The relevance of tumor mutation profiling in interpretation of NGS data from cell-free DNA in non-small cell lung cancer patients

Anine Larsen Ottestad, Sissel G.F. Wahl, Bjørn Henning Grønberg, Frank Skorpen, Hong Yan Dai

PII: S0014-4800(19)30609-4

DOI: https://doi.org/10.1016/j.yexmp.2019.104347

Reference: YEXMP 104347

To appear in: *Experimental and Molecular Pathology*

Received date: 7 August 2019

Revised date: 11 November 2019

Accepted date: 19 November 2019

Please cite this article as: A.L. Ottestad, S.G.F. Wahl, B.H. Grønberg, et al., The relevance of tumor mutation profiling in interpretation of NGS data from cell-free DNA in non-small cell lung cancer patients, *Experimental and Molecular Pathology*(2018), https://doi.org/10.1016/j.yexmp.2019.104347

The relevance of tumor mutation profiling in interpretation of NGS data from cell-free DNA in non-small cell lung cancer patients

Anine Larsen Ottestad[1,2,*] anine.l.ottestad@ntnu.no, Sissel G. F. Wahl[1,3], Bjørn Henning Grønberg[1,2], Frank Skorpen[1,2], Hong Yan Dai[1,3]

[1]Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

[2]Cancer Clinic, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway.

[3]Department of Pathology, Clinic of Laboratory Medicine, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway

[*]Corresponding author at: The Cancer Clinic, St. Olav's Hospital, Trondheim University Hospital, P.O. box 3250 Sluppen, 7006 Trondheim, Norway

Abstract

Studies have indicated that detection of circulating tumor DNA (ctDNA) prior to treatment is a negative prognostic marker in non-small cell lung cancer (NSCLC). ctDNA is currently identified by detection of tumor mutations. Commercial next-generation sequencing (NGS) assays for mutation analysis of ctDNA for routine practice usually include small gene panels and are not suitable for general mutation analysis. In this study, we investigated whether mutation analysis of cfDNA could be performed using a commercially available comprehensive NGS gene panel and bioinformatics workflow. Tumor DNA, plasma DNA and peripheral blood leukocyte DNA from 30 NSCLC patients were sequenced. In two patients (7%), tumor mutations in cfDNA were immediately called by the bioinformatic workflow. In 13 patients (43%), tumor mutations were not called, but were present in ctDNA and were identified based on the known tumor mutation profile. In the remaining 15 patients (50%), no concordant mutations were detected. In conclusion, we were able to identify tumor mutations in ctDNA from 57% of NSCLC patients using a comprehensive gene panel. We demonstrated that sequencing paired tumor DNA was helpful to interpret data and confirm ctDNA, and thus increased the ratio of patients with detectable ctDNA. This approach might be feasible for mutation analysis of ctDNA in routine diagnostic practice, especially in case of suboptimal plasma quality and quantity.

Keywords: non-small cell lung cancer (NSCLC), next-generation sequencing (NGS), circulating tumor DNA (ctDNA)

Introduction

Tumors release DNA, known as circulating tumor DNA, which makes up a small fraction of total cell-free DNA (cfDNA) in the blood.[1] Many studies have shown that tumor mutations can be detected in cfDNA from patients with non-small cell lung cancer (NSCLC).[2,3] It has also been shown that pre-

treatment detection of tumor mutations in cfDNA is a negative prognostic factor. cfDNA is thereby emerging as an important biomarker.

A requirement for using cfDNA as a prognostic biomarker is to detect at least one tumor mutation in cfDNA prior to treatment. Because the mutation spectrum is diverse in NSCLC, it is often necessary to analyze many different genes to ensure mutation detection.[4] Commercial cfDNA next-generation sequencing (NSG) assays for routine practice usually contain small gene panels with mainly targetable genes which are not suitable for this purpose.

Mutation analysis of cfDNA using a large gene panel is challenging. It generates large data sets and sometimes the data quality is poor. Suboptimal quality and quantity of plasma are the main factors that affect the quality of sequencing data due to insufficient amount and low quality of input DNA used for NGS library preparation. This is a frequent issue, especially in the routine practice.

