DNA repair genes are associated with tumor tissue differentiation and immune environment in lung adenocarcinoma: a bioinformatics analysis based on big data

Jiayin Li[1], Jingxu Zhou[1], Jing Zhang[1], Zhiwei Xiao[1], Wenping Wang[1], Hanrui Chen[1], Lizhu Lin[1], Qiuye Yang[2]

[1]Cancer Center, The First Affiliated Hospital to Guangzhou University of Chinese Medicine, Guangzhou, China; [2]Department of Medical Technologic, The First Affiliated Hospital to Guangzhou University of Chinese Medicine, Guangzhou, China

Contributions: (I) Conception and design: J Li, Q Yang; (II) Administrative support: J Zhou, J Zhang; (III) Provision of study materials or patients: Z Xiao; (IV) Collection and assembly of data: W Wang, H Chen, L Lin; (V) Data analysis and interpretation: J Li, Q Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Qiuye Yang. No. 16 Airport Road, Guangzhou, China. Email: qiuyeyang@126.com.

Background: Lung adenocarcinoma (LUAD) is the most common type of lung cancer. DNA repair genes (DRGs) is important in lung cancer. The relationship between the immune environment and the expression levels of DRGs in LUAD remains unclear. The purpose of this study is to assess the relationship between DRGs and the immune environment and clinical characteristics of LUAD.

Methods: Data of 169 LUAD cases were obtained from cbioportal. The RNA-seq data came from the The Cancer Genome Atlas (TCGA) database. We collected DRGs from the Reactom database (KW0037, Reactom.org). The 302 genes expressed in each sample were analyzed by hierarchical clustering and grouped using the Gene Cluster 3.0 program. The Java Treeview program was used to generate heat maps of cluster indications and tumor staging patterns. GraphPad Prism 8 was used to draw survival curves and compare overall survival (OS). For single genes, an OS difference analysis between low and high expression populations was performed in GraphPad Prism 8.

Results: Matrix clustering showed no difference in the prognosis of the two clusters. The comparison of subgroups showed that Subcluster 1 (SC1) had the best prognosis, and Subcluster 2 (SC2) had the worst. There was a significant difference in tumor grades between Cluster 1 and Cluster 2 (P=0.01). There were significant differences in smoking status, histological grade and adenocarcinoma subtype among subgroups. In Subcluster 3 (SC3), the proportion of poorly differentiated cases was highest. Immunological index analysis showed that there were significant differences between Cluster 1 and Cluster 2 in interferon, macrophages, monocytes, neutrophils, natural killer (NK) cells, and T cells. Tumor purity, interferon, macrophages, monocytes, neutrophils, NK cells, T cells, translation, and proliferation all showed significant differences between subgroups. In SC2, the proliferation index increased (0.082 *vs.* 0.070); the protein translation index decreased (0.134 *vs.* 0.137); and the interferon level increased (0.099 *vs.* 0.097). In SC3, the proliferation index decreased (0.076 *vs.* 0.071); the protein translation index decreased (0.140 *vs.* 0.136); and the level of neutrophils increased (0.083 *vs.* 0.086).

Conclusions: The differences of DRGs in LUAD are related to tissue differentiation and immune indicators but not to prognosis.

Keywords: DNA repair genes; lung adenocarcinoma (LUAD); tissue differentiation; prognosis

Introduction

Lung cancer is the cancer with the highest morbidity and mortality in the world (1-3). Non-small cell lung carcinoma (NSCLC) accounts for about 80% of lung cancers. Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma are the two most common histological types of NSCLC (4,5). More than half of patients cannot benefit from targeted therapy (6-8). Tumor tissue differentiation is closely related to tumor staging, treatment options and prognosis (9,10). The differentiation of tumor tissue is related to many factors, including various environmental factors and genetic factors (11,12). We know that the occurrence and development of lung cancer is closely related to smoking (13,14). Eighty percent of lung cancers are caused by smoking (15). However, amongst all smokers, only about 15% will eventually develop lung cancer, and a considerable number of people who have never smoked will also develop lung cancer, e.g., as a result of inhalation of second-hand smoke (16). These observations show that the occurrence of lung cancer is related to individual susceptibility. DNA repair genes (DRGs) are amongst the factors that play an important role in the susceptibility of different individuals to lung cancer (17-19).

In recent years, due to the development of big data technology, more and more researchers store original data in public databases, allowing researchers to further use these data for research, which has greatly promoted the development of clinical research. Recent immunotherapy has provided new ideas for the treatment of lung cancer. The immune environment is an important factor in determining the treatment effect (20). The relationship between the immune environment and the expression levels of DEGs in LUAD is still unclear. This study is based on the clinical data of the East Asian LUAD RNA-seq data published in Nature Genetics in 2020, which analyzed the expression and distribution of DNA repair-related genes in LUAD and studied the relationship between DEGs and LUAD immune environment clinical features. We present the following article in accordance with the REMARK reporting checklist (available at https://dx.doi.org/10.21037/jtd-21-949).

Methods

Samples

According to a previously described approach, data from 169 LUAD cases were obtained, and the relevant data were downloaded from The Cancer Genome Atlas (TCGA) database platform cbioportal, including RNA sequencing data and the corresponding clinical records relating to these patients (21,22). All these patients were pathologically diagnosed as LUAD. The RNA-seq data came from the East Asian LUAD data included in the TCGA database and were analyzed using the cbioportal platform. This research was conducted in accordance with the "Declaration of Helsinki" (as revised in 2013).

