DETECTION OF MUTATIONS AND FUSIONS IN LUNG ADENOCARCINOMA USING ION TORRENT SEQUENCING TECHNOLOGY

Valentina STRATAN, Valeri TUTUIANU, Victor SITNIC, Corneliu PREPELITA, Cristina POPA, Valeriu BILBA, Sergiu BRENISTER

Oncological Institute, Chisinau, the Republic of Moldova

Corresponding author: Cristina Popa, e-mail: cristina.popa18@yahoo.com

Introduction. Adenocarcinomas are the most common lung tumors and are often diagnosed in advanced stages when the tumor has a polyclonal form with a wide variety of genetic alterations and activated mutational processes. Comprehensive analysis of mutational status from FFPE tissue samples in such patients can provide a therapeutic perspective and contribute considerably to clinical decisions thereby increasing the overall survival rate. *Material and methods.* 22 genes were analyzed for mutations and 4 genes for fusion transcripts using two Ion AmpliSeq panels. DNA was isolated from paraffin-embedded tissue sections while RNA analysis was performed using two types of samples: paraffin blocks and fresh tumor tissue. Key variant detection and data analysis was performed using next platforms: Ion Reporter, ONCOMINE, R language. *Results.* The study of 30 tumor samples allowed the detection of 147 mutations and 4 fusions in 19 genes, and the therapeutically actionable variants were associated with different drugs clinically approved or in the phase of clinical trials. The most genetic variants were identified in the TP53, EGFR and NOTCH1 genes with a prevalence of over 50% in the TP53 gene, while all 4 detected fusions (one fusion per sample) represent the association of the ALK gene with other partners: EML4(13)-ALK(20) – present in 2 samples; EML4(6)-ALK(20); and an ALK fusion with an unknown partner gene. *Conclusions.* Analyzing the mutational status of tumor samples from patients with lung adenocarcinoma it has been ascertained the therapeutic utility of gene panel sequencing covering point mutations, INDELs and SNVs, as well as gene fusions, using FFPE tissue as primary material. This is valid for both targeted monotherapies and combined therapies.

DETECȚIA MUTAȚIILOR ȘI A FUZIUNILOR ÎN ADENOCARCINOMUL PULMONAR FOLOSIND TEHNOLOGIA DE SECVENȚIERE *ION TORRENT*

Introducere. Adenocarcinoamele sunt cel mai frecvent atestate tumori pulmonare și adesea sunt diagnosticate în stadiile avansate, când tumoarea are formă policlonală, cu o varietate mare de alterații genetice și procese mutaționale activate. Analiza comprehensivă a statutului mutațional din probe de țesut FFPE la astfel de pacienți poate oferi o perspectivă terapeutică și contribuie considerabil la luarea deciziilor clinice, astfel determinându-se creșterea ratei de supraviețuire globală. *Material și metode.* Au fost analizate 22 gene pentru determinarea mutațiilor și 4 gene pentru identificarea transcripților de fuziune în baza a două paneluri Ion AmpliSeq. ADN-ul a fost izolat din secțiuni de țesut parafinat, în timp ce analiza ARN-ului a fost realizată utilizându-se două tipuri de probe: blocuri de parafină și țesut tumoral proaspăt. Detectarea variantelor cheie și analiza datelor a fost efectuată cu ajutorul platformelor Ion Reporter, ONCOMINE și limbajul R. *Rezultate.* Studierea a 30 de probe tumorale a permis detectarea a 147 mutații și 4 fuziuni în 19 gene, iar variantele care pot fi acționate terapeutic au fost asociate cu diferiți agenți medicamentoși aprobați clinic sau aflați în faza de studii clinice. Cele mai multe variante genetice au fost identificate în genele TP53, EGFR și NOTCH1 cu o prevalență de peste 50% în gena TP53, iar cele 4 fuziuni (câte o fuziune per probă) reprezintă în totalitate asocierea genei ALK cu alți parteneri: EML4(13)-ALK(20) – prezentă în 2 probe; EML4(6)-ALK(20); și o fuziune ALK cu o genă partener necunoscută. *Concluzii.* Analizând statutul mutațional al probelor tumorale de la pacienții cu adenocarcinom pulmonar, s-a constatat utilitatea secvențierii panelurilor de gene care acoperă atât mutațiile punctiforme, INDEL-urile (inserții și deleții mici) și SNV-urilor (variațiile unui nucleotid), cât și a fuziunilor genice, cu utilizarea ca material primar a țesutului FFPE, pentru selectarea celor mai potrivite terapii atât din perspectiva aplicării monoterapiilor țintite, cât și a terapiilor combinate.

