type.

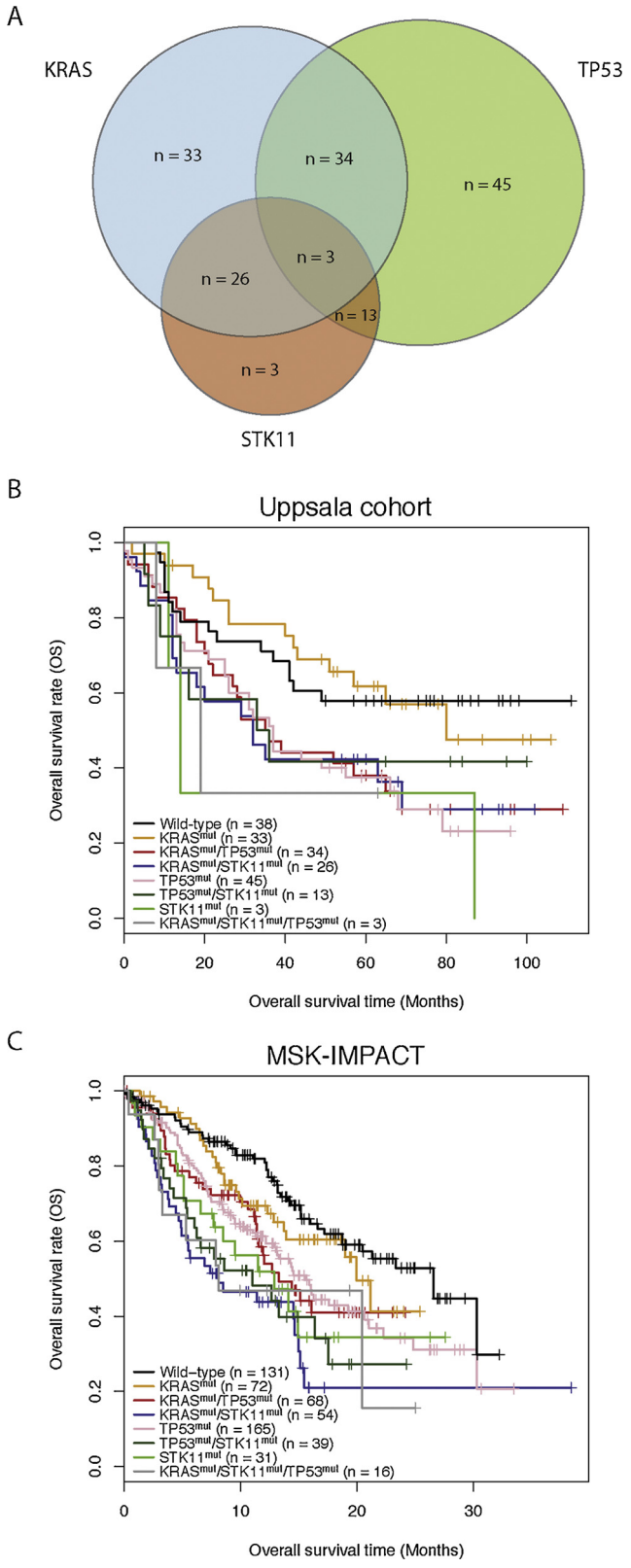

Fig. 4. (A) Venn-diagram showing the overlap of mutations in the genes *KRAS*, *TP53* and *STK11* for the 195 *EGFR* negative adenocarcinomas included in this study. (B) Kaplan-Meier analysis of overall survival stratified by the mutation status in the three genes *KRAS*, *TP53* and *STK11* for the Uppsala NSCLC cohort. (C) Kaplan-Meier analysis of overall survival of 576 adenocarcinomas (stage I-IV) from the MSK-IMPACT study, stratified by the mutation status in the three genes *KRAS*, *TP53* and *STK11*. The data from the MSK-IMPACT study were retrieved from the cBioportal and only lung adenocarcinoma patients with available survival data and complete mutation data in the genes *KRAS*, *TP53*, *STK11* and *EGFR* were included. Patients with an activating *EGFR* mutation were excluded from the analysis.