Gene set construction

We collected DRGs from the Reactom database (KW0037, Reactom.org); the gene set ID was: R-HSA-73894. Three hundred and two effective genes with complete sequencing and prognostic data were included in the final study sample. The corresponding expression level of each gene is shown as an mRNA z-score. According to differences in the expression of different genes, genes with similar expression trends were divided into clusters.

Bioinformatics

Hierarchical cluster analysis was performed on the 302 genes expressed in each sample, and samples with similar gene expression patterns were grouped. The gene expression differences related to DNA repair between different groups were determined from the entire data set, and the hierarchical clustering algorithm was used to group them using the Gene Cluster 3.0 program. The Java Treeview program was used to generate heat maps of cluster indications and tumor staging patterns.

Prognostic implication analysis

In order to evaluate the relationship between DNA repair-related genes and the clinical outcome of patients with LUAD, we used GraphPad Prism 8 (GraphPad Software, Inc., California, United States; Version 6.01, 2012) for Windows to plot survival curves and overall survival (OS) period for comparison. In addition, for single genes, in GraphPad Prism 8, OS difference analysis between the between low and high expression populations was carried out.

Statistical analysis

SPSS version 22.0 (IBM, Inc., USA) was used for statistical

processing. A Kaplan-Meier survival curve was drawn and tested by log-rank in GraphPad Prism version 8.0. Fisher's exact tests and Pearson correlations were used to analyze the correlations between clinical features and variables. One-way analysis of variance (ANOVA) was used to analyze the differences in gene expression between different groups. In Gene Set Enrichment Analysis (GSEA), gene sets with false discovery rates (FDR) <0.25 are regarded as significantly enriched gene sets to screen out the most meaningful scientific hypotheses and reduce the risk of false positives. P<0.05 was considered to indicate statistical significance.

Results

Subgroup analysis of DNA repair gene expression

A matrix was prepared based on the data, and there were two clusters and four subclusters after clustering and stratification (*Figure 1A*). There was no difference in the prognosis of clusters, and the difference in prognosis between clusters was compared. Subcluster 1 had the best prognosis and Subcluster 2 had the worst prognosis (SC1, SC2) (*Figure 1B*). When the survival of SC1 and SC2 with SC3 and SC4 was compared, although the curves were different, the P value was >0.05 (*Figure 1C*). Comparisons of the clinical data of Cluster 1 and Cluster 2 showed that there was a significant difference in tumor grade (P=0.01) (*Table 1*). When each subgroup was compared, significant differences in smoking status, historical grade and WHO2015 adenocarcinoma subtype were found. Amongst the subgroups, Subcluster 3 had the highest proportion of poorly differentiated cases (*Table 2*).

Differentially expressed genes (DEGs)

According to the above results, SC2 had the worst prognosis, while SC3 had the worst degree of differentiation. We further analyzed the DRGs in the worst prognosis group (SC2) and the worst differentiated group (SC3) compared with other patients. Compared with other subclusters, SC2 had 149 differentially expressed genes (DEGs), and SC3 had 213 DEGs. *Figure 2A* lists the top 50 genes with the most highly significant differences. Comparisons of the two groups of specifically expressed genes showed that there were 107 overlapping DEGs in both groups (*Figure 2B*). Protein-protein interaction (PPI) analysis of these genes indicated that the differential genes were functionally related to each other (*Figure 2C*).

One hundred and seven genes were functionally enriched. Gene Ontology (GO)-biological progress analysis showed that these genes were mainly involved in DNA metabolic processes, GO-molecule function analysis demonstrated that these genes were mainly involved in catalytic activity, and GO-cell component analysis showed that these genes were mainly distributed in the nucleoplasm. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis showed that these genes were enriched in the Fanconi anemia pathway (*Figure 2D*).

Cellular and molecular expression differences

When comparing the differences in the relative amounts of immune cells between the groups, it was found that many immune indicators were significantly different. Comparisons of Cluster 1 and Cluster 2 showed that there were significant differences in interferon, macrophages, monocytes, neutrophils, natural killer (NK) cells, T cells, and translation (*Table 3*, *Figure 3A*); comparisons of subclusters showed tumor purity, interferon, macrophages, monocytes, neutrophils, NK cells, T cells, translation, and proliferation were all significantly different (*Table 4*, *Figure 3B*).

In addition, we analyzed the indexes of histological differences in the worst-prognosis group (SC2) and the worst-differentiated group (SC3) compared with other patients (*Table 5*). It was found that the proliferation indexes and the protein translation indexes were significantly different (*Figure 3C,D*). In the worst-prognosis group (SC2), the proliferation index increased (0.082 *vs.* 0.070); the protein translation index decreased (0.134 *vs.* 0.137); and the interferon level increased (0.099 *vs.* 0.097) (*Figure 3C*). In the worst-differentiated group (SC3), the proliferation index decreased (0.076 *vs.* 0.071); the protein translation index decreased (0.140 *vs.* 0.136); and the neutrophil level increased (0.083 *vs.* 0.086) (*Figure 3D*).

