Discover Oncology

Research

Comparison of targeted next generation sequencing assays in non-small cell lung cancer patients

Ieva Drejeriene[1,3] · Jurate Gruode[3,6] · Saulius Cicenas[4] · Charalambos Loizides[2] · Alexia Eliades[2] · Achilleas Achilleos[2] · Elena Kypri[2] · Kyriakos Tsangaras[2,7] · Marios Ioannides[2] · George Koumbaris[2] · Diana Stanciute[4] · Arnoldas Krasauskas[1,4] · Philippos C. Patsalis[2,5]

Received: 27 March 2024 / Accepted: 18 September 2024
Published online: 18 December 2024

Abstract

Non-small cell lung cancer (NSCLC) is the most prevalent type of lung cancer the mutational spectrum of which has been extensively characterized. Treatment of patients with NSCLC based on their molecular profile is now part of the standard clinical care. The aim of this study was firstly to investigate two different NGS-based tumor profile genetic tests and secondly to assess the clinical actionability of the mutations and their association with survival and clinicopathological characteristics. Overall, 52 mutations were identified in 31 patients by either one or both assays. The most frequently mutated genes were TP53 (40.4%), KRAS (13.46%) and EGFR (9.62%). TP53 and KRAS mutations were associated with worst overall survival while KRAS was positively correlated with adenocarcinoma. The two methods showed a high concordance for the commonly covered genomic regions (97.14%). Ten mutations were identified in a genomic region exclusively covered by the MEDICOVER Genetics custom tumor profile assay. Likewise, one MET mutation was identified by the Ion Amliseq assay in a genomic region exclusively covered by Ion Amliseq. In conclusion both assays showed highly similar results in the commonly covered genomic areas, however, the MEDICOVER Genetics assay identified additional clinically actionable mutations that can be applied in clinical practice for personalized treatment decision making for patients with NSCLC.

Keyword NSCLC · mutational profile · clinical utility · next generation sequencing

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12672-024-01640-7.

✉ Ieva Drejeriene, ieva.drejeriene@kulig.lt; ✉ Alexia Eliades, alexia.eliades@medicover.com; ✉ Philippos C. Patsalis, philippos.patsalis@medicover.com; Jurate Gruode, jurate.gruode@kulig.lt; Saulius Cicenas, saulius.cicenas@nvi.lt; Charalambos Loizides, charalambos.loizides@medicover.com; Achilleas Achilleos, achilleas.achilleos@medicover.com; Elena Kypri, elena.kypri@medicover.com; Kyriakos Tsangaras, kyriakos.tsangaras@medicover.com; Marios Ioannides, marios.ioannides@medicover.com; George Koumbaris, george.koumbaris@medicover.com; Diana Stanciute, diana.stanciute3@nvi.lt; Arnoldas Krasauskas, arnoldas.krasauskas@nvi.lt | [1]Faculty of Medicine, Vilnius University, Vilnius, Lithuania. [2]MEDICOVER Genetics, Nicosia, Cyprus. [3]Klaipeda University Hospital, Klaipeda, Lithuania. [4]National Cancer Institute, Vilnius, Lithuania. [5]Department of Human Genetics, School of Medicine, University of Nicosia Medical School, Nicosia, Cyprus. [6]Faculty of Medicine, Klaipeda University, Klaipeda, Lithuania. [7]Department of Life and Health Sciences, University of Nicosia, Nicosia, Cyprus.

1 Introduction

Lung cancer is the third leading type of cancer globally with an incidence of 22 cases per 100,000 people annually and has the highest mortality rate [1]. NSCLC accounts for approximately 85% of the lung cancer cases [2]. There is an increasing incidence of lung cancer in women with a more favorable prognosis as compared to NSCLC in men [3]. In addition, female patients with NSCLC have an improved benefit compared to men in regard to their treatment response with EGFR inhibitors versus chemotherapy [4]. Molecular targeted therapies against driver mutations of patients with NSCLC are already improving patients' survival over traditional chemotherapy. Consequently, molecular testing is now applied as part of routine clinical practise [5].

An essential part of the diagnostic procedure in guiding the appropriate treatment for NSCLC is the molecular characterization of the tumor. A variety of techniques are employed to detect molecular alterations including protein-based methods (immunohistochemistry), fluorescence *in situ* hybridization (FISH) and chromogenic *in situ* hybridization (CISH) [6]. Most of these methods predominantly depend on a qualitative assessment and can therefore represent a challenge to standarise [6]. NGS testing has been increasingly applied to clinical practice in recent years and is now recommended by professional guidelines for NSCLC molecular profiling [5, 7]. NGS is a sensitive and sufficiently quick method which can simultaneously identify a large number of driver mutations in oncogenes that are associated with targeted therapy and acquired drug-resistance and has been increasingly applied to clinical practice in recent years [8–10]. Patients with EGFR sensitizing mutations in the tumor can now receive EGFR tyrosine kinase inhibitor (TKI) treatment as a first line therapy [5]. Following TKI treatment, if EGFR resistant mutations are identified (e.g EGFT T790M), then osimertinib treatment is recommended [11, 12]. Likewise patients with BRAF V600E mutated tumors can receive dabrafenib, a BRAF inhibitor and trametinib, a MEK inhibitor as first line therapy [5, 13]. MET exon 14 skipping mutated tumors show sensitivity to capmatinib, a MET inhibitor [5, 14]. Moreover, several genes are now emerging as potential biomarkers to identify novel targeted therapies for patients with NSCLC and serve as inclusion criteria in clinical trials [15]. For example, patients with ERBB2 (HER2) exon 20 mutations show clinical response to pan-HER2 blocking drugs [16]. Likewise, activating mutations in the JAK2 gene are shown to confer sensitivity to both JAK2 inhibitors and anti-PD1 immunotherapy in patients with NSCLC [17].

Multi-gene NGS assays have also enabled the simultaneous analysis of multiple genes from a single tumor. It is now clear that a subset to NSCLC patients carry co-existing driver mutations that could explain the heterogeneity in clinical outcomes upon targeted treatment [18].

The biggest challenge for molecular testing is tissue availability of the patient's tumor. In many cases the amount of formalin fixed paraffin embedded (FFPE) tumor tissue remaining after pathologists' analysis is limited. The formalin fixation and paraffin -embedding process reduces DNA quality via fragmentation, cross-linking and chemical modifications that introduce DNA damage [19]. These limitations highlight the importance of selecting the appropriate molecular test that can overcome these challenges by introducing quality control checkpoints to ensure high quality molecular data. In addition, the selection of the most appropriate test is of outmost importance to ensure robust diagnostic power by covering the highest number of clinically actionable mutations.

In this study the main objective was the comparison of two different NGS-based tests to investigate their sensitivity and clinical utility in identifying clinically actionable mutations in female patients diagnosed with NSCLC. To this end, we analyzed FFPE tissue samples using two different NGS-based assays: i) Ion AmpliSeq Colon and Lung Cancer Research V2 Panel (Ion Torrent PGM platform), an amplicon-based assay that covers hotspot regions in 22 genes associated with lung and colorectal cancer and ii) MEDICOVER Genetics custom- tumor profile assay, that relies on hybrid capture technology and covers hotspot regions and selected targeted regions in 49 genes primarily associated with NSCLC. Sequencing data were compared between the two methods for the commonly covered genomic regions. Moreover, mutations identified in genomic areas exclusively covered by each method are also reported. The diagnostic yield and clinical utility of each assay as well as correlation of molecular findings with clinicopathological parameters is discussed.

2 Results

2.1 Experimental strategy and patient characteristics

Sections were cut from the FFPE block of the primary tumor tissue biopsy for all 51 patients and were subjected to DNA extraction. For 39 patients, the same DNA sample was analyzed by both assays. For 12 patients, different DNA samples

(extracted from different sections of the same FFPE block) were analyzed by only one assay, either the Ion Ampliseq Colon Lung v2 assay or the MEDICOVER Genetics custom tumor profile assay (Fig. 1A, B). Out of 51 patients, 39 patients had samples that met quality control criteria and were subjected to targeted sequencing with the two different assays (Fig. 1C). A total of 12 DNA samples were excluded from analysis by both assays due to low or poor DNA quality. The average age of patients was 60.8 years (±9.2). Out of 39 patients, 38.5% were smokers. Most of the patients were diagnosed with adenocarcinoma (74.3%) and 25.7% of them with squamous cell carcinoma. Five patients (12.8%) were diagnosed with stage IV disease, 35.9% of the patients were diagnosed with stage III, 25.6% with stage II and 25.7% with stage I. Out of 29 adenocarcinoma patients 18 (62%) and out of 10 squamous cell carcinoma 8 (80%) patients had either lymph node or distant metastasis (Table 1).