It is promising that recent studies using large gene panels found concordant mutation in tumor DNA and cfDNA in 50-100% of early stage NSCLC patients.[3,5,6] In these studies, customized gene panels and own-developed bioinformatic workflows were used for mutation detection and interpretation of cfDNA sequencing data. These approaches may not be directly transferable to most routine diagnostic laboratories in which commercial assays are usually applied.

In this study we used a commercial comprehensive gene panel that included hot spots in 275 genes. We sequenced matched tumor DNA, peripheral blood leukocyte (PBL) DNA and cfDNA from 30 patients with NSCLC. The aim was to investigate whether tumor mutations could be detected in cfDNA using a large commercially available gene panel and bioinformatic workflow.

Methods

Patient material and approvals

Tumor tissue and blood samples were retrieved from Biobank1, a local lung cancer biobank of tumor tissue, cytological specimens, blood samples and clinical data from more than 900 patients with all stages and subtypes of lung cancer. The biobank is approved by the Norwegian Regional Committee

for Medical and Health Research Ethics (REC) Central, the Norwegian Health Department, and the Norwegian Data Protection Authority. The REC central has approved the present study.

Patients with NSCLC and both tumor tissue and blood samples available for histological examination and for DNA extraction were included in this study. Blood samples were collected before treatment commenced. Tumor specimens were reviewed, classified and subtyped according to the 2015 World Health Organization classification of lung tumors by a lung cancer pathologist (SGFW).[7] Disease stage was assessed according to the 8th TNM Classification of Malignant Tumors for lung cancer.[8]

DNA extraction

Formalin-fixed (formaldehyde solution 4% phosphate buffered) paraffin-embedded (FFPE) tumor blocks were used for isolation of tumor DNA. Tissue sections of 10 µm were cut from areas with the highest proportion of tumor cells. Tumor cell content ranged from 5-50%. From seven tumors, two replicating DNA extracts were prepared by cutting the same tumor tissue twice. Both DNA extracts were sequenced. DNA was extracted using GeneRead DNA FFPE kit (Qiagen, Valencia, CA) or QIAamp DNA FFPE Tissue kit (Qiagen, Valencia, CA) according to the respective protocols. Treatment of DNA with uracil-DNA glycosylase was performed either during extraction or after final elution to remove uracil. Spontaneous deamination of cytosine to uracil occurs over time, and this would lead to C>T and G>A sequence artifacts.

Plasma was prepared from 10 mL whole blood with EDTA or citrate anticoagulant immediately after sampling. Blood samples were either centrifuged once at 2,500x g for 10 min, or first at 1,500x g for 15 min and then at 10,000x g for 10 min. Plasma was transferred to cryotubes and stored at -80 °C. cfDNA was isolated using QIAamp Circulating Nucleic Acid kit (Qiagen, Valencia, CA) from 2.5-6 mL plasma and eluted with 30-50 µL supplied buffer. It was further concentrated by precipitation with 0.3M sodium acetate, 0.05-1µg/µL glycogen (Thermo Fisher Scientific, Waltham, MA) and 2.5 x volume of 96% ethanol and finally resuspended in 10 µL molecular biology grade water.

DNA from PBLs was extracted from whole blood using QIAsymphony DSP DNA kit (Qiagen, Valencia, CA). DNA concentration was measured by Qubit® (Thermo Fisher Scientific, Waltham, MA), using dsDNA HS Assay Kit for cfDNA and dsDNA BR Assay Kit for tumor DNA and PBL DNA.

NGS library preparation

NGS libraries were prepared from 30-210 ng tumor DNA, 20 ng cfDNA and 20-40 ng PBL DNA using QIAseq Comprehensive Cancer Panel (Qiagen, Valencia, CA). The panel included 836,670 bases in 275 genes (Table S4). Briefly, DNA was fragmented, end repaired and a 5' adenine overhang was made. Synthetic fragments with a sequencing adapter, unique molecular index (UMI, a 12-nucleotide random sequence) and a 3' thymine overhang were then ligated to all DNA fragments. An estimated size of 4^{12} different UMI sequence combinations ensured that all DNA fragments were tagged with a unique UMI.

In the first polymerase chain reaction (PCR) of six cycles, PCR products covering the genomic areas of interest were synthesized using gene-specific primers and primers complementary to the 5' sequencing adapter. There were a total of 11,311 different gene-specific primers, which all had a common sequence, a primer site, at their 5' end. The PCR products were then used as templates for the second PCR, in which complete NGS libraries were made using primers for the 5' sequencing adapters and 3'sequencing adapters complementary to the common 5' end of the templates.