Discussion

This bioinformatics analysis showed that the differences in the expression profiles of DNA repair-related genes in LUAD were related to tissue differentiation and immune indicators but not to prognosis.

Clinical studies have found that different individuals with the same clinical characteristics show great heterogeneity in clinical manifestations, treatment sensitivity, disease recurrence, and survival outcomes. We have previously reported on individual differences in genetic susceptibility,
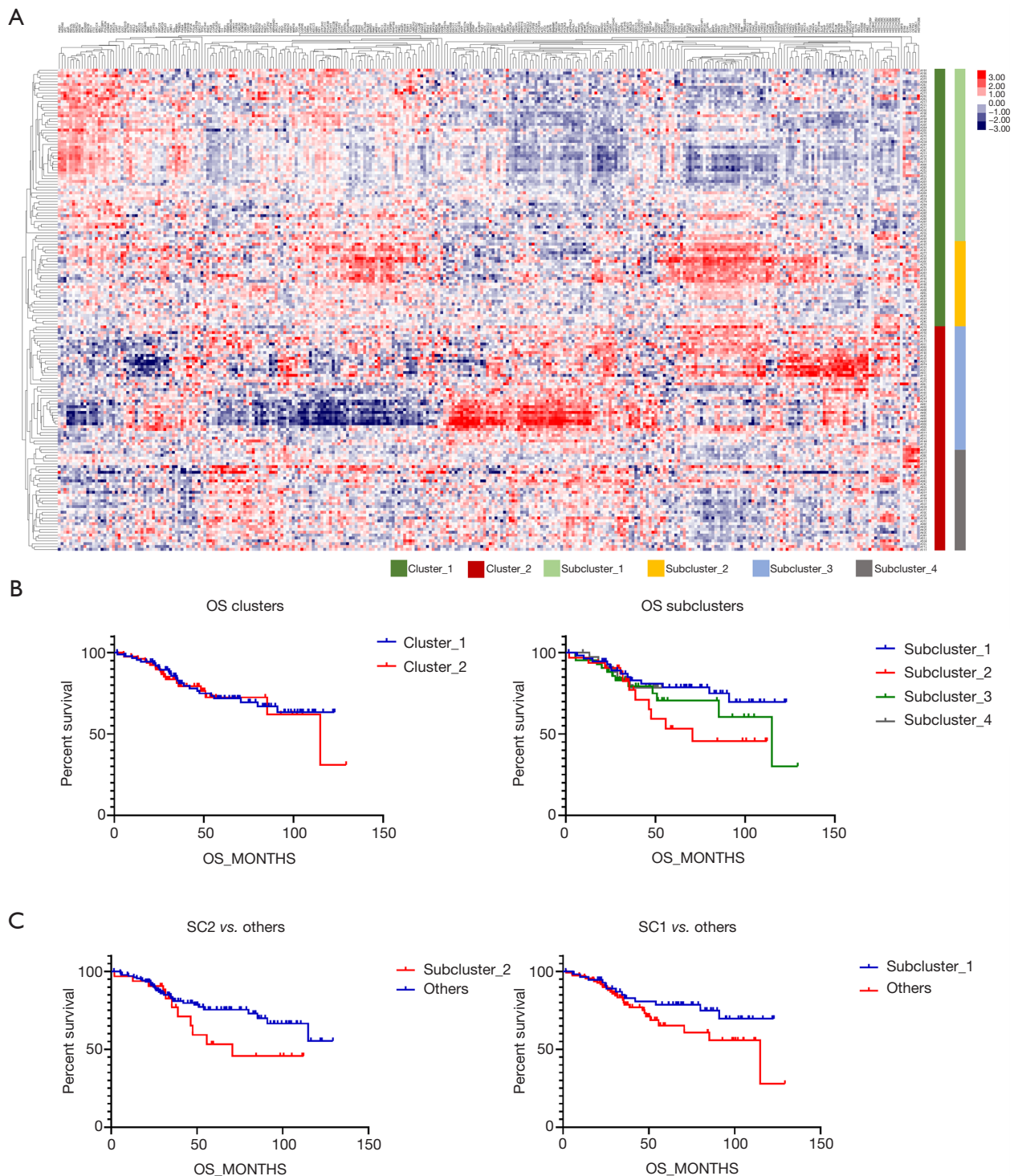
Figure 1 Clusters analysis and subclusters analysis of DRGs. (A) In total, 169 patients were primarily divided into two clusters with two subclusters respectively. The expression values of 302 effective genes corresponding to the individual patient were arrayed in the columns according to the expression affinity. Patients with similar gene expression patterns were clustered and grouped using a hierarchical clustering algorithm and arrayed in rows. (B) For overall survival, there were no significant differences between Cluster 1 and 2, but such differences existed between subclusters. (C) Subcluster 1 had the best prognosis, while Subcluster 2 had the worst prognosis. DRGs, DNA repair genes.