2.2 Concordance analysis

Overall 31 patients (79.5%) had mutations identified in their tumor specimens by either assay, 27 of which had mutations identified by both assays (69.23%) (Fig. 2A, B). Concordance analysis was performed on sequencing data generated by both assays as described above for the genomic region commonly covered by both methods using the same DNA samples (originating from the same FFPE sections). This commonly covered region consists of hotspot regions of 18 genes with a total genomic size of 8.6 Kb (Fig. 1A, B). Thirty-four variants in 25 patients were identified by both methods while one extra variant -the KRAS G13C at 5.9% VAF- was identified only by the MEDICOVER tumor profile assay in patient 1 (Table 2). The IonAmpliseq assay failed to detect this variant above the minimum acceptable threshold of 5%, however, it was detected at VAF=3.78%. Hence, concordance between the two assays was estimated at 97.14% (Fig. 2C). The frequencies of the concordant mutations in both assays were highly similar (r^2=0.9156, Fig. 2C).

Data are compared from samples analyzed by both assays and originating from the same DNA sample. NF=not found above VAF≥5% threshold.

2.3 Intra-tumor variability

The processing of DNA samples originating from different sections at different layers of the same patients' FFPE tissue biopsy block enabled the investigation of intra-tumor heterogeneity. As shown in Fig. 2D, two commonly covered variants, TP53 V227G and BRAF V600E were not identified in all sections tested from the same tumor (Figure 2D, Table S1). Given that the two methods have shown performance similarities, this variability does not necessarily reflect differences in the sensitivity of each method but instead could be a result of the heterogeneity of the tissue biopsy samples with different spatial origin within the primary tumor. However, to conclusively characterize the extend of intra-tumor heterogeneity and investigate its clinical importance a large-scale study is required.

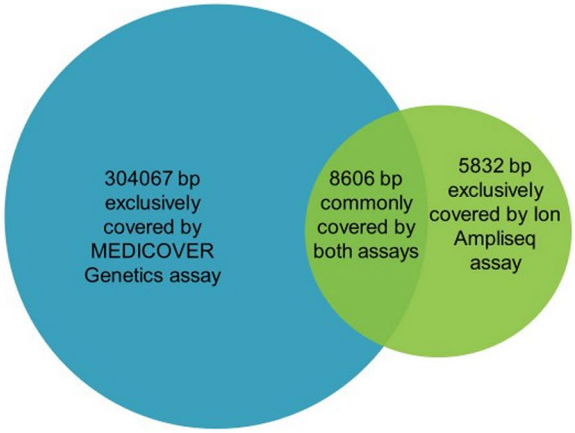
2.4 Assessment of the molecular profile for NSCLC patients

Data from all the regions covered from the two assays combined (including overlapping and exclusively covered regions) were used to assess the mutation profile of these patients. A total of 52 mutations were identified in 31 patients. Sixteen patients (51.6%) had just 1 mutation identified in their tumor biopsy (Figure 3A). The total number of mutations did not show significant correlation with tumor content in the FFPE specimens (Supplemental figure 1A). However, there's a statistically significant increase allele frequency in FFPE samples with tumor content >70% as compared with FFPE samples with tumor content ≤70% (Supplemental Figure 1b). The most frequent mutations were identified in TP53 (21 patients, 40.4%), KRAS (7 patients, 13.46%), EGFR (5 patients, 9.62%) and PIK3CA (4 patients, 7.7%) (Fig. 3B). Mutations in TP53 were predominantly identified in the DNA binding domain (amino acids 98–292) [20], a region known to harbor the majority of deleterious mutations in this gene (Fig. 4A). KRAS mutations were predominantly localized on codon 12, a widely studied recurrent region in multiple cancer types while 1 mutation was identified in codon 13 (Fig. 4B). All KRAS mutations identified were associated with reduced sensitivity to TKI inhibitors [21]. EGFR mutations were identified in the tyrosine kinase domain (Fig. 4C); four in-frame exon19 deletions and one exon 21 L858R substitution, all known to confer sensitivity to EGFR inhibitors such as gefitinib, afatinib and erlotinib [22].

A.

Ampliseq panel (22 genes)		MEDICOVER Genetics panel (49 genes)			
AKT1	KRAS	AKT1	ERBB2	HOXB13	NPM1
ALK	MAP2K1	ALK	ERBB3	IDH1	NRAS
BRAF	MET	APC	ERBB4	IDH2	PDGFRA
CTNNB1	NOTCH1	AR	ESR1	JAK2	PIK3CA
DDR2	NRAS	ARAF	FBXW7	KEAP1	PIK3CB
EGFR	PIK3CA	ATM	FLT3	KIT	PTEN
ERBB2	PTEN	BRAF	FOXA1	KRAS	RAF1
ERBB4	SMAD4	BRCA1	FOXL2	MAP2K1	RET
FBXW7	STK11	BRCA2	GATA3	MAP3K1	RUNX1
FGFR1	TP53	CTNNB1	GNA11	MET	SMAD4
FGFR2		DDR2	GNAQ	MPL	SPOP
FGFR3		EGFR	GNAS	MTOR	STK11
					TP53

B.



C.

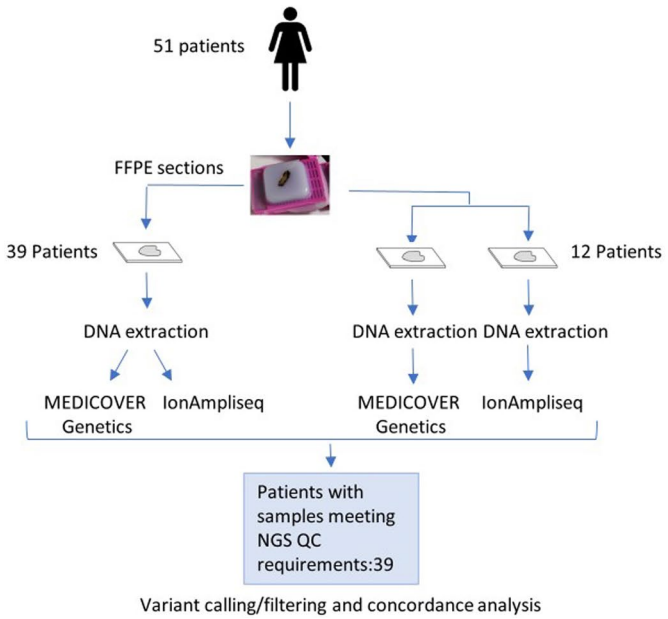


Fig. 1 Description of the assays and experimental strategy used. **A.** List of genes covered in the two assays used for molecular profiling of NSCLC tumor specimens. The IonAmliseq assay covers hotspot regions of 22 genes while the MEDICOVER Genetics custom tumor profile assay covers hotspots and other exonic regions of 49 genes. With red are the genes commonly covered by both assays. **B.** Venn diagram showing the size of the genomic area covered by both assays as well as their overlapping genomic coverage. **C.** Experimental strategy followed: FFPE specimens from 51 patients diagnosed with NSCLC were subjected to sectioning. For 39 patients, adequate amount of DNA was extracted from the same set of sections and sent to two labs for subsequent analysis with the Ion Ampliseq and the MEDICOVER assays. For 12 patients DNA derived from different sections of the same FFPE block was sent to the two labs for downstream processing. A total of 39 patients met QC parameters and proceeded to NGS. Available sequencing data were used for concordance analysis and estimation of the molecular profile of each tumor sample

Ten mutation and one variant were identified in regions exclusively covered by only one assay. A MET N375S variant (with no clinical actionability) was identified in patient 42 by the Ion Amliseq assay, while 10 more mutations of clinical significance in STK11, RET, PTEN, GNAS, TP53, BRCA2, PALB2, CHEK2, and PIK3CA were exclusively covered and identified by the MEDICOVER tumor profile assay in 10 patients (Fig. 3A, Table S2). Moreover, oncodriver mutations in BRAF, EGFR, KRAS and exon 10/exon21 PIK3CA were found to be mutually exclusive in this cohort (Table 2**,** Fig. 3).