adenocarcinoma patients. The Cancer Genome Atlas and Clinical Lung Cancer Genome Project both report mutated *KRAS* in approximately 30% of lung adenocarcinomas [5,29]. In more homogenous populations, a Dutch study describe *KRAS* mutations in 37% of the adenocarcinomas [30], and in a previous Swedish study on NSCLC of mixed histologic subtypes the *KRAS* mutation rate in codon 12 and 13 was 29% [31]. As *KRAS* mutations are linked to smoking, we believe that the high rate of *KRAS* mutations in our cohort is explained by a high fraction of smokers (89%). Another explanation might be the high sensitivity of our assay, in combination with the ability to assess hotspot regions outside codon 12/13. Indeed, mutations in codons 61 and 146 accounted for 8% of the detected *KRAS* mutations.

The mutational status of a gene may have prognostic implications. In our study, after adjustment for confounding clinical parameters, we found that adenocarcinoma patients harboring mutations in *TP53*, *STK11* or *SMARCA4* had a worse overall survival. Some studies of NSCLC of mixed histologic subtypes do not show any prognostic impact of mutated *TP53* [32,33]. Other reports, including ours, show worse survival in lung adenocarcinoma patients with mutations in *TP53* [34,35]. However, separating the true prognostic and predictive impact of TP53, in relation to adjuvant therapy, has proven to be difficult [36]. Indeed, we could not confirm a significant prognostic impact of TP53 after adjustment for adjuvant therapy in multivariate analysis. The finding that mutations in *STK11* conferred poor prognosis is supported by a trend for a worse survival in one study [37], while another paper showed poor prognosis for patients with mutations in exon 1 and 2 of *STK11* in comparison to mutations in exon 3-9 [38]. *SMARCA4*, a subunit in the SWI/SNF chromatin remodeling complex, is frequently mutated in lung adenocarcinoma. We observed a worse prognosis for patients harboring a *SMARCA4* mutation. So far, to our knowledge, the mutational status of *SMARCA4* and connection to prognosis have not been evaluated in NSCLC. Interestingly, there is data linking low gene expression of *SMARCA4* to poor survival and results from *in-vivo* models showing good response to an OXPHOS inhibitor in *SMARCA4* deficient xenografts [39,40]. In squamous cell carcinoma, mutations in *MLL2*, a histone methyltransferase, was linked to worse overall survival, in concordance to previous reports [41]. Little is known about the function of *CSMD3* (CUB and Sushi multiple domains 3), a predictive membrane protein with a role in adhesion. One report links mutations in this gene to better survival in patients with esophageal squamous cell carcinoma [42]. To our knowledge, the mutational status of CSMD3 and its impact on prognosis have not been evaluated in NSCLC before. Notably, in our study it represents the gene with the most profound independent impact on prognosis (HR 0.17 and 0.3, with and without adjustment for adjuvant therapy in multivariate analysis). We could not detect an impact on overall survival for any other gene with recurrent mutations in our gene panel. A limitation of this exploratory study is that adjustment for multiple testing was not performed, creating a risk for false positive results. We here interpret the results based on known functional roles of genes and mutations in addition to supportive evidence from other studies. However, further studies are needed to confirm the described clinical impact of the respective mutations.