Table 1 Clinical characteristics of each primary cluster

Characteristic	Subgroup	Cluster 1 (n=90)	Cluster 2 (n=79)	P value
Age (years)	Mean	63.32	64.77	0.32
	SD	9.89	8.97	
Sex (n)	Female	51	43	0.45
	Male	39	36	
Smoking status (n)	No	53	55	0.10
	Yes	37	24	
Stage (n)	I	58	44	0.47
	II	17	13	
	III	13	18	
	IV	1	3	
	NA	1	1	
Histological grade (n)	Well differentiated	10	4	0.01
	Well to moderately differentiated	2	2	
	Moderately differentiated	53	56	
	Moderately to poorly differentiated	2	2	
	Poorly differentiated	7	13	
	NA	16	2	
Adenocarcinoma subtype WHO 2015 (n)	Acinar adenocarcinoma (85513)	42	45	0.16
	Acinar adenocarcinoma with mucin production	1	0	
	Adenocarcinoma, mixed subtypes	4	1	
	Adenocarcinoma, NOS (81403)	17	5	
	Invasive mucinous adenocarcinoma	1	0	
	Invasive mucinous adenocarcinoma (82533)	2	1	
	Lepidic adenocarcinoma (82503)	3	5	
	Micropapillary adenocarcinoma (82653)	2	1	
	Minimally invasive adenocarcinoma, non-mucinous (82502)	0	1	
	Papillary adenocarcinoma (82603)	11	8	
	Solid adenocarcinoma (82303)	6	12	
	NA	1	0	
Tyrosine kinase inhibitor treatment (n)	No	69	59	0.45
	Yes	21	20	
Chemotherapy (n)	No	64	52	0.28
	Yes	26	27	
Smoking pack years (n)	N	37	24	0.40
	Mean	33.30	41.53	
	SD	38.59	34.97	

Table 2 Clinical characteristics of each secondary cluster

| Characteristic	Subgroup	Subcluster 1 (n=58)	Subcluster 2 (n=32)	Subcluster 3 (n=42)	Subcluster 4 (n=37)	P value
Age (years)	Mean	63.98	62.13	64.74	64.81	0.62
	SD	8.85	11.61	10.16	7.54	
Sex (n)	Female	35	16	18`	25	0.12
	Male	23	16	24	12	
Smoking status (n)	No	38	15	25	30	0.03
	Yes	20	17	17	7	
Stage (n)	I	37	21	21	23	0.70