INTRODUCTION

Although lung adenocarcinomas (LUAD) are the most common lung tumors, they are often diagnosed late when local tissue invasion is well advanced and metastasis is present. Genetic alterations such as somatic mutations or gene fusions are often associated with the development of these tumors and molecular profiling allows to approach the correct treatment strategy. According to The Cancer Genome Atlas (1), patients with this type of tumors show mutations in the tumor suppressor TP53 overlapping with other oncogenic driver alterations such as KRAS, EGFR, BRAF, ERBB2 mutations or ALK, RET, ROS1 gene fusions all of which have potential therapeutic implications. The complete genetic profile of LUAD is not simple to establish in routine clinical practice since this often involves the use of invasive techniques, the availability of a well-equipped molecular biology laboratory with qualified personnel and, last but not least, involves high costs when performing the analyses. Although a few years ago experts in the field recommended only EGFR testing in advanced cancers for this type of tumor, more recent recommendations include testing for EML4-ALK fusions. However, advances in molecular profiling suggest the need to increase the number of molecular targets that need to be tested in any stage of lung adenocarcinoma (2). This will provide a more comprehensive genetic picture of the tumor and, respectively, a more effective treatment strategy.

Aim: in our study we proposed the use of Ion Torrent sequencing technology for the detection of mutations and fusions in the tumors of patients with lung adenocarcinomas, and for fusion transcripts – two types of RNA: the first, isolated from FFPE samples (Formalin-fixed, paraffin-embedded), and the second – from fresh tumor tissue. The research was carried out in the Cancer Biology Scientific Laboratory within the Oncological Institute of Moldova during the years 2021-2022.

MATERIAL AND METHODS

The study analyzed 30 tumors taken from 30 patients with the histological diagnosis of lung adenocarcinoma with a content of more than 20% tumor cells per sample. DNA was isolated from paraffin-embedded tumor tissue sections (3 10-μm sections), and RNA from both paraffin-embedded sections (4 15-μm sections) and fresh tissue. Purification of nucleic acids was performed using the *RecoverAll Total Nucleic Acid Isolation Kit (Invitrogen, ThermoFisher Scientific)*. The *Ion AmpliSeq Colon and Lung Cancer Research Panel v2* was used for mutation detection, which includes 22 genes (KRAS, EGFR, BRAF, PIK3CA, AKT1, ERBB2, PTEN, NRAS, STK11, MAP2K1, ALK, DDR2, CTNNB1, MET, TP53, SMAD4, FBX7, FGFR3, NOTCH1, ERBB4, FGFR1, FGFR2) and for the identification of fusion transcripts – the *Ion AmpliSeq RNA Fusion Lung Cancer Research Panel*, which allows the determination of more than 70 transcripts in the ALK, RET and ROS1 genes in only 1% of RNA used. The performance of the genomic investigations was evaluated by assessing the number of reads, the average coverage of the bases (Coverage) in the regions of interest, as well as the imbalance of the expression of the 3'/5' amplicons in the case of translocations. For all samples, a maximum performance was obtained, the number of reads exceeding 100,000 and the average base coverage being over 500x. To determine the fusions, the 3'/5' imbalance thresholds recommended by *Ion Reporter* (3, 4) were respected. The average charge density of the chip as well as the ISP (Ion Sphere Particles) details of the experiment are shown in Figure 1.

Detection of key variants and data analysis was performed using Ion Reporter platforms (AmpliSeq Colon and Lung Cancer v2 single sample and AmpliSeq RNA Lung Fusion – w1.2 – Single Sample workflows), ONCOMINE (5), R language (6, 7).