A large fraction of lung adenocarcinomas in western populations harbor an activating mutation in *KRAS* [43]. These *KRAS* mutations are frequently accompanied by mutations in the tumor suppressor genes *TP53* and *STK11*. We observed that patients with an isolated *KRAS* mutation had an overall survival similar to the wild-type group, whereas patients with co-occurring mutations in either *TP53* or *STK11* had a worse overall survival in comparison to the wild-type group and

the *KRAS* only group. Our results were further supported by data from the MSK-IMPACT study. Other studies looking at the prognostic impact of co-occurring mutations in *TP53* or *STK11* together with *KRAS* have been inconclusive [37,44]. Gene expression analysis in lung adenocarcinoma has shown that *TP53* mutations enhance signatures related to cell proliferation in contrast to *STK11* mutations that boost expression of genes induced by *KRAS* mutations and suppress signatures related to immune function [45]. Thus, from a functional perspective the negative prognostic impact of co-mutations in *TP53* or *STK11* might be explained by proliferative drive and poor immune function respectively in the two subgroups. At the same time, from the perspective of immune therapy, *TP53* and *STK11* appear to have completely opposite impact. Tumors with *TP53* mutations, especially with co-occurring *KRAS* mutations, generally demonstrate dense infiltration of CD8+ T-cells in combination with high expression of PD-L1, and tend to respond well to checkpoint inhibitor therapy [46,47]. Conversely *STK11* mutations were linked to decreased immune cell infiltration, low PD-L1 expression, and poor response to PD-L1 inhibition. Indeed, recent studies on larger patient cohorts confirm STK11 mutations as a prominent predictor of resistance to checkpoint inhibitor therapy, partly independent of PD-L1 expression and tumor mutation burden [48,49]. In view of the described complex prognostic and predictive effects we believe that it might be relevant to include analysis of *TP53* and *STK11* in molecular testing algorithms. Interestingly, these genes are covered in most established small targeted NGS panels routinely used in reflex testing algorithms for driver mutations in NSCLC.

5. Conclusion

We have analyzed mutations in 82 genes associated with lung cancer development in a Swedish NSCLC cohort. Bias due to sample selection, FFPE versus fresh-frozen tissue, was avoided by technical adaptation of the targeted sequencing method. At large the mutation spectra resemble findings in other western lung cancer populations, with an exception of a high *KRAS* mutation rate in adenocarcinoma. A poor prognosis could be linked to mutations in *TP53*, *STK11 or SMARCA4* in adenocarcinoma and *MLL2* mutations in squamous cell carcinoma, while mutations in *CSMD3* were linked to a better prognosis in squamous cell carcinoma. Interestingly, we show that *TP53* and *STK11* mutations confer poor survival in the *KRAS*-positive subgroup, but not *KRAS* mutations themselves as an isolated aberration, suggesting that co-mutations beside commonly tested driver mutations in the MAP-kinase pathway are important determinants of aggressive behavior after surgery of lung cancer, as well as strong predictors of immune therapy response.

Acknowledgements

This study was supported by research grants from the Swedish Cancer Society (2013/711, 2016/827), the Lions Cancer Foundation (Uppsala, Sweden) and the County council of Uppsala (Uppsala University Hospital, Sweden).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lungcan.2019.01.003.