Tumor DNA libraries and cfDNA libraries were pooled separately with 5 µL from each individual tumor DNA and cfDNA library, respectively. The pooled libraries were then purified using QIAquick PCR Purification kit (Qiagen, Valencia, CA). To acquire a library pool with even and proper fragment lengths, electrophoresis of both library pools was performed using DNA 300 Chip (Caliper Life Sciences, Hopkinton, MA) on LabChip XT (Caliper Life Sciences, Hopkinton, MA). Fragments with a length of 295 to 445 bp were extracted according to the procedure "Extract and Pause". The two pools were then quantified using KAPA Library Quantification Kit – Illumina ABI Prism19 (KAPA Biosciences, Wilmington, MA). Libraries were sequenced by NextSeq 550 (Illumina, San Diego, CA)

with 151 paired-end reading. NGS libraries of PBL DNA were prepared, but without the step of fragment purification, and sequenced separately in the same manner as tumor DNA and cfDNA.

Bioinformatic analysis

Biomedical Genomics Workbench version 5.0.1 (Qiagen, Valencia, CA) with the gene panel-specific plugin QIAseq Targeted Panel Analysis version 1.2 was applied for detection of variants. Briefly, adapters and UMIs were removed and the reads annotated with their UMI to allow mapping to the reference genome (hg19). Reads with same UMI (i.e PCR duplicates) were grouped into a "UMI family". A consensus sequence was assembled from each UMI family with the variants that were present in ≥75% of the duplicates. The subsequent steps were performed with the consensus sequence. Sequences of two DNA fragments that were generated during adapter/UMI-ligation in the NGS library preparation, were removed. Inserts and deletions were then detected, followed by local realignment and removal of primer sequences. Variants were then called in the panel target area with the tool "Low frequency variant detection".

For each patient, variants that were detected in ≥10 different UMI families in PBL DNA were removed from both tumor DNA and cfDNA. Tumor mutations were defined as follows; coding, non-synonymous, not located in a homopolymer, detected equally in read 1 and 2, and detected in ≥5 big UMI families (defined as families with ≥3 duplicates). cfDNA mutations were defined as follows; coding, non-synonymous, not located in a homopolymer, detected equally in read 1 and 2, and detected in ≥3 big UMI families.

The lists of mutations in tumor DNA passing these filters were uploaded to SNPnexus to evaluate pathogenicity.[9] A mutation was classified as pathogenic if it was registered in the catalogue of somatic mutations in cancer (COSMIC)[10] or predicted to be pathogenic by both SIFT[11] and PolyPhen.[12] Only mutations classified as pathogenic were included for concordance analysis.

Mutation spectrum was compared between paired tumor DNA and cfDNA. Tumor mutation spectrum from two synchronous tumors were separately compared to the matched cfDNA in one patient (patient 6). The BAM files created after UMI family creation in the cfDNA samples were

manually inspected for presence of matched tumor mutations that were not called by the bioinformatic workflow.

Cases with no concordant mutations were classified as undetected. Patients with concordant mutations were divided into two categories based on the method of detection. Cases with concordant mutations that were immediately called by the bioinformatic workflow in cfDNA were classified in the first category. Cases where concordant mutations were identified in cfDNA with prior knowledge from tumor DNA were classified in the other category.

Results

Thirty-one patients diagnosed with NSCLC at St. Olav's Hospital, Trondheim University Hospital, Norway between 2009 and 2016 were included (Table S1). One patient was excluded from the analyses because of failed PBL DNA library sequencing. Patient characteristics of the remaining 30 patients are shown in Table 1. Histological diagnoses were adenocarcinoma (n=15, 50%), squamous cell carcinoma (n=13, 43%), adenosquamous carcinoma (n=1, 3%) and NSCLC-not otherwise specified (NSCLC-NOS, n=1, 3%). Twenty two patients (77%) had stage I-II disease and seven patients (23%) had stage III-IV.