RESULTS

Following the isolation of the RNA samples, a mean concentration of 20.92 ng/μL (non-normal distribution, SD=13.73; Median=15.90; IQR=12.80) was obtained for the paraffin-embedded tissue samples and a mean 11.17 ng/μL (non-normal distribution, SD=7.12; Median=10.20; IQR=4.64) for fresh tissue. The mean DNA concentration was 7.57 ng/μL (non-normal distribution, SD=7.26; Median=4.40; IQR=8.38). Regarding the average read length, this was 69 nucleotides (non-normal distribution, SD=15.95; Median=72; IQR=29) for the RNA extracted from FFPE sections, 111 nucleotides (non-normal distribution normal, SD=25.54; Median=102;

IQR=36) for RNA from fresh tissue and 117 for DNA (non-normal distribution (fig. 2), SD=4.11;
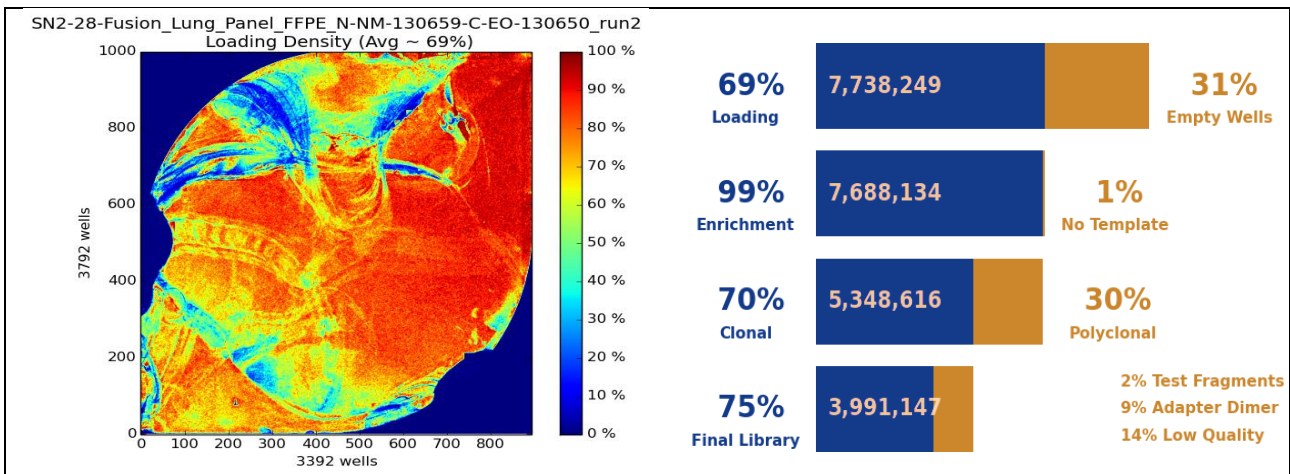
Median=118; IQR= 2.75).

Figure 1. Details of the sequencing experiment (left – average chip load density, right – ISP details).
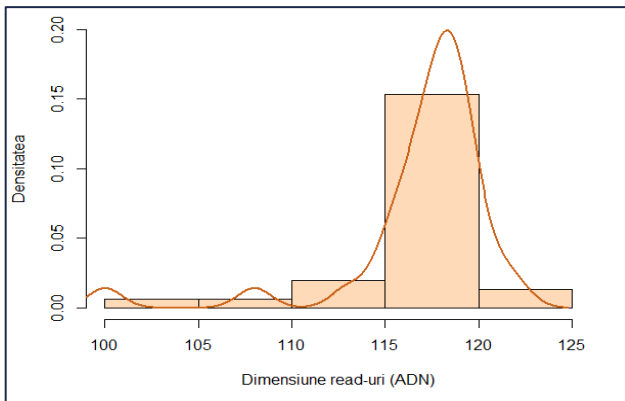

Figure 2. Data distribution with reference to size of reads (DNA).