References

[1] D.S. Ettinger, D.E. Wood, W. Akerley, L.A. Bazhenova, H. Borghaei, D.R. Camidge, et al., National comprehensive cancer n. Non-small cell lung Cancer, version 6.2015, J. Compr. Canc. Netw. 13 (2015) 515–524.
[2] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, et al., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int. J. Cancer 136 (2015) E359–386.
[3] Z. Chen, C.M. Fillmore, P.S. Hammerman, C.F. Kim, K.K. Wong, Non-small-cell lung cancers: a heterogeneous set of diseases, Nat. Rev. Cancer 14 (2014) 535–546.
[4] L.B. Alexandrov, S. Nik-Zainal, D.C. Wedge, S.A. Aparicio, S. Behjati, Signatures of mutational processes in human cancer, Nature 500 (2013) 415–421.
[5] Cancer Genome Atlas Research Network, Comprehensive molecular profiling of lung adenocarcinoma, Nature 511 (2014) 543–550.
[6] Cancer Genome Atlas Research Network, Comprehensive genomic characterization of squamous cell lung cancers, Nature 489 (2012) 519–525.
[7] J.G. Paez, P.A. Janne, J.C. Lee, S. Tracy, H. Greulich, S. Gabriel, et al., EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy, Science 304 (2004) 1497–1500.
[8] E.L. Kwak, Y.J. Bang, D.R. Camidge, A.T. Shaw, B. Solomon, R.G. Maki, et al., Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer, N. Engl. J. Med. 363 (2010) 1693–1703.
[9] K. Bergethon, A.T. Shaw, S.H. Ou, R. Katayama, C.M. Lovly, N.T. McDonald, et al., ROS1 rearrangements define a unique molecular class of lung cancers, J. Clin. Oncol. 30 (2012) 863–870.
[10] A. Drilon, L. Wang, A. Hasanovic, Y. Suehara, D. Lipson, P. Stephens, et al., Response to Cabozantinib in patients with RET fusion-positive lung adenocarcinomas, Cancer Discov. 3 (2013) 630–635.
[11] D. Planchard, B. Besse, H.J. Groen, P.J. Souquet, E. Quoix, C.S. Baik, et al., Dabrafenib plus trametinib in patients with previously treated BRAF(V600E)-mutant metastatic non-small cell lung cancer: an open-label, multicentre phase 2 trial, Lancet Oncol. 17 (2016) 984–993.
[12] M.G. Kris, D.R. Camidge, G. Giaccone, T. Hida, B.T. Li, J. O'Connell, et al., Targeting HER2 aberrations as actionable drivers in lung cancers: phase II trial of the pan-HER tyrosine kinase inhibitor dacomitinib in patients with HER2-mutant or amplified tumors, Ann. Oncol. 26 (2015) 1421–1427.
[13] J. Mazieres, S. Peters, B. Lepage, A.B. Cortot, F. Barlesi, M. Beau-Faller, et al., Lung cancer that harbors an HER2 mutation: epidemiologic characteristics and therapeutic perspectives, J. Clin. Oncol. 31 (2013) 1997–2003.
[14] G.M. Frampton, S.M. Ali, M. Rosenzweig, J. Chmielecki, X. Lu, T.M. Bauer, et al., Activation of MET via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors, Cancer Discov. 5 (2015) 850–859.
[15] P.K. Paik, A. Drilon, P.D. Fan, H. Yu, N. Rekhtman, M.S. Ginsberg, et al., Response to MET inhibitors in patients with stage IV lung adenocarcinomas harboring MET mutations causing exon 14 skipping, Cancer Discov. 5 (2015) 842–849.
[16] P.S. Hammerman, M.L. Sos, A.H. Ramos, C. Xu, A. Dutt, W. Zhou, et al., Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer, Cancer Discov. 1 (2011) 78–89.