Sequencing coverage, i.e. the mean number of sequencing reads covering each position in the target area, was on average 2,364x (range 1,031x-4,338x) across all tumor DNA samples, 2,143x (range 1,055x-3,625x) across all cfDNA samples and 748x (range 515x-1,630x) across all PBL DNA samples (Table S2).

The UMI family consensus sequence represents one original DNA fragment, and by extension, one haploid genome. Mean haploid genome coverage was on average 743x (range 225x-1643x) in all tumor DNA samples, 207x (range 69x-443x) in all cfDNA samples and 493x (range 209x-1061x) in all PBL DNA samples (Table S1). The median number of PCR duplicates in each UMI family was ≤ 3 in all samples.

Mutation analyses

At least one mutation was detected in tumor from 29/30 patients (97%), with a mean allele frequency (AF) of 24.4% (range 1.4-73.8%). Tumors that were sequenced in duplicates contained both concordant and discordant variants, but pathogenic mutations were concordant in all cases. Patient 6 was the only exception. Two synchronous tumors were sequenced from this patient and no mutations were concordant between the two tumors.

Mutations detected in tumor DNA were compared to those detected in the matching cfDNA (Figure 1, Table S3). Tumor mutations were immediately called by the bioinformatic workflow in two patients. In one of these patients, all four tumor mutations were detected in cfDNA with a mean AF of 8.6% (range 5.5-12.2%) These mutations were detected in 17 different UMI families on average (range 13-24). In the other patient, nine out of ten tumor mutations were detected in cfDNA with a mean AF of 5.4% (range 2.6-12.2%). The mutations were detected in 8 UMI families on average (range 3-13). Variants were called by the bioinformatic workflow in the remaining 28 patients, but these were not confirmed in the matching tumor DNA.

Next, we manually investigated if tumor mutations were present in cfDNA but not called by the bioinformatic workflow. Through this approach, we identified tumor mutations in cfDNA from another 14 patients. The median AF of tumor mutations in cfDNA was 0.9% (range 0.26-15.2%). One example was an L858R mutation in *Epidermal Growth Factor Receptor* (*EGFR*) that was detected in tumor DNA from patient 12. It was detected in two UMI families in the matching cfDNA (AF 1.0%). One family was made from two reads, the other was a singleton UMI. In both cases, the mutation was only detected in read 1 because it was located in a non-overlapping area in the DNA fragment. Another example was a mutation in *Signal Transducer And Activator Of Transcription 3* (*STAT3*) that was detected in tumor DNA from patient 2. The same mutation was detected in four UMI families in the matching cfDNA (AF 2.2%). These UMI families were made from total 11 reads.

ctDNA and patient characteristics

In total, concordant mutations were detected in 15/29 patients (52%). Eight patients (53%) had squamous cell carcinoma, five (33%) had adenocarcinoma, one (6%) had NSCLC-NOS and one (6%)

had adenosquamous carcinoma. Necrosis was observed in tumor from nine out of 14 patients (64%) that had evaluable tumor tissue and concordant mutations, and in four out of 14 patients (29%) with no concordant mutations. Concordant mutations were detected in 43% of patients with stage I-II disease and in 67% of patients with stage III-IV disease. In patient 6 with two synchronous tumors, a mutation from one tumor was identified in cfDNA by manual inspection. The mean haploid genome coverage in this patient was 116x.

Discussion

In this study of 30 NSCLC-patients, we sequenced matched tumor DNA and plasma cfDNA and compared mutation profiles. At least one concordant mutation was detected in 15/29 patients (52%) with mutations in tumor DNA. In 2/15 patients, tumor mutations were immediately called by the bioinformatic workflow in cfDNA. These cases were also different in that virtually all tumor mutations were detected in cfDNA. In 13/15 patients, concordant mutations were present in cfDNA, but were not called by the bioinformatic workflow. These mutations could be identified because they were detected in the matching tumor DNA. The remaining 14 patients (47%) had mutations in their tumor DNA, but these mutations were not detected in the matching cfDNA.