Table 1 Patient characteristics. Clinical details for the 39 patients subjected to NGS analysis

Characteristic	n (%)
Age (years)	
Mean (SD)	60,8 (9,2)
Median (range)	61 (40–78)
Non-small cell lung cancer Pathological diagnosis	39 (100%)
Adenocarcinoma	29 (74,3%)
Squamous cell carcinoma	10 (25,7%)
Tumor stage	
I	2 (5,2%)
IA	1 (2,6%)
IB	7 (17,9%)
IIA	5 (12,8%)
IIB	5 (12,8%)
IIIA	9 (23,1%)
IIIB	5 (12,8%)
IV	5 (12,8%)
Adenocarcinoma metastatic status	
Non-metastatic (early-stage)	11 (38%)
Metastatic	18 (62%
Squamous cell carcinoma metastatic status	
Non-metastatic (early-stage)	2 (20%)
Metastatic	8 (80%)
Smoking status	
Non-smoking	24 (61,5%)
Smoking	15 (38,5%)

2.5 Comparison of clinical utility

In total, 52 unique mutations were identified, 50 mutations with the MEDICOVER tumor profile assay and 40 mutations with the Ion Ampliseq assay (Table 2, Table S1, Table S2). These, mutations were assessed based on i) their clinical utility including their association with sensitivity to an approved therapy, ii) resistance to an approved therapy (contraindicated therapy), iii) association with approved therapies in a different cancer type and iv) their investigation in clinical trials (either for their prognostic or potential therapeutic significance). The MEDICOVER Genetics custom tumor profile assay identified more mutations associated with approved therapy in NSCLC as compared to the Ion Ampliseq (6 versus 5 mutations respectively). In addition, 7 mutations associated with resistance to approved therapy (contra-indicated for use) were identified by the MEDICOVER Genetics assay as compared to 6 mutations with the Ion Ampliseq assay. Furthermore, the MEDICOVER Genetics assay identified 8 mutations associated with approved therapy for a different cancer type compared with 3 mutations identified with the Ion Amlpiseq assay. Thirteen mutations identified by the MEDICOVER assay were included in NCCN guidelines for NSCLC as compared to 11 mutations identified with the Ion Ampliseq assay. Finally, 49 mutations associated with clinical trials were identified by the MEDICOVER Genetics assay compared to 39 mutations identified by the Ion Amliseq assay (Figure 5, Table S3).

2.6 Association of mutations with patient survival and clinicopathological characteristics

The prognostic significance with respect to overall survival (OS) of the mutations identified in these patients was explored. Patients were divided in early (stage I-II) and late (III-IV) stage NSCLC with or without mutations in TP53 or KRAS. Although there was no statistically significant difference between the 4 groups, a trend was observed for improved OS in late-stage patients with wild-type KRAS or TP53 (Fig. 6). Due to the limited number of patients in the cohort it was not possible to investigate the prognostic significance of each gene separately. Next, associations of mutations in the 4 most frequently mutated genes with clinicopathological characteristics such as age, histology, smoking status were evaluated.

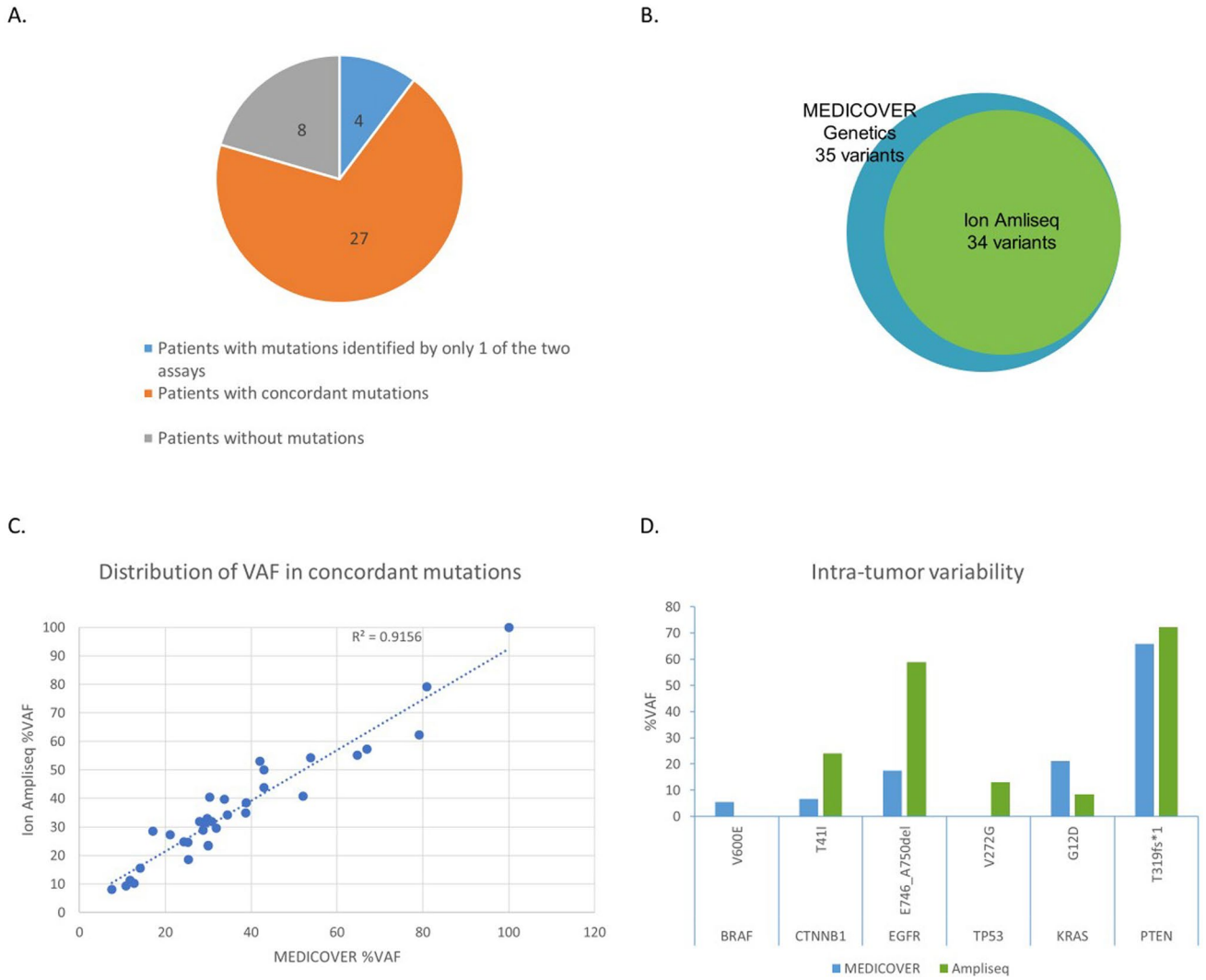


Fig. 2 Concordance analysis in the commonly covered regions by both assays. **A.** Frequency of NSCLC tumors with mutations identified by either one or both methods (concordant mutations). **B.** Venn diagram indicating the number of variants identified by either assay. 34 variants were commonly identified by both assays, while one extra variant was identified by the N MEDICOVER Genetics custom tumor profile assay (the KRAS G13C variant was detect below threshold level at VAF < 5% with the IonAmliseq and thus was excluded). **C.** Distribution of variant allele frequencies (VAF) for the concordant mutations identified by each assay. **D.** Variability of mutation detection and VAF for NGS data originating from different sections for the same FFPE biopsy

A statistically significant association between KRAS status and histology was observed (p=0.03099). Mutated KRAS was positively correlated with adenocarcinoma as opposed to squamous cell carcinoma (Table 3). However, due to the small size of the patient cohort, the results should be interpreted with caution.