	II	12	5	9	4	
	III	8	5	9	9	
	IV	1	0	2	1	
	NA	0	1	1	0	
Histological grade (n)	Well differentiated	8	2	2	2	0.00
	Well to moderately differentiated	2	0	1	1	
	Moderately differentiated	32	21	26	30	
	Moderately to poorly differentiated	0	2	1	1	
	Poorly differentiated	3	4	11	2	
	NA	13	3	1	1	
Adenocarcinoma subtype WHO 2015 (n)	Acinar adenocarcinoma (85513)	29	13	23	22	0.05
	Acinar adenocarcinoma with mucin production	0	1	0	0	
	Adenocarcinoma, mixed subtypes	4	0	1	0	
	Adenocarcinoma, NOS (81403)	8	9	4	1	
	Invasive mucinous adenocarcinoma	0	1	0	0	
	Invasive mucinous adenocarcinoma (82533)	2	0	0	1	
	Lepidic adenocarcinoma (82503)	2	1	2	3	
	Micropapillary adenocarcinoma (82653)	0	2	0	1	
	Minimally invasive adenocarcinoma, non-mucinous (82502)	0	0	0	1	
	Papillary adenocarcinoma (82603)	9	2	3	5	
	Solid adenocarcinoma (82303)	3	3	9	3	
	NA	1	0	0	0	
Tyrosine kinase inhibitor treatment (n)	No	45	24	29	30	0.63
	Yes	13	8	13	7	
Chemotherapy (n)	No	41	23	24	28	0.30
	Yes	17	9	18	9	
Smoking pack years (n)	N	20	17	17	7	0.80
	Mean	33.74	32.77	44.32	34.75	
	SD	49.60	20.71	36.65	32.10	
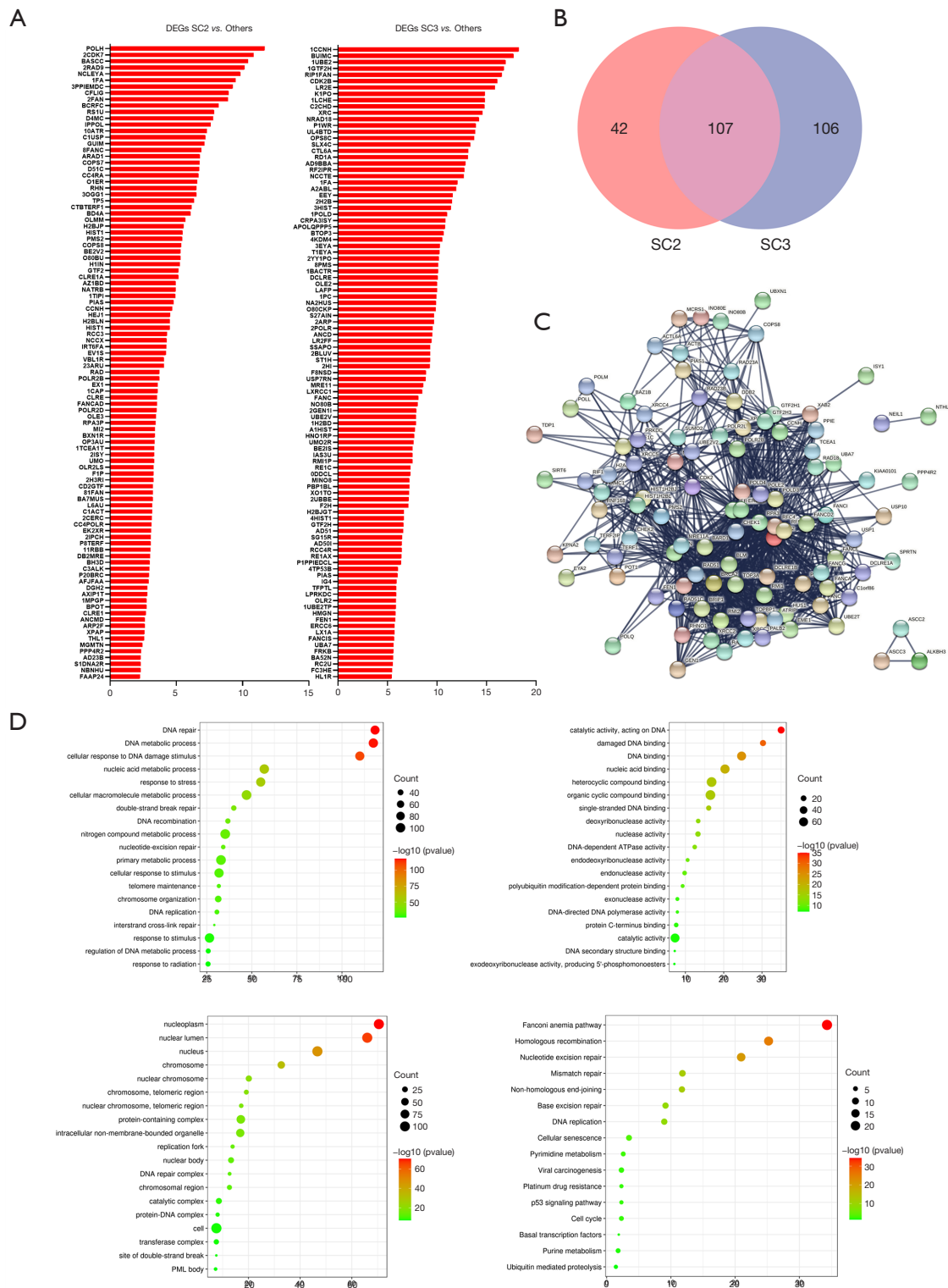
Figure 2 Analysis of DEGs. (A) The top 50 genes with the most significant differences. (B) There were 107 overlapping genes between the two groups. (C) PPI indicates differential gene interaction. (D) The functional enrichment of 107 genes, upper left: biological progress, upper right: molecular function, lower left: cell component, lower right: KEGG pathway. DRGs, DNA repair genes; PPI, protein-protein interaction; KEGG, Kyoto Encyclopedia of Genes and Genomes .