The result demonstrates the benefit of using a large gene panel for sequencing tumor DNA from NSCLC patients. In a previous study we used a 26-gene panel and detected mutations in only a subset of tumors from NSCLC patients *(unpublished data)*, and studies that used panels of <60 genes detected mutations in 63-78% of patients with early stage NSCLC.[3,13–16] In this study, we detected at least one tumor tissue mutation in 97% of patients with the 275-gene panel. Similarly, a recent study detected mutations in 94% of NSCLC tumor samples with a 546-gene panel.[17]

We observed mutation concordance in tumor DNA and cfDNA in 52% of patients. This is consistent with previous research. Studies on stage I-III NSCLC reported concordance in 33-50% of patients,[3,5,13–17] though, Chaudhuri *et al.*[17] detected concordant mutations in 100% of patients with available tumor tissue using their Cancer Personalized Profiling by deep Sequencing (CAPP-Seq)

assay. A possible explanation is that all mutations, not only known pathogenic or driver mutations, were included in the study.

Most patients with concordant mutations in our study had squamous cell carcinoma histology. The study by Abbosh *et al.*[5] also found that non-adenocarcinoma histology was an independent predictor of mutation detection in cfDNA. These tumors are more necrotic and thereby release more DNA.[1,18] In line with this reasoning, we found that 69% of evaluable squamous cell carcinomas were necrotic compared to 20% of adenocarcinomas.

Although the large gene panel ensured mutation detection in most patients, applying large panels for cfDNA sequencing is challenging for the bioinformatic analysis. The main issue is the low fraction of circulating tumor DNA in cfDNA, and high haploid genome coverage is necessary to detect the variants with low AF. The average haploid genome coverage on most cfDNA samples in our study was not high (mean 207x). It is possible that size selection of the cfDNA NGS libraries excluded fragments from the tumor since circulating tumor DNA is shorter than other cfDNA.[19] Tumor mutations were called only in the two cfDNA samples with relatively high fraction of tumor DNA. The median AF of the called tumor mutations was 5.5%, while the median AF of the mutations identified with help from the matching tumor DNA was 0.9%.

Variants were called in all cfDNA samples, but only in two patients the variants were confirmed in the matching tumor tissue. In the remaining samples, the called variants could not be confirmed in the matching tumor tissue. The origin of these variants is unknown. The same observation has been reported in other studies.[13,20] Although such variants have been attributed to tumor heterogeneity, a recent study found that most discordant mutations were technical artifacts.[21] Some suggest that such variants originate from clonal hematopoiesis of indeterminate potential.[22]

In this study, no AF cut-off was applied in the bioinformatic filtering, but rather a criterium that the mutation must be present in ≥3 UMI families. This is stringent and a dynamic AF cut-off could be applied. However, finding the optimal cut-off is challenging, especially in data sets from large gene panels. Low AF cut-off setting generates big data sets with too many variants that makes it difficult to identify true tumor mutations. Suboptimal sequencing data with low coverage is even more difficult to

interpret. High AF cut-off setting will miss calling of true tumor mutations and thus decrease the sensitivity of detection.

The low number of duplicates in each UMI family was a drawback in this study. Tumor mutations in cfDNA in most patients did not pass the quality filters with the set cut-off for number of duplicates (≥3) but were identified directly in the UMI sequences because the true tumor mutation profile was known. Therefore, through this study we showed that it was still possible to identify true tumor mutations in cfDNA by using the mutation information of matching tumor DNA in spite of suboptimal cfDNA samples and sequencing data.

Several factors can result in suboptimal cfDNA sequencing data, such as access to limited plasma quantity and variable plasma quality. This directly affects the amount and quality of input DNA for NGS library preparation, which further affects the NGS library quality. Low sequencing coverage may also lead to suboptimal data but increasing sequencing coverage may not always be an option due to limited resources.

In conclusion, we were able to identify tumor mutations in cfDNA from NSCLC patients using a commercially available, comprehensive NGS gene panel and bioinformatic workflow. We also show that it is possible to obtain mutation information from suboptimal cfDNA sequencing data by sequencing tumor DNA and PBL DNA from the same patients.

Acknowledgement

The library preparation and sequencing were provided by the Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU). GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and The Central Norway Regional Health Authority. This study was supported, in part, by The Central Norway Regional Health Authority and the Joint Research Committee between St. Olav's hospital and the Faculty of Medicine and Health Sciences, NTNU.

F.S, H.D and B.H.G designed and planned the study. S.G.F.W evaluated all histological material. A.L.O optimized experimental protocol and performed experiments together with GCF.