3 Discussion

The purpose of this study was firstly to compare two different NGS-based tests for their clinical utility in patients with NSCLC and secondly to assess the clinical significance of the mutations identified and their associated with clinicopathological parameters. To this end, we employed two different assays; The Ion Ampliseq Colon Lung v2 assay (22 genes) and the MEDICOVER Genetics custom tumor profile assay (49 genes). We first compared the results of both assays in the commonly covered regions with highly similar results and high concordance of VAF. The MEDICOVER Genetics assay identified one additional mutation in the KRAS gene in this region, a KRAS G13C mutation of high diagnostic significance as it is associated with resistance to TKI EGFR inhibitors [23]. The IonAmpliseq assay failed to

Table 2 List of mutations in commonly covered regions

SampleID	Gene	AA change	CDS mutation	Type of alteration	COSMIC ID	MEDICOVER (%VAF)	Ampliseq (%VAF)
FEPE1	KRAS	p.G13C	c.37G>T	SNV	COSM527	5.90	NF
FEPE2	PIK3CA	p.E545K	c.1633G>A	SNV	COSM763	14.09	15.58
FEPE2	TP53	p.G266*	C.796G>T	SNV	COSM44891	42.98	43.75
FEPE3	KRAS	p.G12A	c.35G>C	SNV	COSM522	38.70	34.92
FEPE4	KRAS	p.G12D	c.35G>A	SNV	COSM521	10.79	9.23
FEPE5	TP53	p.R280I	c.839G>T	SNV	COSM11287	7.53	7.98
FEPE6	TP53	p.R273L	c.818G>T	SNV	COSM10779	12.69	10.26
FEPE8	TP53	p.G279E	c.836G>A	SNV	COSM43714	53.81	54.19
FEPE10	TP53	p.A159P	c.475G>C	SNV	COSM43836	30.00	23.29
FEPE13	TP53	p.Y163C	c.488A>G	SNV	COSM10808	28.86	31.33
FEPE15	KRAS	p.G12D	c.35G>A	SNV	COSM521	27.98	31.85
FEPE16	TP53	p.H178Tfs* 69	c.532del	INDEL	COSM111495	29.97	23.54
FEPE21	PIK3CA	p.E545K	c.1633G>A	SNV	COSM763	29.17	31.28
FEPE25	MET	p.T10101	c.3029C>T	SNV	COSM707	100	99.97
FEPE25	TP53	p.R280T	c.839G>C	SNV	COSM10724	38.84	38.38
FEPE29	TP53	p.R110L	c.329G>T	SNV	COSM10716	33.77	39.66
FEPE36	EGFR	p.L858R	c.2573 T>G	SNV	COSM6224	25.34	18.47
FEPE36	TP53	p.A159V	c.476C>T	SNV	COSM11148	29.75	32.86
FEPE38	TP53	p.H214R	c.641A>G	SNV	COSM43687	42.02	52.97
FEPE40	EGFR	p.E746_A750del	c.2236_2250del	INDEL	COSM6225	11.73	11.21
FEPE45	CTNNB1	p.G34E	c.1010G>A	SNV	COSM5671	34.46	34.15
FEPE45	EGFR	p.E746_A750del	c.2235_2249del	INDEL	COSM6223	25.30	24.56
FEPE45	TP53	p.I255F	c.763A>T	SNV	COSM43651	30.81	31.90
FEPE46	TP53	p.R337L	c.1010G>T	SNV	COSM11411	21.14	27.24
FEPE47	EGFR	p.L747_753delinsS	c.2240_2257del	INDEL	COSM12370	66.92	57.19
FEPE48	KRAS	p.G12A	c.35G>C	SNV	COSM522	30.36	40.32
FEPE48	PIK3CA	S.T1025=	c.3075C>T	SNV	COSM21451	28.83	28.91
FEPE48	TP53	p.P177=	c.531C>T	SNV	COSM43679	17.04	28.48
FEPE48	TP53	p.E349*	c.1045G>T	SNV	COSM10770	64.62	55.11
FEPE49	MET	p.T1010I	c.3029C>T	SNV	COSM707	52.05	40.78
FEPE49	TP53	p.V157F	c.469G>T	SNV	COSM10670	43.02	50.04
FEPE51	TP53	p.F212Sfs*3	c.635_636del	INDEL	COSM44162	24.26	24.76
FEPE55	TP53	p.E285K	c.853G>A	SNV	COSM10722	79.18	62.29
FEPE55	TP53	p.M246I	c.738G>A	SNV	COSM44310	80.84	79.06
FEPE57	KRAS	p.G12D	c.35G>A	SNV	COSM521	31.81	29.60

detect it above threshold (set at ≥5%) in the same DNA sample, but was detected in lower (VAF=3.78%). There are different possible explanations of what we observed. Even with the same DNA material some form of intra-individual variability is to be expected. Another possible explanation regarding this difference may originate form the different methodology these two assays employed (hybrid capture vs amplicon) and their linearity with regards to the true (unknown) VAF. The discrepancy in this specific variant is due to being marginally over the cutoff in one assay and slightly below cutoff in the other. An essential part of MEDICOVER Genetics capture technology is the design of TACS [24]. These are specifically designed to tolerate the presence of mismatches without compromising hybridisation efficiency and enrichment uniformity. Additionally, TACS capture flanking regions that may not be easily captured with amplicon based assays [25]. Most importantly, they ensure capture of all fragments thus providing a better representation of the complexity of the original DNA in the patient's tumor. These fundamental differences between the two methods can potentially explain the difference observed in regard to the KRAS G13C mutation

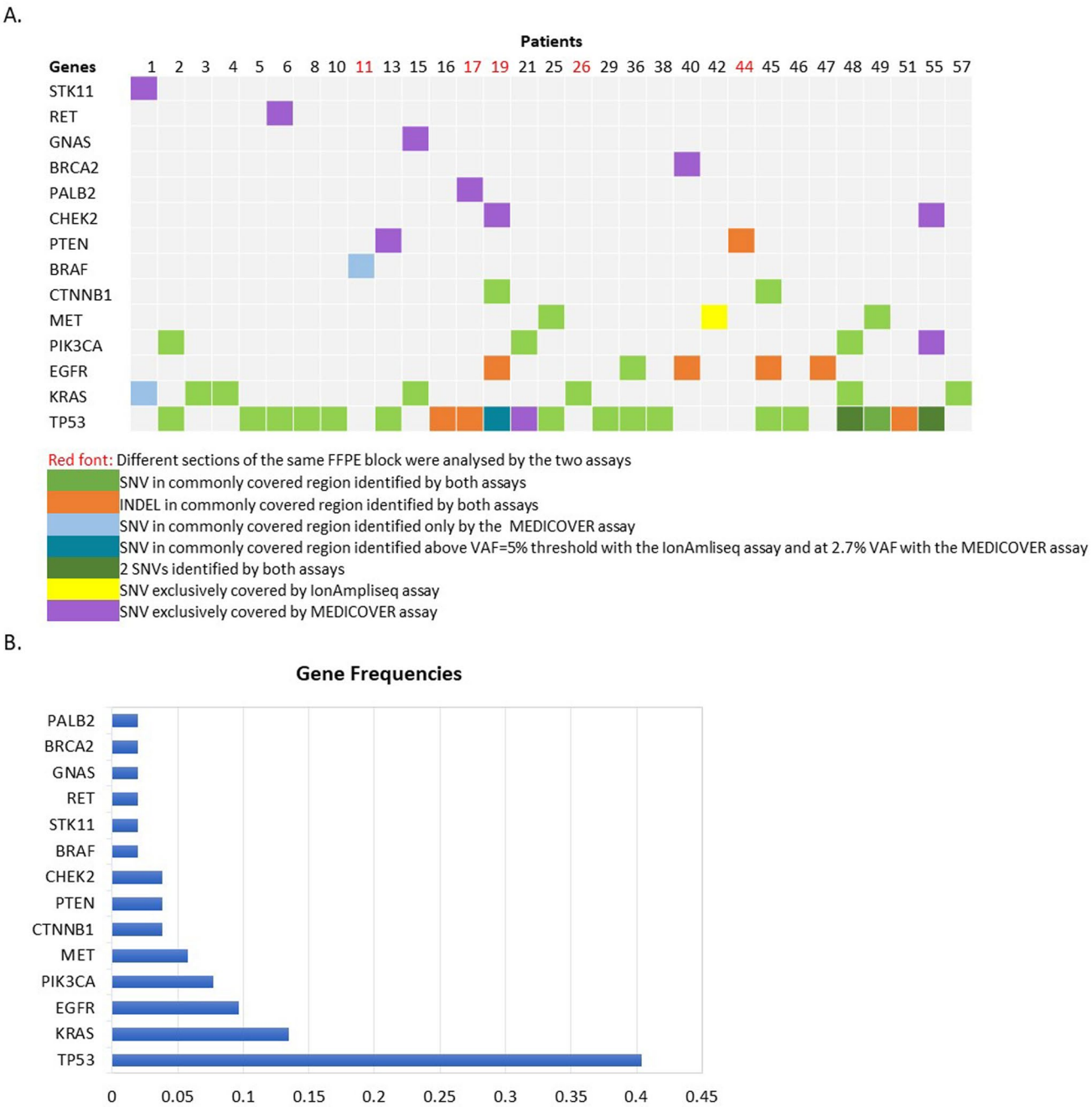


Fig. 3 Mutational profile of NSCLC tumors. **A.** Diagram depicting all the mutations identified by either both or at least one of the two assays. **B.** Frequency of mutations per gene identified in all patients

VAF represents the proportion of DNA fragments with a specific mutation relative to the total DNA fragments in a sample, and higher VAF values generally indicate a higher tumour burden, which can provide insights into the extent of tumour presence and heterogeneity. The sensitivity of mutation detection methods is influenced by VAF, with higher VAF values making it easier to detect mutations, while lower VAF values may challenge detection sensitivity and increase the risk of false negatives. VAF levels can impact treatment decisions by indicating the significance of mutations in disease progression; high VAF mutations may be more actionable and relevant for targeted therapies, whereas low VAF mutations might require more sensitive assays to monitor and could suggest emerging resistance.