Table 3 Immune related features of each primary cluster

| Variable	Cluster 1 (n=90)		Cluster 2 (n=79)		P value
	Mean	SD	Mean	SD	
Purity	0.3133	0.17	0.3622	0.19	0.080
Imsig B cells	0.0620	0.01	0.0607	0.02	0.572
Imsig interferon	0.0989	0.01	0.0959	0.01	0.000
Imsig macrophages	0.0993	0.01	0.0938	0.01	0.000
Imsig monocytes	0.0962	0.01	0.0920	0.01	0.001
Imsig neutrophils	0.0890	0.01	0.0816	0.01	0.000
Imsig NK cells	0.0371	0.01	0.0322	0.01	0.003
Imsig plasma cells	0.0893	0.01	0.0862	0.02	0.202
Imsig proliferation	0.0723	0.01	0.0725	0.01	0.849
Imsig T cells	0.0833	0.01	0.0797	0.01	0.016
Imsig translation	0.1338	0.00	0.1403	0.01	0.000
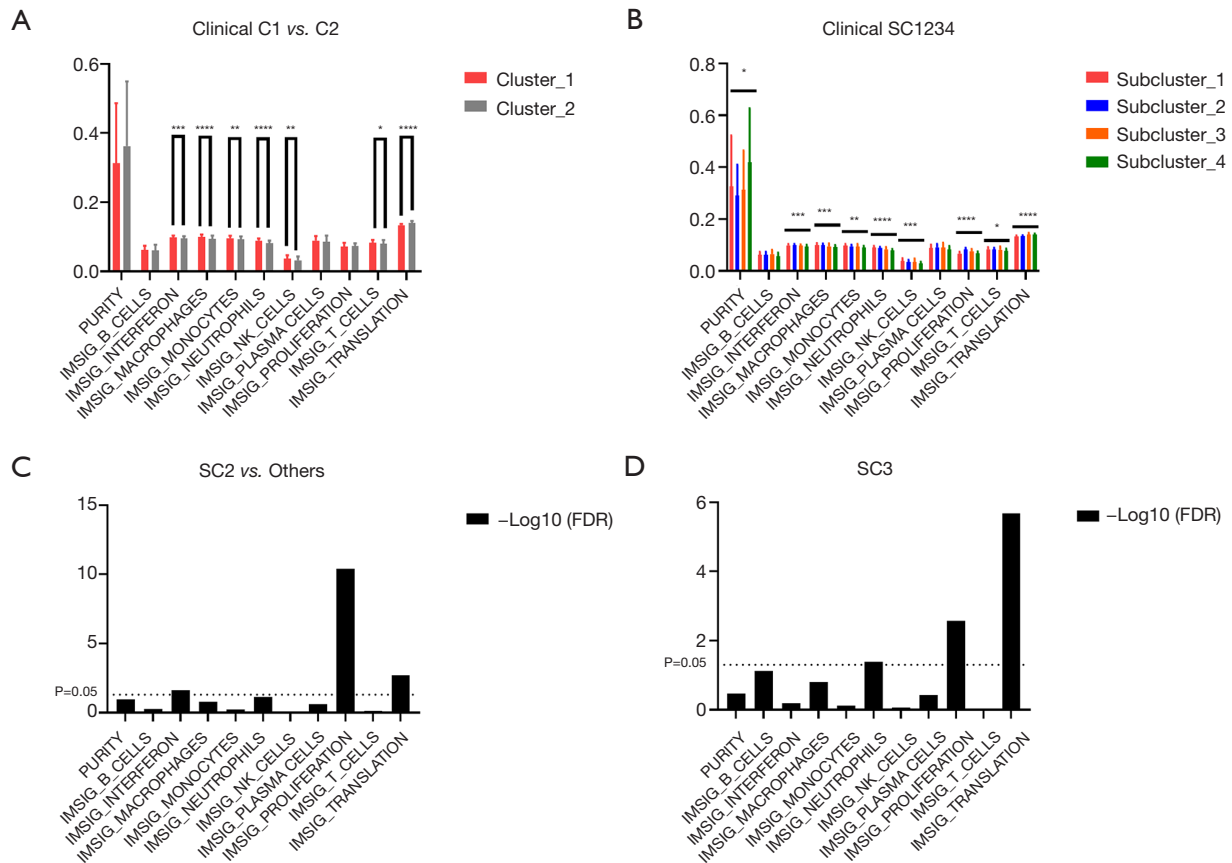
Figure 3 Comparison of indicators of tissue and cells between clusters and subclusters (A,B). Patients in the worst-prognosis group (SC2, A) and the worst-differentiated group (SC3, B) showed significant differences in the proliferation index and protein translation index (C,D). *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

Table 4 Histiocytic features of each secondary cluster

Variable	Subcluster 1 (n=58)		Subcluster 2 (n=32)		Subcluster 3 (n=42)		Subcluster 4 (n=37)		P value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Purity	0.326034	0.19612	0.290313	0.119393	0.312857	0.151163	0.418108	0.20899	0.013
Imsig B cells	0.061532	0.012517	0.062726	0.011672	0.064767	0.01686	0.056101	0.014777	0.051
Imsig interferon	0.098493	0.005062	0.099522	0.004944	0.097137	0.005234	0.094556	0.006515	0.001
Imsig macrophages	0.099593	0.007961	0.098798	0.006909	0.095012	0.011305	0.092458	0.007962	0.001
Imsig monocytes	0.09687	0.007711	0.094981	0.005941	0.093905	0.010367	0.089917	0.007385	0.001
Imsig neutrophils	0.089694	0.006995	0.087872	0.005179	0.083368	0.009114	0.079589	0.00559	0.000
Imsig NK cells	0.038136	0.010567	0.03517	0.00763	0.035036	0.013462	0.028897	0.006977	0.001
Imsig plasma cells	0.088463	0.013168	0.090685	0.013425	0.089689	0.01921	0.082299	0.014289	0.083
Imsig proliferation	0.066816	0.007723	0.082105	0.007643	0.076295	0.008699	0.068281	0.00703	0.000
Imsig T cells	0.083916	0.007961	0.082072	0.007845	0.081547	0.012608	0.077567	0.008866	0.019
Imsig translation	0.133648	0.003655	0.134081	0.003654	0.140297	0.006607	0.140246	0.004049	0.000

Table 5 Histiocytic features in subcluster 2 and subcluster 3

Variable	Subcluster 2 specific					Subcluster 3 specific				
	Subcluster 2 (n=32)		Others (n=137)		P value	Subcluster 3 (n=42)		Others (n=127)		P value
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Purity	0.290313	0.119393	0.346861	0.191033	0.111	0.312857	0.151163	0.343858	0.189419	0.337
Imsig B cells	0.062726	0.011672	0.061057	0.014839	0.553	0.064767	0.01686	0.060251	0.013196	0.075
Imsig interferon	0.099522	0.004944	0.097014	0.005728	0.024	0.097137	0.005234	0.097605	0.00581	0.644
Imsig macrophages	0.098798	0.006909	0.096262	0.009541	0.158	0.095012	0.011305	0.097314	0.008271	0.158
Imsig monocytes	0.094981	0.005941	0.094083	0.008934	0.590	0.093905	0.010367	0.094368	0.007744	0.759
Imsig neutrophils	0.087872	0.005179	0.085026	0.008486	0.071	0.083368	0.009114	0.086291	0.007538	0.040
Imsig NK cells	0.03517	0.00763	0.03469	0.011338	0.821	0.035036	0.013462	0.034697	0.009702	0.859
Imsig plasma cells	0.090685	0.013425	0.087174	0.01572	0.245	0.089689	0.01921	0.087227	0.013856	0.369
Imsig proliferation	0.082105	0.007643	0.070118	0.008845	0.000	0.076295	0.008699	0.071095	0.009849	0.003
Imsig T cells	0.082072	0.007845	0.081475	0.010099	0.755	0.081547	0.012608	0.081602	0.008578	0.975
Imsig translation	0.134081	0.003654	0.137468	0.005822	0.002	0.140297	0.006607	0.135679	0.004763	0.000