Thus, 4 samples were identified with a 3'/5' imbalance score indicating the presence of 4 fusions (one fusion per sample): EML4(13)-AK(20) – present in 2 samples; EML4(6)-ALK(20); and an ALK fusion with an unknown gene. Only one sample had an equivocal score for RET gene translocation. Regarding DNA sequencing, of the 22 genes included in the panel, mutations were detected in 18 (EGFR, TP53, NOTCH1, CTNNB1, MET, BRAF, KRAS, STK11, DDR2, SMAD4, NRAS, ERBB4, PTEN, AKT1, PIK3CA, FGFR3, FGFR2, FGFR1) and in 13 (EGFR, TP53, NOTCH1, CTNNB1, MET, BRAF, KRAS, STK11, ERBB4, PTEN, AKT1, PIK3CA, FGFR3) – the presence of so-called key variants. Key variants in cancer are biologically relevant mutations that induce cell proliferation. A summary of the identified mutations is presented in Table 1.

Table 1. Summary of variants identified in samples following DNA sequencing.

The total number of genetic variants	147
Types of variants detected	91
Key variants	77
Non-key variants	14
Maximum number of variants detected per sample	28
Minimum number of variants detected per sample	2
Average number of variants detected per sample	4.9

Most variants were identified in the TP53 gene with 51% of the total number of mutations detected. The EGFR gene follows with 17.7% and NOTCH1 with 7.5% (fig. 3).

The top of the key variants by the number of samples in which they appear is shown in Figure 4.