Allele frequency in the context of tumour profiling refers to the proportion of DNA fragments carrying a particular mutation relative to the total DNA fragments in the sample. In samples with higher tumour content, the allele frequency

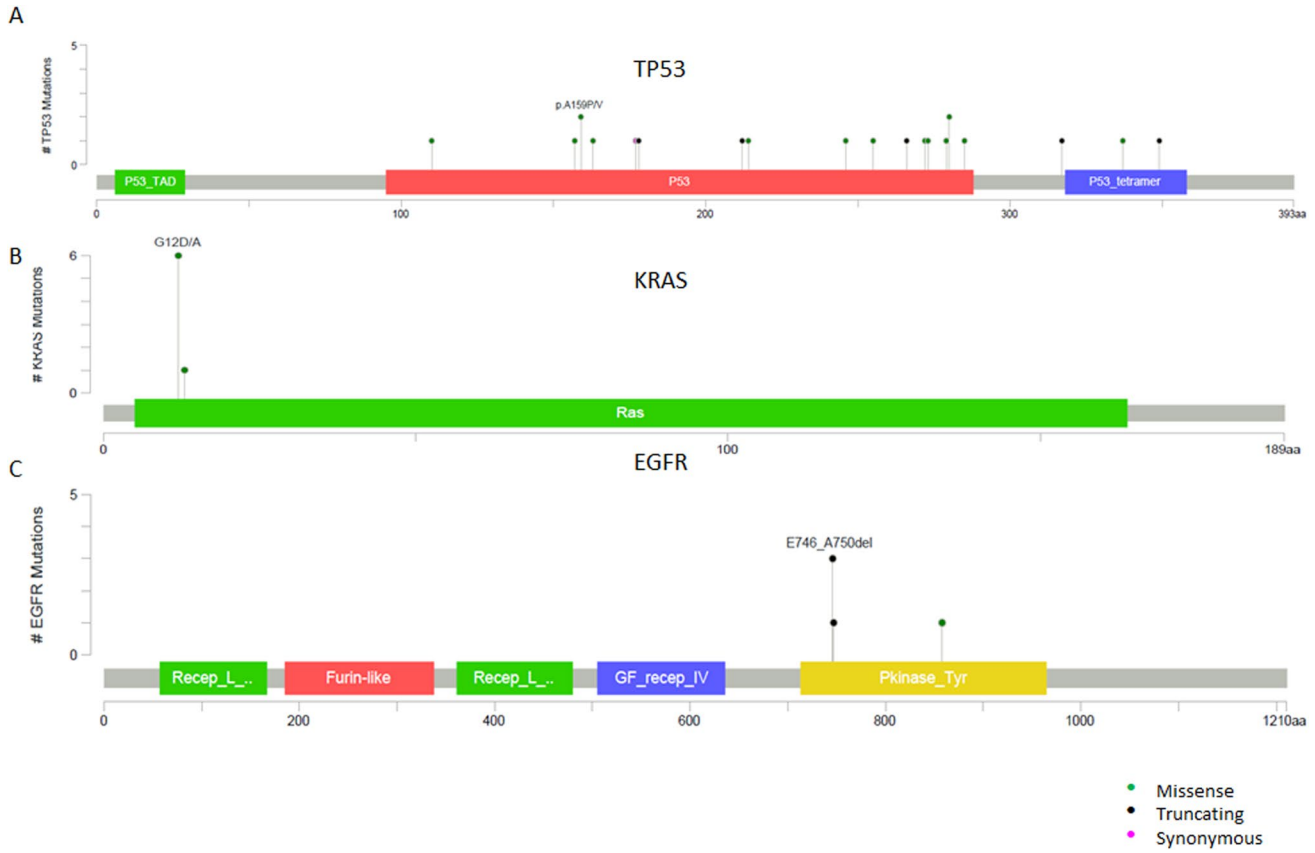


Fig. 4 Distribution of mutations in the most frequently mutated genes. Lollipop plots showing the distribution of mutations across the coding region of **A.** TP53 **B.** KRAS **C.** EGFR

of mutations will generally be higher. This is because a larger proportion of the DNA originates from tumour cells, which are more likely to contain the mutations of interest. When the tumour content is high, the higher allele frequency makes it easier to detect mutations, as the signal from the mutation is stronger relative to the non-tumour DNA. Higher allele frequencies in a sample can indicate a higher tumour burden, which might correlate with more aggressive disease. Understanding this can help in predicting how a patient might respond to certain treatments. It highlights why accurately assessing tumour content is crucial for reliable mutation detection and for making informed treatment decisions. This understanding is essential for both clinical practice and research in oncology.

Tumor heterogeneity is a challenge in clinical practice using FFPE tissue sections. It is well known that FFPE sections can only provide a snapshot of the tumor's molecular profile and cannot capture intra-tumor heterogeneity. Molecular heterogeneity is a well-known event in non-small cell lung cancer that can be attributed to different mechanisms related to structural chromosomal instability, somatic mutations, tumor mutational burden and genomic instability [8, 26]. In our results, we evaluated the degree of intra-tumor heterogeneity by analyzing different sets of FFPE sections from the patient's block by either assay. We observed differences in detection rate and allele frequencies of detectable variants that do not necessarily highlight differences between the two methods but are suggestive of intra-tumor heterogeneity. For example, two commonly covered variants, TP53 V227G in patient 19 and BRAF V600E in patient 11 were not identified in all sections tested from the same tumor whereas, KRAS G12D in patient 26 was identified in varying allelic balance (21.06% versus 8.46%) in different sections (Fig. 2D, Table S1).

Due to the comprehensive genomic coverage of the MEDICOVER Genetics custom tumor profile assay, a higher number of mutations were identified as compared to the Ion Amliseq method. The additional mutations identified are clinically actionable mutations with either available approved treatments or ongoing trials that investigate their prognostic and/or therapeutic significance. For example, a PIK3CA V344G mutation was identified in a region covered exclusively by the MEDICOVER Genetics' assay. This mutation resides in the C2 domain of the membrane-binding region of PI3K p110a and has shown sensitivity to p110a/PIK3CA-specific inhibitor alpelisib, a drug recently approved for the treatment of PIK3CA mutated HER2-negative metastatic breast cancer [27, 28]. Furthermore, mutations identified in BRCA2 and PALB2, genes