NK, natural killer.

especially DNA repair ability (23-25). DNA repair genes play a vital role in keeping cells and organisms alive (26-28). DNA repair mainly includes nucleotide excision repair (NER), base excision repair (BER), DNA double strand break repair (DSBR) and mismatch repair (MMR) pathways. It can repair DNA damage caused by ultraviolet rays and drugs and use complementary strands as templates to replicate and repair damaged DNA to maintain the integrity of the genome (29,30). DNA repair gene single nucleotide polymorphisms (SNPs) are related to DNA repair ability. The functional SNPs of DNA repair genes affect the natural progression of lung cancer (pathological stage and grade, disease progression rate, metastatic tendency) and sensitivity to platinum drugs, thereby affecting the survival

outcome of patients (31-33). However, a large number of research results are inconsistent, and there is no consensus on whether differences in DNA repair gene expression can be used as prognostic indicators. The reason may be that in current clinical practice, the treatment of tumors based on genetic testing is not widespread, especially the selection of treatments based on the precise expression of specific DNA repair genes in different patients and different tumors. In short, the translation of this approach from research results into clinical practice still has a long way to go. A number of studies have found that some biomarkers are related to DNA damage, which provides support for choosing treatments that protect against DNA damage (34-36).

Our results did not show a correlation between DNA repair genes and the prognosis of patients with LUAD, but the prognosis of patients was indeed related to the differentiation and immune indicators of LUAD. The difference in the immune environment is the influencing factor that causes the difference in the prognosis of lung cancer and the sensitivity to treatment (37). The incidence of LUAD is increasing year by year, with a tendency of patients to be younger, with few early symptoms, rapid onset, high mortality and poor prognosis. Therefore, the diagnosis and prognosis of LUAD is very important (1,4,5). At present, because the underlying molecular mechanism of LUAD cannot be determined, its early diagnosis and prognostic treatment are relatively difficult, and most patients are already at an advanced stage when they are diagnosed (38,39). With the rapid development of medical technology and molecular biology technology, the treatment of LUAD has gradually transitioned to molecular targeted therapy (39). Identifying targets that are closely related to the biological behavior of tumors and that play an important role in the prognosis of tumor patients is the key to targeted therapy (6,40). We believe that different expressions of DNA repair genes will lead to differences in tumor biology, which is manifested in different differentiation of tumor tissues. Due to varying histological differentiation, the immune system's response to tumors also varies, which in turn leads to differences in related immune indicators. However, there are currently several problems. First, the biological processes of tumors are relatively complicated. Whether the treatment of DNA repair gene defects can play a key role or not is still unclear. Second, there are many types of DNA repair genes. Different tumors and types have different DNA repair genes. There are not many drugs that specifically target different DNA repair genes. Third, currently very few patients are actually treated based

on DNA repair gene expression, and there are not many clinical data on targeted treatments. Fourth, most of the previous research results are based on excised tumor tissue specimens, with limited applicability to patients in early and late disease stages. Simpler methods and indicators are needed to detect changes in tumor-related DNA repair genes, which are beneficial to clinical practice.

The limitations of this study are its small sample size, and that it included only patients whose tumors had been partially surgically removed. Study findings are not generalizable to the broader population of patients with LUAD, especially patients who were diagnosed at a later stage and patients who were not treated with surgery. Secondly, this study did not screen out specific DNA repair genes and related biomarkers. This requires further exploration and prospective observation and large sample follow-up studies as the next step in this line of research.

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://dx.doi.org/10.21037/jtd-21-949

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/jtd-21-949). Dr. JL reported receiving funding supports from Scientific Research Project of Traditional Chinese Medicine in Guangdong Province No. 20191106; Scientific Research Project of Traditional Chinese Medicine in Guangdong Province No. 20201113 and Youth Research Fund Project "Innovation and Strengthening Hospital Fund" of in the First Affiliated Hospital of Guangzhou University of Traditional Chinese Medicine (No. 2019QN02). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This research was conducted in accordance with the "Declaration of Helsinki" (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020;70:7-30.

2. Shankar A, Saini D, Dubey A, et al. Feasibility of lung cancer screening in developing countries: challenges, opportunities and way forward. Transl Lung Cancer Res 2019;8:S106-21.

3. Yang D, Liu Y, Bai C, et al. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. Cancer Lett 2020;468:82-7.

4. Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. Clin Chest Med 2020;41:1-24.

5. Chen J, Yang H, Teo ASM, et al. Genomic landscape of lung adenocarcinoma in East Asians. Nat Genet 2020;52:177-86.

6. Hirsch FR, Scagliotti GV, Mulshine JL, et al. Lung cancer: current therapies and new targeted treatments. Lancet 2017;389:299-311.