A.L.O analyzed data and interpreted data with B.H.G, S.G.F.W and H.D. A.L.O wrote the manuscript and incorporated feedback from all authors. A.L.O is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

This work was supported by The Central Norway Regional Health Authority and the Joint Research Committee between St. Olav's hospital and the Faculty of Medicine and Health Sciences, NTNU.

References

1. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clin Chim Acta Int J Clin Chem, 2001, 313:139–42
2. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, Brem H, Cameron JL, Lee CC, Fecher LA, Gallia GL, Gibbs P, Le D, Giuntoli RL, Goggins M, Hogarty MD, Holdhoff M, Hong S-M, Jiao Y, Juhl HH, Kim JJ, Siravegna G, Laheru DA, Lauricella C, Lim M, Lipson EJ, Marie SKN, Netto GJ, et al. Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. Sci Transl Med, 2014, 6:224ra24-224ra24
3. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, Anagnostou V, Fiksel J, Cristiano S, Papp E, Speir S, Reinert T, Orntoft MBW, Woodward BD, Murphy D, Parpart-Li S, Riley D, Nesselbush M, Sengamalay N, Georgiadis A, Li QK, Madsen MR, Mortensen FV, Huiskens J, Punt C, van Grieken N, Fijneman R, Meijer G, Husain H, Scharpf RB, Diaz LA, Jones S, Angiuoli S, Ørntoft T, Nielsen HJ, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med, 2017, 9:eaan2415
4. Ono A, Isaka M, Serizawa M, Omata K, Kojima H, Nakashima K, Omori S, Wakuda K, Kenmotsu H, Naito T, Murakami H, Urakami K, Nagashima T, Sugino T, Kusuhara M, Takahashi T, Yamaguchi K, Ohde Y. Genetic alterations of driver genes as independent prognostic factors for disease-free survival in patients with resected non-small cell lung cancer. Lung Cancer, 2019, 128:152–7
5. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, Le Quesne J, Moore DA, Veeriah S, Rosenthal R, Marafioti T, Kirkizlar E, Watkins TBK, McGranahan N, Ward S, Martinson L, Riley J, Fraioli F, Al Bakir M, Grönroos E, Zambrana F, Endozo R, Bi WL, Fennessy FM, Sponer N, Johnson D, Laycock J, Shafi S, Czyzewska-Khan J, Rowan A, Chambers T, Matthews N, Turajlic S, Hiley C, Lee SM, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature, 2017, 545:446–51
6. Chaudhuri AA, Lovejoy AF, Chabon JJ, Newman A, Stehr H, Say C, Aggarwal S, Carter JN, West RB, Neal JW, Wakelee HA, Loo BW, Alizadeh A, Diehn M. CAPP-Seq Circulating Tumor DNA Analysis for Early Detection of Tumor Progression After Definitive Radiation Therapy for Lung Cancer. Int J Radiat Oncol, 2016, 96:S41–2
7. Travis WD, Brambilla E, Burke A, Marx A, Nicholson A, editors. WHO classification of tumours of lung, pleura, thymus and heart. 4. ed. Lyon, France, International Agency for Research on Cancer, 2015
8. Brierley J, Gospodarowicz MK, Wittekind C, editors. TNM classification of malignant tumours. 8th ed. Chichester, West Sussex, UK ; Hoboken, NJ, John Wiley & Sons, Inc, 2017
9. Dayem Ullah AZ, Oscanoa J, Wang J, Nagano A, Lemoine NR, Chelala C. SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. Nucleic Acids Res, 2018, 46:W109–13

10. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ, Forbes SA. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res, 2019, 47:D941–7
11. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc, 2009, 4:1073–81
12. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods, 2010, 7:248–9
13. Guo Q, Wang J, Xiao J, Wang L, Hu X, Yu W, Song G, Lou J, Chen J. Heterogeneous mutation pattern in tumor tissue and circulating tumor DNA warrants parallel NGS panel testing. Mol Cancer, 2018, 17:131
14. Chen K, Zhang J, Guan T, Yang F, Lou F, Chen W, Zhao M, Zhang J, Chen S, Wang J. Comparison of plasma to tissue DNA mutations in surgical patients with non–small cell lung cancer. J Thorac Cardiovasc Surg, 2017, 154:1123-1131.e2
15. Chen K-Z, Lou F, Yang F, Zhang J-B, Ye H, Chen W, Guan T, Zhao M-Y, Su X-X, Shi R, Jones L, Huang XF, Chen S-Y, Wang J. Circulating Tumor DNA Detection in Early-Stage Non-Small Cell Lung Cancer Patients by Targeted Sequencing. Sci Rep, 2016, 6
16. Guo N, Lou F, Ma Y, Li J, Yang B, Chen W, Ye H, Zhang J-B, Zhao M-Y, Wu W-J, Shi R, Jones L, Chen KS, Huang XF, Chen S-Y, Liu Y. Circulating tumor DNA detection in lung cancer patients before and after surgery. Sci Rep, 2016, 6
17. Zhang B, Niu X, Zhang Q, Wang C, Liu B, Yue D, Li C, Giaccone G, Li S, Gao L, Zhang H, Wang J, Yang H, Wu R, Ni P, Wang C, Ye M, Liu W. Circulating tumor DNA detection is correlated to histologic types in patients with early-stage non-small-cell lung cancer. Lung Cancer, 2019. https://doi.org/10.1016/j.lungcan.2019.05.034
18. Caruso R, Parisi A, Bonanno A, Papalo D, Quattrocchi E, Branca G, Scardigno M, Fedele F. Histologic coagulative tumour necrosis as a prognostic indicator of aggressiveness in renal, lung, thyroid and colorectal carcinomas: A brief review. Oncol Lett, 2012, 3:16–8
19. Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F, Gongora C, Thierry AR. High Fragmentation Characterizes Tumour-Derived Circulating DNA. PLoS ONE, 2011, 6:e23418
20. Jamal-Hanjani M, Wilson GA, Horswell S, Mitter R, Sakarya O, Constantin T, Salari R, Kirkizlar E, Sigurjonsson S, Pelham R, Kareht S, Zimmermann B, Swanton C. Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer. Ann Oncol Off J Eur Soc Med Oncol, 2016, 27:862–7
21. Stetson D, Ahmed A, Xu X, Nuttall BRB, Lubinski TJ, Johnson JH, Barrett JC, Dougherty BA. Orthogonal Comparison of Four Plasma NGS Tests With Tumor Suggests Technical Factors are a Major Source of Assay Discordance. JCO Precis Oncol, 2019:1–9
22. Hu Y, Ulrich BC, Supplee J, Kuang Y, Lizotte PH, Feeney NB, Guibert NM, Awad MM, Wong K-K, Jänne PA, Paweletz CP, Oxnard GR. False-Positive Plasma Genotyping Due to Clonal Hematopoiesis. Clin Cancer Res, 2018, 24:4437–43

Figure legends

Figure 1, Summary of patient characteristics and concordance in tumor DNA and matched plasma cell-free DNA (cfDNA). Matched tumor DNA, peripheral blood leukocyte DNA and cfDNA were sequenced from 30 patients with non-small cell lung cancer. Patients were divided into three categories. Cases with concordant mutations that were called by the bioinformatic workflow in cfDNA were classified in the first category (n=2). Cases where concordant mutations were detectable in cfDNA only with prior knowledge from tumor DNA where classified in the second category (n=13). Patients with no detected concordant mutations were classified in the last category (n=14). One patient had no detectable mutations in tumor (Patient 1). *Patient 6 had two synchronous tumors that were sequenced separately. cfDNA; cell-free DNA, NSCLC-NOS; non-small cell lung cancer - not otherwise specified.

Table 1. Characteristic of the patients analyzed in the study

Characteristics	Number of patients (%)
In total	30
Age, years:	
Mean: 70	
Range: 57-81	
Sex:	
Female	11 (37)
Male	19 (63)
Smoking history:	
Smoker/former smoker	28 (93)
Never smoker	2 (7)
NSCLC subtype:	
Adenocarcinoma	15 (50)
Squamous cell carcinoma	13 (43)
Adenosquamous carcinoma	1 (3)
NSCLC-not otherwise specified	1 (3)

NSCLC; non-small cell lung cancer

Highlights

- ctDNA sequencing was performed with a large commercially available gene panel
- Detection rate of ctDNA was increased by using mutation profile of tumor tissue
- This approach is useful in case of suboptimal plasma DNA quality and quantity