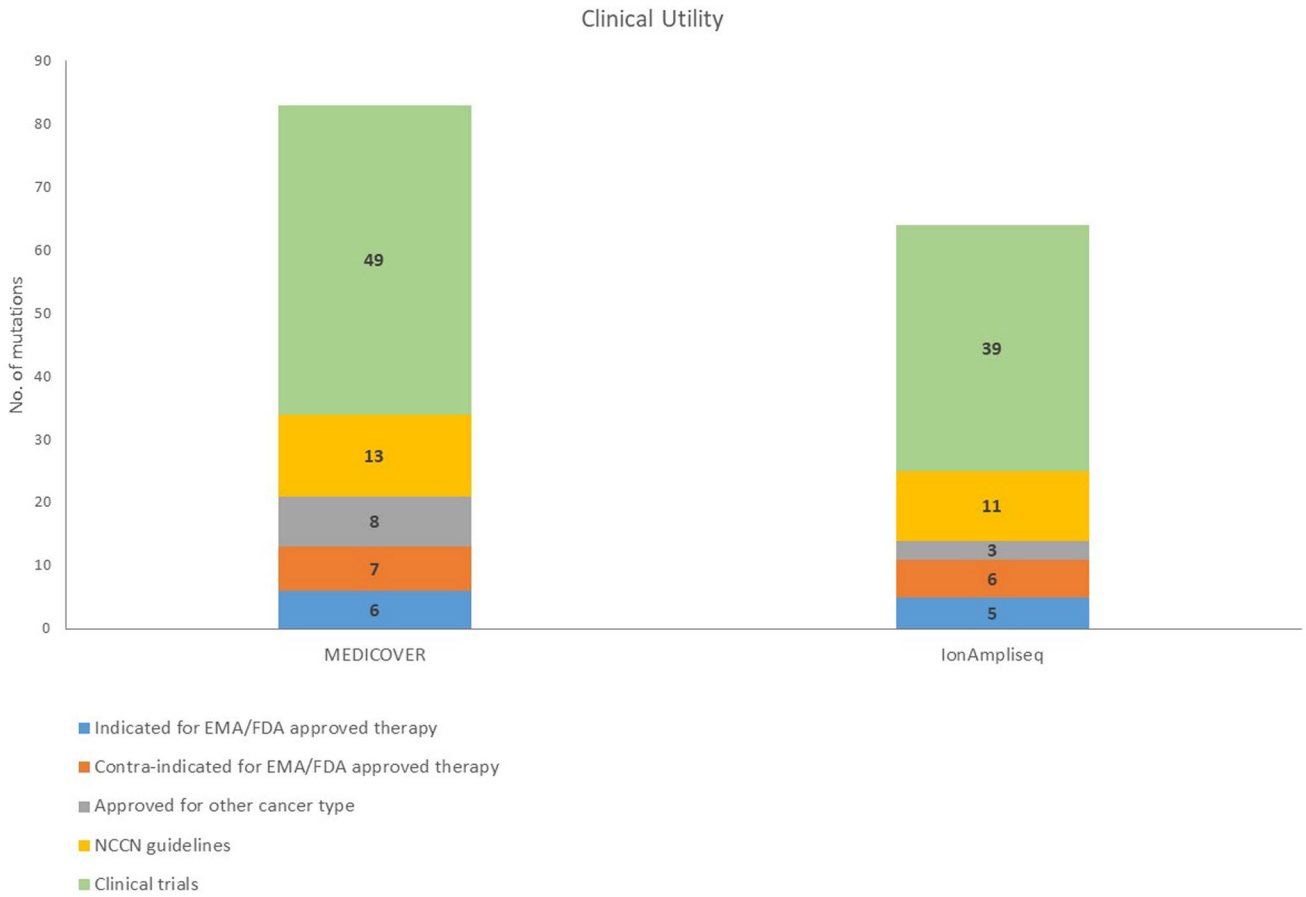


Fig. 5 Analysis of clinical utility of the two assays. The number of genes with mutations associated with sensitivity to approved therapy, resistance, or low sensitivity to therapy (contra-indicated), sensitivity to drugs approved for other cancer types and inclusion in clinical trials is indicated

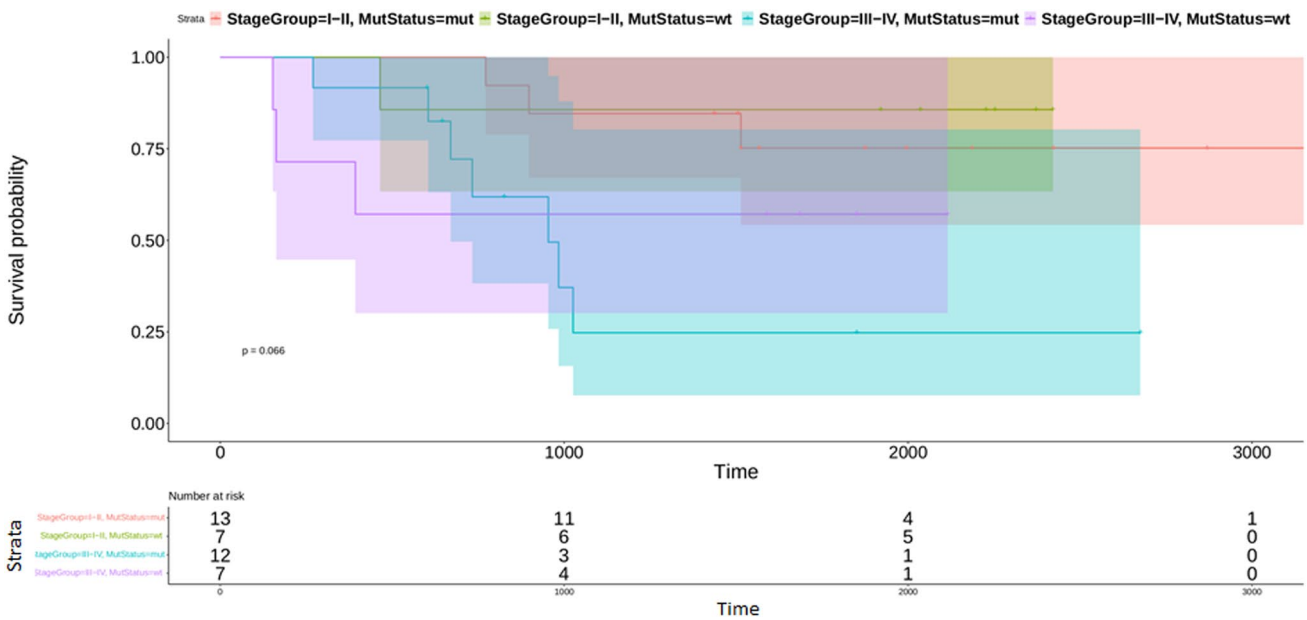


Fig. 6 Overall survival (OS) in patients with and without mutations in KRAS or TP53 in early (I-II) and late (III-IV) stage NSCLC. Red = Stage I-II/mutated, Green = StageI-II/wild type, blue = stageIII-IV/mutated, purple = stage III-IV/wild type

Table 3 Association of mutations with clinicopathological parameters

Characteristics	Gene	Group	Wild type	Mutant	p. value
Age	TP53	40–53	2(40%)	3(60%)	0.6983
		54–65	11(58%)	8(42%)	
		>65	7(47%)	8(53%)	
	KRAS	40–53	3(60%)	2(40%)	0.3364
		54–65	2(10.5%)	17(89.5%)	
		>65	12(80%)	3(20%)	
	EGFR	40–53	4(80%)	1(20%)	0.6235
		54–65	16(84%)	3(16%)	
		>65	14(93%)	1(7%)	
	PIK3CA	40–53	4(80%)	1(20%)	0.05566
		54–65	19(100%)	0(0%)	
		>65	12 (80%)	3(20%)	
Smoking	TP53	Yes	8(47%)	9(53%)	0.6427
		No	12(55%)	10(45%)	
	KRAS	Yes	15(88%)	2(12%)	0.368
		No	17(77%)	5(23%)	
	EGFR	Yes	15(88%)	2(12%6)	0.8619
		No	19(86%6)	3(14%)	
	PIK3CA	Yes	15(88%)	2(12%)	0.7857
		No	20(91%)	2(9%)	
Histology	TP53	ADC	17(59%)	12(41%)	0.1148
		SCC	3(30%)	7(70%)	
	KRAS	ADC	22(76%)	7(24%)	**0.03099**
		SCC	10(100%)	0(0%)	
	EGFR	ADC	25(86%)	4(14%)	0.7515
		SCC	9(90%)	1(10%)	
	PIK3CA	ADC	27(93%)	2(7%)	0.2675
		SCC	8(80%)	2(20%)	

Bold value indicates the statistically significant association between KRAS status and histology was observed (p=0.03099)

that play a critical role in the homologous recombination repair (HRR) mechanism, could represent potential therapeutic targets for poly(ADP-ribose) polymerase (PARP) 1, 2, 3 inhibitors such as rucaparib. A phase 2 study is currently investigating rucaparib for the treatment of solid tumors including lung cancer associated with deleterious mutations in HRR genes (NCT04171700) [29]. In addition, a GNAS R844C mutation has been identified by the MEDICOVER Genetics' assay. This mutation lies within a GTP binding region of the Gnas protein resulting in a loss of the GTPase activity and consequently leading to constitutive downstream pathway [30]. The GNAS R844C is shown to associate with resistance to targeted therapy in colorectal patients treated with vemurafenib, cetuximab and irinotecan combination treatment [31]. The clinical significance of this gene in NSCLC is under investigation [32]. Other mutations identified in genes such as PTEN and STK11 are also under investigation for their significance as potential targets of targeted therapy [29].

Furthermore, associations of NGS findings with OS and clinicopathological characteristics were investigated. KRAS or TP53 mutated NSCLC exhibited worse OS in the late-stage NSCLC. This finding is in agreement with previous studies showing that TP53 and KRAS are correlated with adverse prognosis in NSCLC [33, 34]. KRAS mutated tumors were also correlated with adenocarcinoma in our study. KRAS mutations predominately occur in lung adenocarcinomas with a frequency of 17% and are more rare in squamous cell carcinomas (4%) according with data retrieved from the COSMIC database [35]. The small sample size of this study does not allow for a conclusive association of molecular findings with clinical characteristics and highlights the need for larger validation studies.