[17] S.H. Lim, J.M. Sun, Y.L. Choi, H.R. Kim, S. Ahn, J.Y. Lee, et al., Efficacy and safety of dovitinib in pretreated patients with advanced squamous non-small cell lung cancer with FGFR1 amplification: a single-arm, phase 2 study, Cancer 122 (2016) 3024–3031.
[18] C.E. Greer, S.L. Peterson, N.B. Kiviat, M.M. Manos, PCR amplification from paraffin-embedded tissues. Effects of fixative and fixation time, Am. J. Clin. Pathol. 95 (1991) 117–124.
[19] C. Williams, F. Ponten, C. Moberg, P. Soderkvist, M. Uhlen, J. Ponten, et al., A high frequency of sequence alterations is due to formalin fixation of archival specimens, Am. J. Pathol. 155 (1999) 1467–1471.
[20] P. Micke, J.S. Mattsson, D. Djureinovic, B. Nodin, K. Jirstrom, L. Tran, et al., The impact of the fourth edition of the WHO classification of lung tumours on histological classification of resected pulmonary NSCCs, J. Thorac. Oncol. 11 (2016) 862–872.
[21] Tumours of the lung, in: W.D. Travis, BE, A.P. Burke, A. Marx, A.G. Nicholson (Eds.), WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart, 4th (ed), IARC Press, Lyon, France, 2015, pp. 9–151.
[22] L.N. Moens, E. Falk-Sorqvist, V. Ljungstrom, J. Mattsson, M. Sundstrom, L. La Fleur, et al., HaloPlex targeted resequencing for mutation detection in clinical formalin-fixed, paraffin-embedded tumor samples, J. Mol. Diagn. 17 (2015) 729–739.
[23] T. Tikkanen, B. Leroy, J.L. Fournier, R.A. Risques, J. Malcikova, T. Soussi, Seshat: a Web service for accurate annotation, validation, and analysis of TP53 variants generated by conventional and next-generation sequencing, Hum. Mutat. 39 (2018) 925–933.
[24] R. Frank, M. Scheffler, S. Merkelbach-Bruse, M.A. Ihle, A. Kron, M. Rauer, et al., Clinical and pathological characteristics of KEAP1- and NFE2L2-mutated non-small cell lung carcinoma (NSCLC), Clin. Cancer Res. 24 (2018) 3087–3096.
[25] J.S. Mattsson, H. Brunnstrom, V. Jabs, K. Edlund, K. Jirstrom, S. Mindus, et al., Inconsistent results in the analysis of ALK rearrangements in non-small cell lung cancer, BMC Cancer 16 (2016) 603.
[26] A. Zehir, R. Benayed, R.H. Shah, A. Syed, S. Middha, H.R. Kim, et al., Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients, Nat. Med. 23 (2017) 703–713.
[27] E. Cerami, J. Gao, U. Dogrusoz, B.E. Gross, S.O. Sumer, B.A. Aksoy, et al., The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data, Cancer Discov. 2 (2012) 401–404.
[28] J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, et al., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, Sci. Signal. 6 (2013) l1.
[29] Clinical Lung Cancer Genome Project, Network Genomic Medicine, A genomics-based classification of human lung tumors, Sci. Transl. Med. 5 (2013) 209ra153.
[30] A.J. Smits, J.A. Kummer, J.W. Hinrichs, G.J. Herder, K.C. Scheidel-Jacobse, N.M. Jiwa, 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. Dordr. (Dordr) 35 (2012) 189–196.
[31] A. Hallqvist, F. Enlund, C. Andersson, H. Sjogren, A. Hussein, E. Holmberg, et al.,