7. Wang Q, Shen B, Qin X, et al. Akt/mTOR and AMPK signaling pathways are responsible for liver X receptor agonist GW3965-enhanced gefitinib sensitivity in non-small cell lung cancer cell lines. Transl Cancer Res 2019;8:66-76.

8. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature 2018;553:446-54.

9. Tanoue LT. Lung Cancer Staging. Clin Chest Med 2020;41:161-74.

10. Baisi A, Raveglia F, De Simone M, et al. TNM Staging

11. Xu JY, Zhang C, Wang X, et al. Integrative Proteomic Characterization of Human Lung Adenocarcinoma. Cell 2020;182:245-61.e17.

12. Yu Y, Tian X. Analysis of genes associated with prognosis of lung adenocarcinoma based on GEO and TCGA databases. Medicine (Baltimore) 2020;99:e20183.

13. Burch PR. Smoking and lung cancer: an overview. Cancer Res 1986;46:3200-3.

14. Klebe S, Leigh J, Henderson DW, et al. Asbestos, Smoking and Lung Cancer: An Update. Int J Environ Res Public Health 2019;17:258.

15. Parascandola M, Xiao L. Tobacco and the lung cancer epidemic in China. Transl Lung Cancer Res 2019;8:S21-30.

16. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007;7:778-90.

17. Parry EM, Gable DL, Stanley SE, et al. Germline Mutations in DNA Repair Genes in Lung Adenocarcinoma. J Thorac Oncol 2017;12:1673-8.

18. Jiang M, Jia K, Wang L, et al. Alterations of DNA damage repair in cancer: from mechanisms to applications. Ann Transl Med 2020;8:1685.

19. Perera D, Poulos RC, Shah A, et al. Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. Nature 2016;532:259-63.

20. Ojlert AK, Halvorsen AR, Nebdal D, et al. The immune microenvironment in non-small cell lung cancer is predictive of prognosis after surgery. Mol Oncol 2019;13:1166-79.

21. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.

22. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6:pl1.

23. Li W, Zhang M, Huang C, et al. Genetic variants of DNA repair pathway genes on lung cancer risk. Pathol Res Pract 2019;215:152548.

24. Tian D, Tang J, Geng X, et al. Targeting UHRF1-dependent DNA repair selectively sensitizes KRAS mutant lung cancer to chemotherapy. Cancer Lett 2020;493:80-90.

25. Lok BH, Rudin CM. Epigenetic targeting of DNA repair in lung cancer. Proc Natl Acad Sci U S A 2019;116:22429-31.

26. Chen LL, Xiong Y. Tumour metabolites hinder DNA repair. Nature 2020;582:492-4.

27. Tubbs A, Nussenzweig A. Endogenous DNA Damage

as a Source of Genomic Instability in Cancer. Cell 2017;168:644-56.

28. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. Science 2008;319:1352-5.

29. Motegi A, Masutani M, Yoshioka KI, et al. Aberrations in DNA repair pathways in cancer and therapeutic significances. Semin Cancer Biol 2019;58:29-46.

30. Yeh CD, Richardson CD, Corn JE. Advances in genome editing through control of DNA repair pathways. Nat Cell Biol 2019;21:1468-78.

31. Wu H, Li S, Hu X, et al. Associations of mRNA expression of DNA repair genes and genetic polymorphisms with cancer risk: a bioinformatics analysis and meta-analysis. J Cancer 2019;10:3593-607.

32. Boige V, Mollevi C, Gourgou S, et al. Impact of single-nucleotide polymorphisms in DNA repair pathway genes on response to chemoradiotherapy in rectal cancer patients: Results from ACCORD-12/PRODIGE-2 phase III trial. Int J Cancer 2019;145:3163-72.

33. Koberle B, Koch B, Fischer BM, et al. Single nucleotide polymorphisms in DNA repair genes and putative cancer risk. Arch Toxicol 2016;90:2369-88.

34. Cleary JM, Aguirre AJ, Shapiro GI, et al. Biomarker-

Guided Development of DNA Repair Inhibitors. Mol Cell 2020;78:1070-85.

35. Mouw KW, Goldberg MS, Konstantinopoulos PA, et al. DNA Damage and Repair Biomarkers of Immunotherapy Response. Cancer Discov 2017;7:675-93.

36. Stover EH, Konstantinopoulos PA, Matulonis UA, et al. Biomarkers of Response and Resistance to DNA Repair Targeted Therapies. Clin Cancer Res 2016;22:5651-60.

37. Chen F, Yang Y, Zhao Y, et al. Immune Infiltration Profiling in Nonsmall Cell Lung Cancer and Their Clinical Significance: Study Based on Gene Expression Measurements. DNA Cell Biol 2019;38:1387-401.

38. Kuhn E, Morbini P, Cancellieri A, et al. Adenocarcinoma classification: patterns and prognosis. Pathologica 2018;110:5-11.

39. Song Q, Shang J, Yang Z, et al. Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. J Transl Med 2019;17:70.

40. Bedard PL, Hyman DM, Davids MS, et al. Small molecules, big impact: 20 years of targeted therapy in oncology. Lancet 2020;395:1078-88.

(English Language Editor: B. Meiser)