In conclusion, both assays exhibited similar technical performance both in pre-analytical and post-analytical parameters in the commonly covered genomic areas. However, the more comprehensive coverage of the MEDICOVER Genetics custom tumor profile assay in clinically significant genes remarkably expands the potential for identifying additional

clinically actionable mutations. These additional mutations were mostly associated with approved therapies in other cancer types as well as clinical trials. Considering the rapid advancements in the molecular etiology of NSCLC and corresponding advancement in molecularly-targeted therapies, the expanded coverage of the custom-made MEDICOVER Genetics assay could potentially allow a more personalized clinical management for an increased number of patients.

4 Materials and methods

4.1 Patients

The study (No. 158200-13-688-219) has been approved by Vilnius Regional Biomedical Research Ethics Committee (Vilnius, Lithuania). All participants of the study have signed the informed consent to participate before study specific procedures started. Tumor tissue samples were collected at National Cancer Institute (Vilnius, Lithuania) and Vilnius University Hospital Santaros Klinikos (Vilnius, Lithuania).

4.2 DNA preparation

Fresh frozen or FFPE tumor tissue biopsy were collected from each patient. DNA extraction was performed using the Qiagen DNeasy blood and tissue kit and the QIAamp DNA FFPE Tissue Kit (Qiagen) for fresh frozen and FFPE tissue respectively, following the manufacturers' instructions. Tumor content was evaluated on FFPE specimens as the percentage of tumor cells in the total number of nucleated cells using hematoxylin/eosin staining. Minimum tumor content was 10% and maximum tumor content was 95%. DNA was quantified using a spectrophotometric assay (Cary 60 UV-Vis, Agilent Technologies) for fresh frozen tissue derived DNA and a fluorometric based assay for FFPE tissue-derived DNA (Qubit flex fluorometer, Qubit dsDNA high sensitivity assay, Thermo Scientific). A minimum of 10 ng of DNA and a minimum DNA concentration of 1 ng/µl (as measured by a fluorometric based method for FFPE samples) were used as thresholds for library preparation for both assays. DNA quality was assessed using Agilent D1000 ScreenTape analysis (cat.no. 5067–5582).

4.3 Ion ampliseq colon and lung cancer research panel library preparation and sequencing

Libraries were amplified using Ion AmpliSeq Colon and Lung Cancer Research Panel (Ion Torrent by Life Technologies) which analyzes amplicons in hotspots and target regions of 22 oncogenes (Fig. 1A) covering single nucleotide variants (SNVs) and insertions and deletions (Indels) involved in colon and lung cancers. 10 ng of DNA were amplified using Ion AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific) following the manufacturers' instructions. The library concentration was quantified with Ion Library TaqMan Quantitation Kit (Thermo Fisher Scientific). Each library was diluted to reach a concentration of 100pM and amplified using emulsion PCR. Sequencing was performed on the Ion PGM (Thermo Fisher Scientific) using the Ion PGM 200 Sequencing Kit (Thermo Fisher Scientific). Sequencing libraries were loaded onto a 316 chip following the manufacturers' instructions.

4.4 MEDICOVER Genetics custom tumor profile library preparation and sequencing

DNA libraries were prepared from sheared DNA based on previously established protocols [37]. Briefly, blunt ending and 5' phosphorylation was performed using T4 polymerase and T4 kinase respectively. Following adaptor ligation using T4 Ligase (New England Biolabs, Ipswich, UK), nicks were removed using Bst polymerase (New England Biolabs). Unique barcodes were assigned to all samples in a final PCR reaction using Herculase II Fusion Polymerase (Agilent Technologies, Santa Clara, CA). At each step, products were purified using magnetic beads. DNA enrichment for the genomic regions of interest, was carried out using an in solution- hybridization based method using TACS (TArget Capture Sequences) specifically designed to capture selected loci in the genes of interest. Biotinylated TACS were then immobilized on streptavidin coated magnetic beads for subsequent hybridization with the DNA libraries. Eluted samples were amplified using outer-bound adaptor primers. Enriched DNA libraries were then normalized and subjected to sequencing on an Illumina sequencing platform. The MEDICOVER Genetics tumor profile panel was used for the identification of single nucleotide variants SNVs and indels in hotspot regions and selected targeted genomic loci of 49 genes (Fig. 1A).

4.5 Bioinformatics and data analysis

Sequencing data were de-multiplexed with bcl2fastq (v.2.16.0) and aligned to the human genome build 37 (hg19) to generate alignment (bam) files. Specifically, for each sample, paired-end DNA sequencing reads were processed with cutadapt (v.1.8.1) to remove adapter sequences and poor-quality reads (quality base cutoff 25). Reads with length less than 25bps were also removed from further analysis. The remaining sequences were aligned to the human reference genome build 37 (hg19) using the Burrows-Wheeler alignment algorithm (bwa mem). For the MEDICOVER Genetics custom tumor profile assay duplicate read entries were removed to convert aligned reads to a binary (BAM) file containing uniquely aligned read entries only. Per base allele-specific read-depth information was retrieved from this final BAM file. All samples with a minimum depth of coverage of 250 reads proceeded to variant calling with vardict [38]. For concordance analysis, a list of selected targeted genomic coordinates that were commonly covered by both the Ion Ampliseq method and the MEDICOVER Genetics custom tumor profile assay was used. A threshold for VAF (variant allele frequency) for data generated by either method was set at ≥5%. Analysis was also performed in the genomic regions exclusively covered by either the Ion Ampliseq or the MEDICOVER Genetics custom tumor profile assay. Assessment of somatic/germline status was not assessed as germline DNA was not available for this patient cohort.

4.6 Assessment of clinical utility

Assessment of clinical significance was performed by retrieving information from multiple databases. Specifically, pharmacological information was retrieved from the PharmGKB database which consolidates available data on therapies approved by various regulatory authorities including the Food and drug administration (FDA) and the European Medicines Agency (EMA) [39]. Information on association of genes and clinical trials for the indicated cancer type were retrieved from Clinical.trials.gov, a resource provided by the U.S National Library of Medicine [29].

4.7 Statistical analysis

Survival analysis was carried out in R, using the Kaplan-Meier estimator (survival package in R) [40]. The G-test (R package DescTools) was used to test association for 2x2 contingency tables, specifically between age group, smoking status and histology findings against mutational findings from selected genes [41].

Acknowledgements The authors are indebted to patients and their families for their trust and participation in the present registry. The authors wish to thank Chrysovalanto Marinou, Louisa Constantinou, Charalambos Kkoufou for their laboratory assistance.

Institutional Review Board Statement The study (No. 158200-13-688-219) has been approved by Vilnius Regional Biomedical Research Ethics Committee (Vilnius, Lithuania). All participants of the study have signed the informed consent to participate before study specific procedures started. Tumor tissue samples were collected at National Cancer Institute (Vilnius, Lithuania) and Vilnius University Hospital Santaros Klinikos (Vilnius, Lithuania). The study was conducted according to the guidelines of the Declaration of Helsinki. Informed consent was obtained from all subjects involved in the study.

Author contributions Conceptualization, PCP, JG, GK methodology ID, AE, KT formal analysis ID, AA, CL, AE; in-vestigation, AE resources, PCP, GK, JG data curation, ID, AA, CL, AK, DS writing—original draft preparation, ID, JG writing—review and editing, AE, MI, CL, AA, DS supervision, GK, AE, EK, MI, SC project administration, AK funding acquisition, PCP

Funding This research received no external funding.

Data availability The data presented in this study are available on request from the corresponding author.

Declarations

Competing interests CL: Employed by MEDICOVER Genetics has filed a PCT patent application for Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); AA: Employed by MEDICOVER Genetics; has filed a PCT patent application for the Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); has filed a PCT patent application for the Enrichment of Targeted Genomic Regions for Multiplexed Par-allel Analysis (WO2019/008148A9); EK: Employed by MEDICOVER Genetics; has filed a PCT pa-tent application for the Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); has filed a PCT patent application for the Enrichment of Tar-geted Genomic Regions for Multiplexed Parallel Analysis (WO2019/008148A9); KT: Employed by MEDICOVER Genetics, has filed a PCT patent application for the Enrichment of Targeted Ge-nomic Regions for Multiplexed Parallel Analysis (WO2019/008148A9); AE: Employed by MEDI-COVER Genetics; has filed a PCT patent application for the Target-enriched

multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); MI: Employed by MEDICOV-ER Genetics; has filed a PCT patent application for the Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1), has filed a PCT patent applica-tion for the Enrichment of Targeted Genomic Regions for Multiplexed Parallel Analysis (WO2019/008148A9); GK: Employed by MEDICOVER Genetics; has filed a PCT patent application for the Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); has filed a PCT patent application for the Enrichment of Targeted Genomic Regions for Multiplexed Parallel Analysis (WO2019/008148A9); PCP: Employed by MEDICOVER Genetics; has filed a PCT patent application for the Target-enriched multiplexed parallel analysis for assessment of tumor biomarkers (WO2019/008172A1); has filed a PCT patent application for the Enrichment of Targeted Genomic Regions for Multiplexed Parallel Analysis (WO2019/008148A9); The rest of the authors declare no conflict of interest.