Mutated KRAS is an independent negative prognostic factor for survival in NSCLC stage III disease treated with high-dose radiotherapy, Lung Cancer Int. 2012 (2012) 587424.
[32] M.S. Tsao, S. Aviel-Ronen, K. Ding, D. Lau, N. Liu, A. Sakurada, et al., Prognostic and predictive importance of p53 and RAS for adjuvant chemotherapy in non small-cell lung cancer, J. Clin. Oncol. 25 (2007) 5240–5247.
[33] X. Ma, V. Rousseau, H. Sun, S. Lantuejoul, M. Filipits, R. Pirker, et al., GSignificance of TP53 mutations as predictive markers of adjuvant cisplatin-based chemotherapy in completely resected non-small-cell lung cancer, Mol. Oncol. 8 (2014) 555–564.
[34] 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.
[35] C. Deben, J. Van den Bossche, N. Van Der Steen, F. Lardon, A. Wouters, K.O. de Beeck, et al., Deep sequencing of the TP53 gene reveals a potential risk allele for non-small cell lung cancer and supports the negative prognostic value of TP53 variants, Tumour Biol. 39 (2017) 1010428317694327.
[36] X. Ma, G. Le Teuff, B. Lacas, M.S. Tsao, S. Graziano, J.P. Pignon, et al., Prognostic and predictive effect of TP53 mutations in patients with non-small cell lung Cancer from adjuvant cisplatin-based therapy randomized trials: a LACE-Bio pooled analysis, J. Thorac. Oncol. 11 (2016) 850–861.
[37] F. Facchinetti, M.V. Bluthgen, G. Tergemina-Clain, L. Faivre, J.P. Pignon, D. Planchard, et al., LKB1/STK11 mutations in non-small cell lung cancer patients: descriptive analysis and prognostic value, Lung Cancer 112 (2017) 62–68.
[38] N. Pecuchet, P. Laurent-Puig, A. Mansuet-Lupo, A. Legras, M. Alifano, K. Pallier, et al., Different prognostic impact of STK11 mutations in non-squamous non-small-cell lung cancer, Oncotarget 8 (2017) 23831–23840.
[39] E.H. Bell, A.R. Chakraborty, X. Mo, Z. Liu, K. Shilo, S. Kirste, et al., SMARCA4/BRG1 is a novel prognostic biomarker predictive of cisplatin-based chemotherapy outcomes in resected non-small cell lung Cancer, Clin. Cancer Res. 22 (2016) 2396–2404.
[40] Y.L. Deribe, Y. Sun, C. Terranova, F. Khan, J. Martinez-Ledesma, J. Gay, et al., Mutations in the SWI/SNF complex induce a targetable dependence on oxidative phosphorylation in lung cancer, Nat. Med. 24 (2018) 1627.
[41] M. Choi, H. Kadara, J. Zhang, E.R. Parra, J. Rodriguez-Canales, S.G. Gaffney, et al., Mutation profiles in early-stage lung squamous cell carcinoma with clinical follow-up and correlation with markers of immune function, Ann. Oncol. 28 (2017) 83–89.
[42] J. Deng, H. Chen, D. Zhou, J. Zhang, Y. Chen, Q. Liu, et al., Comparative genomic analysis of esophageal squamous cell carcinoma between Asian and Caucasian patient populations, Nat. Commun. 8 (2017) 1533.
[43] N. Karachaliou, C. Mayo, C. Costa, I. Magri, A. Gimenez-Capitan, M.A. Molina-Vila, et al., KRAS mutations in lung cancer, Clin. Lung Cancer 14 (2013) 205–214.
[44] K.C. Arbour, E. Jordan, H.R. Kim, J. Dienstag, Sanchez-Vega F. Yu HA, P. Lito, et al., Effects of co-occurring genomic alterations on outcomes in patients with KRAS-mutant non-small cell lung cancer, Clin. Cancer Res. 24 (2018) 334–340.
[45] M.B. Schabath, E.A. Welsh, W.J. Fulp, L. Chen, J.K. Teer, Z.J. Thompson, et al., Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma, Oncogene 35 (2016) 3209–3216.
[46] Z.Y. Dong, W.Z. Zhong, X.C. Zhang, J. Su, Z. Xie, S.Y. Liu, et al., Potential predictive value of TP53 and KRAS mutation status for response to PD-1 blockade immunotherapy in lung adenocarcinoma, Clin. Cancer Res. 23 (2017) 3012–3024.
[47] J. Biton, A. Mansuet-Lupo, N. Pecuchet, M. Alifano, H. Ouakrim, J. Arrondeau, et al., TP53, STK11, and EGFR mutations predict tumor immune profile and the response to Anti-PD-1 in lung adenocarcinoma, Clin. Cancer Res. 24 (2018) 5710–5723.
[48] F. Skoulidis, M.E. Goldberg, D.M. Greenawalt, M.D. Hellmann, M.M. Awad, J.F. Gainor, et al., STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma, Cancer Discov. 8 (2018) 822–835.
[49] H. Rizvi, F. Sanchez-Vega, K. La, W. Chatila, P. Jonsson, D. Halpenny, et al., Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing, J. Clin. Oncol. 36 (2018) 633–641.