The *NOTCH1 V1578del* deletion that is the most frequent key variant, according to the ONCOMINE platform, is not associated with relevant therapies both in lung adenocarcinoma and in other types of cancer, however, from the results obtained, a major frequency of this mutation is observed with tumors positive for ALK fusions, *NOTCH1 V1578del* being present in 3 of 4 ALK positive samples.
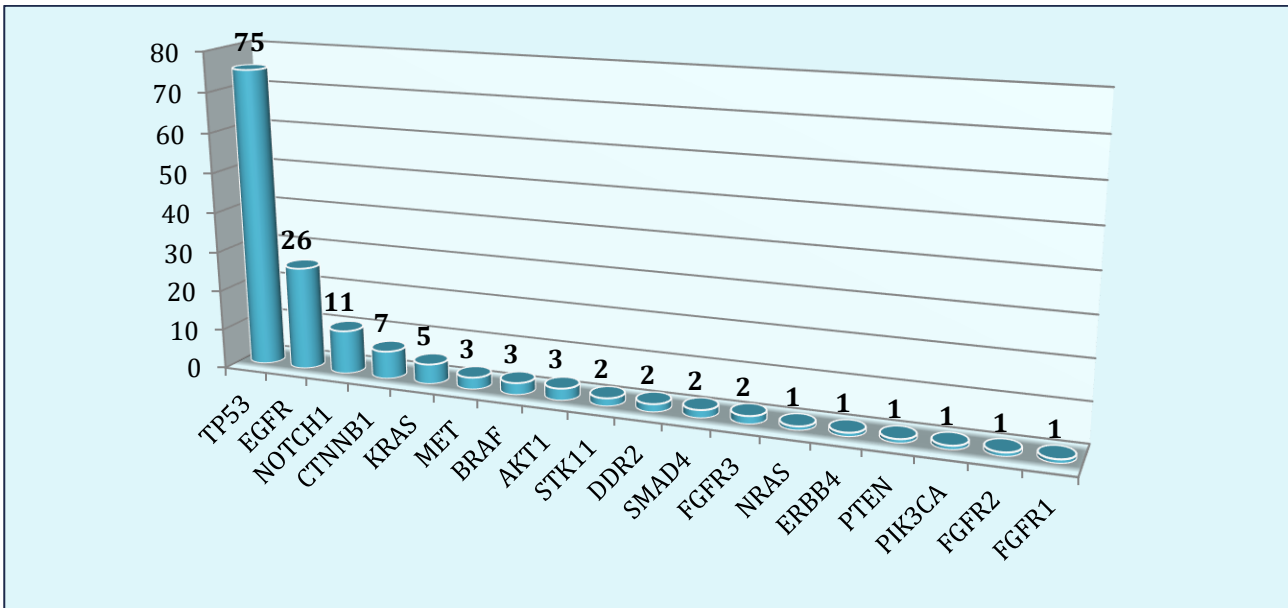
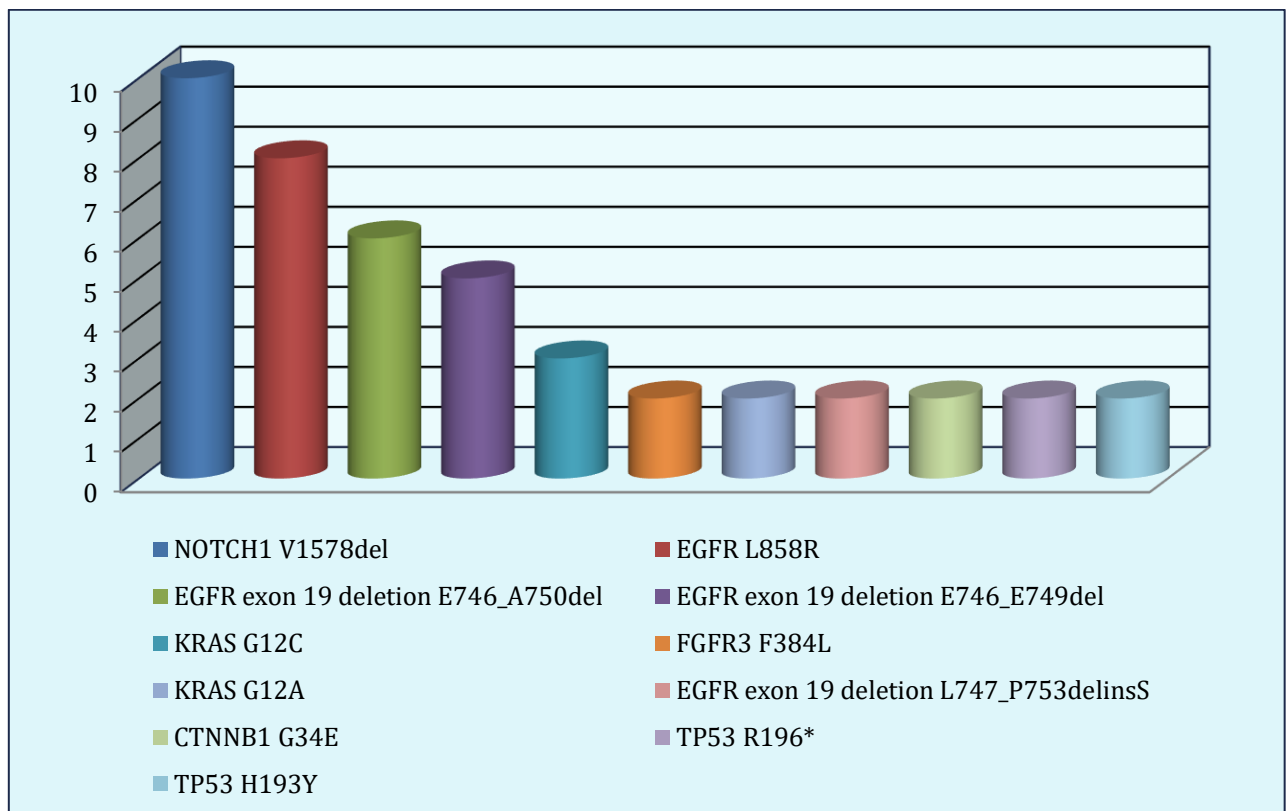Figure 3. Number of mutations identified per gene.

- NOTCH1 V1578del
- EGFR L858R
- EGFR exon 19 deletion E746_A750del
- EGFR exon 19 deletion E746_E749del
- KRAS G12C
- FGFR3 F384L
- KRAS G12A
- EGFR exon 19 deletion L747_P753delinsS
- CTNNB1 G34E
- TP53 R196*
- TP53 H193Y

Figure 4. Top of key variants by number of samples.

DISCUSSIONS

The shorter length of reading units in the case of RNA from paraffin-embedded tissue is explained by its molecular degradation during the process of obtaining paraffin blocks. However, following sequencing, reproducibility of results for fusion transcripts could be observed in both types of RNA samples (FFPE and fresh tissue). According to the specialized literature (8, 9) gene fusions

EML4(13)-ALK(20) and EML4(6)-ALK(20) are the most common EML4-ALK variants that appear in non-small cell lung cancer and constitute around of 75-80% of all translocation variants. Treatment options for ALK-positive lung adenocarcinomas are: Crizotinib (Xalkori or Crizalk), Ceritinib (Zykadia), Alectinib (Alecensa), Brigatinib (Alunbrig), and Lorlatinib (Lorbrena).