References

1. Bray F, et al. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018. https://doi.org/10.3322/caac.21492.
2. Inamura K. Lung cancer: understanding its molecular pathology and the 2015 wHO classification. Front Oncol. 2017. https://doi.org/10.3389/fonc.2017.00193.
3. De Perrot M, et al. Sex differences in presentation, management, and prognosis of patients with non-small cell lung carcinoma. J Thorac Cardiovasc Surg. 2000. https://doi.org/10.1016/S0022-5223(00)70213-3.
4. Pinto JA, et al. Gender and outcomes in non-small cell lung cancer: An old prognostic variable comes back for targeted therapy and immunotherapy? ESMO Open. 2018. https://doi.org/10.1136/esmoopen-2018-000344.
5. Ettinger DS, et al. NCCN guidelines insights: non-small cell lung cancer, version 1.2020. J Natl Compr Cancer Netw. 2019. https://doi.org/10.6004/jnccn.2019.0059.
6. John T, Liu G, Tsao MS. Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors. Oncogene. 2009. https://doi.org/10.1038/onc.2009.197.
7. Planchard D, et al. Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018. https://doi.org/10.1093/annonc/mdy275.
8. Jamal-Hanjani M, et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med. 2017. https://doi.org/10.1056/NEJMoa1616288.
9. Yeh P, et al. DNA-mutation inventory to refine and enhance cancer treatment (DIRECT): a catalog of clinically relevant cancer mutations to enable genome-directed anticancer therapy. Clin Cancer Res. 2013. https://doi.org/10.1158/1078-0432.CCR-12-1894.
10. Feng H, et al. Identification of genetic mutations in human lung cancer by targeted sequencing. Cancer Inform. 2015. https://doi.org/10.4137/CIN.S22941.
11. Yang JCH, et al. Osimertinib in pretreated T790M-positive advanced non-small-cell lung cancer: AURA study phase II extension component. J Clin Oncol. 2017. https://doi.org/10.1200/JCO.2016.70.3223.
12. Oxnard GR, et al. Assessment of resistance mechanisms and clinical implications in patients with EGFR t790m-positive lung cancer and acquired resistance to Osimertinib. JAMA Oncol. 2018. https://doi.org/10.1001/jamaoncol.2018.2969.
13. Planchard D, et al. Dabrafenib in patients with BRAFV600E-positive advanced non-small-cell lung cancer: a single-arm, multicentre, open-label, phase 2 trial. Lancet Oncol. 2016. https://doi.org/10.1016/S1470-2045(16)00077-2.
14. Salgia R, Sattler M, Scheele J, Stroh C, Felip E. The promise of selective MET inhibitors in non-small cell lung cancer with MET exon 14 skipping. Cancer Treat Rev. 2020. https://doi.org/10.1016/j.ctrv.2020.102022.
15. Bernicker EH, Allen TC, Cagle PT. Update on emerging biomarkers in lung cancer. J Thorac Dis. 2019. https://doi.org/10.21037/jtd.2019.01.46.
16. Peters S, et al. Activity of afatinib in heavily pretreated patients (pts) with HER2 mutation-positive (HER2m1) advanced non-small cell lung cancer (NSCLC): findings from a global named patient use (NPU) program. Ann Oncol. 2017. https://doi.org/10.1093/annonc/mdx380.057.
17. Li SD, et al. Cancer gene profiling in non-small cell lung cancers reveals activating mutations in JAK2 and JAK3 with therapeutic implications. Genome Med. 2017. https://doi.org/10.1186/s13073-017-0478-1.
18. Gristina V, et al. Non-small cell lung cancer Harboring concurrent EGFR genomic alterations: a systematic review and critical appraisal of the double dilemma. J Mol Pathol. 2021. https://doi.org/10.3390/jmp2020016.
19. Ascierto PA, Bifulco C, Palmieri G, Peters S, Sidiropoulos N. Preanalytic variables and tissue stewardship for reliable next-generation sequencing (NGS) clinical analysis. J Mol Diagn. 2019. https://doi.org/10.1016/j.jmoldx.2019.05.004.
20. Tanaka T, Watanabe M, Yamashita K. Potential therapeutic targets of TP53 gene in the context of its classically canonical functions and its latest non-canonical functions in human cancer. Oncotarget. 2018. https://doi.org/10.18632/oncotarget.24611.

21. Martin P, Leighl NB, Tsao MS, Shepherd F. KRAS mutations as prognostic and predictive markers in non-small cell lung cancer. J Thorac Oncol. 2013;8:530–42.
22. Griffith M, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet. 2017;49:170–4.
23. Massarelli E, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small cell lung cancer. Clin Cancer Res. 2007. https://doi.org/10.1158/1078-0432.CCR-06-3043.
24. Koumbaris G, et al. Cell-Free DNA analysis of targeted genomic regions in maternal plasma for non-invasive prenatal testing of trisomy 21, trisomy 18, trisomy 13, and fetal sex. Clin Chem. 2016;62:848–55.
25. Samorodnitsky E, et al. Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. Hum Mutat. 2015. https://doi.org/10.1002/humu.22825.
26. Marino FZ, et al. Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications. Int J Med Sci. 2019. https://doi.org/10.7150/ijms.34739.
27. Packer LM, et al. PI3K inhibitors synergize with FGFR inhibitors to enhance antitumor responses in FGFR2mutant endometrial cancers. Mol Cancer Ther. 2017. https://doi.org/10.1158/1535-7163.MCT-16-0415.
28. Lama Tamang TG, et al. Use of alpelisib in the treatment of hormone receptor positive metastatic breast cancer: an institutional experience. J Clin Oncol. 2020. https://doi.org/10.1200/jco.2020.38.15_suppl.e15216.
29. Clinical trials.gov. *U.S. National Institutes of Health* (2015).
30. Huang L, et al. The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. J Am Med Informatics Assoc. 2017. https://doi.org/10.1093/jamia/ocw148.
31. Hong DS, et al. Phase IB study of vemurafenib in combination with irinotecan and cetuximab in patients with metastatic colorectal cancer with BRAFV600E mutation. Cancer Discov. 2016. https://doi.org/10.1158/2159-8290.CD-16-0050.
32. Ritterhouse LL, et al. GNAS mutations in primary mucinous and non-mucinous lung adenocarcinomas. Mod Pathol. 2017. https://doi.org/10.1038/modpathol.2017.88.
33. Svaton M, et al. The prognostic role of kras mutation in patients with advanced NSCLC treated with second-or third-line chemotherapy. Anticancer Res. 2016;36(3):1077.
34. Gu J, et al. TP53 mutation is associated with a poor clinical outcome for non-small cell lung cancer: evidence from a meta-analysis. Mol Clin Oncol. 2016. https://doi.org/10.3892/mco.2016.1057.
35. Forbes SA, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 2017;45:D777–83.
36. Rathi V, et al. Clinical validation of the 50 gene AmpliSeq Cancer Panel V2 for use on a next generation sequencing platform using formalin fixed, paraffin embedded and fine needle aspiration tumour specimens. Pathology. 2017. https://doi.org/10.1016/j.pathol.2016.08.016.
37. Neofytou MC, et al. Targeted capture enrichment assay for noninvasive prenatal testing of large and small size sub-chromosomal deletions and duplications. PLoS ONE. 2017. https://doi.org/10.1371/journal.pone.0171319.
38. Lai Z, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res. 2016. https://doi.org/10.1093/nar/gkw227.
39. Whirl-Carrillo M, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther. 2012. https://doi.org/10.1038/clpt.2012.96.
40. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1992. https://doi.org/10.1007/978-1-4612-4380-9_25.
41. Quine MP, Robinson J. Efficiencies of chi-square and likelihood ratio goodness-of-fit tests. Ann Stat. 2007. https://doi.org/10.1214/aos/1176349550.