Regarding the *L858R* substitution and deletions in EGFR, they are associated with the following therapies: afatinib, bevacizumab + erlotinib, dacomitinib, erlotinib, erlotinib + ramucirumab, gefitinib, osimertinib, atezolizumab + bevacizumab + chemotherapy, bevacizumab+gefitinib, gefitinib + chemotherapy, Osimertinib + chemotherapy. *KRAS G12C* and *KRAS G12A* genetic variants associate with cabozantinib in other cancers, and only KRAS G12C can be targeted with sotorasib in non-small cell lung cancer (5).

The research carried out allowed the evaluation of the most important gene mutations and fusions in 30 samples of lung adenocarcinoma as well as their association with different drugs clinically approved or in the phase of clinical trials (5). Analysis of RNA from two types of samples (FFPE and fresh tissue) was performed and the same results were obtained. We can mention that Ion Torrent sequencing is a suitable technique for evaluating genetic alterations in lung adenocarcinoma and can contribute to the best clinical decisions. Advances in tumor genome sequencing and the identification of druggable molecular targets favor the selection of the most effective therapies in the treatment of cancer, however there are some cost barriers to the widespread translation of these technologies into clinical practice. In order to obtain maximum benefits from the exploration of the tumor genome, well-equipped molecular biology laboratories are needed, with qualified personnel both in terms of laboratory analysis and for the use of bioinformatics algorithms so that the molecular data can be effectively interpreted, integrated with the clinical profile and reflected in clinical reports.

CONCLUSIONS

1. Research allowed reproducibility of results for fusion transcripts in both types of RNA samples (FFPE and fresh tissue).

2. The obtained results confirm the need for clinical testing for EML4-ALK fusions in patients with lung adenocarcinoma.

3. As KRAS mutations are among the top key variants after EGFR mutations, testing for actionable genetic variants of the KRAS gene in patients with lung adenocarcinoma is of increased interest.

4. Further studies are needed to confirm or refute a possible relationship between ALK gene fusions and the NOTCH1 V1578del mutation.

CONFLICT OF INTERESTS
The authors have no conflict of interest to declare.

ETHICAL APPROVAL
This study was carried out with the approval of the National Ethical Expertise Committee of Clinical Study No. 1086/28.04.2021 (2 SOFT/1.2/207 LUNGNEX-RD) and of the Ethics/Bioethics Committee no. 14, 05-18/21/01.02.2021 of PMSI Oncological Institute, Republic of Moldova (State Program 20.80009.8007.02).

REFERENCES

1. NCI. The Cancer Genome Atlas Research Network. Available at: https://www.cancer.gov/tcga [Accessed 14.09.2022].
2. Chalela R, Curull V, Enríquez C, Pijuan L, Bellosillo B, Gea J. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J Thorac Dis*. 2017;9(7):2142-2158. doi:10.21037/jtd.2017.06.20
3. GRC. Data. Human Genome Assembly GRCh37.p5. Available at: https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh37.p5 [Accessed 11.09.2022].
4. GRC. Data. Human Genome Assembly GRCh38.p2. Available at: https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh38.p2 [Accessed 11.09.2022].
5. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004;6(1):1-6. doi:10.1016/s1476-5586(04)80047-2
6. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2021. https://www.R-project.org/
7. Raivo Kolde. pheatmap: Pretty Heatmaps. R package version 1.0.12, 2019. https://CRAN.Rproject.org/package=pheatmap
8. Sabir SR, Yeoh S, Jackson G, Bayliss R. EML4-ALK Variants: Biological and Molecular Properties, and the Implications for Patients. *Cancers (Basel)*. 2017;9(9):118. doi:10.3390/cancers9090118
9. Choi YL, Takeuchi K, Soda M, Inamura K, Togashi Y, Hatano S, et al. Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. *Cancer Res*. 2008;68(13):4971-6. doi:10.1158/0008-5472.CAN-07-6158