

Genetic and tumor microenvironment differences between cell cycle progression pathway altered/non-altered patients with lung adenocarcinoma

Guangyao Shan

Fudan University

Guoshu Bi

Fudan University

Yunyi Bian

Fudan University

Besskaya Valeria

Fudan University

Dejun Zeng

Fudan University

Huan Zhang

Fudan University

Guangyu Yao

Fudan University

Yi Zhang

Fudan University

Hong Fan

Fudan University

Cheng Zhan (✉ czhan10@fudan.edu.cn)

Fudan University

Research Article

Keywords:

Posted Date: April 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-464643/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Lung cancer is the leading cause of cancer-related death worldwide, among which lung adenocarcinoma (LUAD) is the most common type. Identified as a hallmark of cancer, the dysregulated cell cycle progression plays an important role in the promotion and progression of LUAD. This article aims to elucidate the heterogeneity between CDKN2A-CDK/cyclin-RB1 cell cycle progression pathway altered /non-altered patients with LUAD, thus helping us have a better understanding of the effect of the aberrant cell cycle.

Material and Methods

The data of this study were downloaded from The Cancer Genome Atlas (TCGA) data portal (<https://portal.gdc.cancer.gov/>) and UCSC Xena Browser (<http://xena.ucsc.edu/>), including simple nucleotide variation data, RNA-seq gene expression data, survival data, clinical data, and miRNA expression data. After matching the RNA-seq gene expression data, simple nucleotide variation data, miRNA expression data, and survival data with clinical data, 510 gene and long non-coding RNA expression data, 506 simple nucleotide variation data, 440 microRNA expression data, and 497 survival data were included in this study for further analysis. R software (version 4.0.3) was used for analysis.

Results

After dividing the patients into mutation ($n = 57$) and wild ($n = 453$) groups according to the cell cycle progression pathway status, we found no significant difference in survivorship between them. The mutation group had a higher mutational load and mutational rates of various genes such as tumor protein P53 (TP53) compared to the wild group. Subsequently, we analyzed the differentially expressed genes (DEGs) between the two groups. Among the 58387 genes analyzed, 302 were upregulated, and 354 were downregulated in the mutation group. Enrichment analysis indicated that these DEGs were closely related to metabolism items and cell cycle-related events. After performing immune cell infiltration analysis, we found the two groups have different patterns of immune cell profiling. Albeit the immune and stromal scores were higher in the wild group, we failed to find any significant difference between the two groups. Finally, we build a computational model to predict the cell cycle progression pathway-related gene mutation by LASSO-binary logistic regression analysis, the predictive accuracy of which is 0.88.

Conclusion

In summary, our study compared the genetic and microenvironment differences between cell cycle progression pathway altered /non-altered patients with LUAD by analyzing the data from TCGA datasets.

We hope our findings could improve our understanding of the heterogeneity between the two kinds of patients, thus providing new insight into LUAD patients' treatments.

Introduction

Lung cancer is the leading cause of cancer-related death and the most common type of cancer in men globally¹. The two major types of lung cancer are non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC). The former type accounts for 85% of lung cancer cases², among which lung adenocarcinoma (LUAD) is the most common subtype with more than 500,000 deaths each year worldwide³. According to the report published by the American Cancer Society and the National Cancer Institute, the 5-year relative survival of NSCLC for all stages is only 23%⁴. In the past decades, the utilization of targeted therapies and immunotherapies alone or in combination with conventional therapies such as chemotherapies has greatly improved the prognosis of LUAD patients⁵, and they begin to play an increasingly important role in the treatment of LUAD. These remarkable achievements are attributed to the deep understanding of the genetic and molecular traits of LUAD. In this article, we will talk about the cell cycle progression pathway, which is important in the promotion and progression of LUAD.

The cell cycle, a highly organized process regulated by cyclin and cyclin-dependent kinase (CDK) complexes and other regulators, consists of four sequential phases: G1, S, G2, and M phases⁶. In the G1 phase, cells are busy with biosynthesis prepared for the following steps. S phase is characterized by DNA replication and the synthesis of related proteins like histones. During the G2 phase, cells enhance the lipid synthesis required for membrane construction and guarantee that everything is ready to initiate the mitosis⁶. In the M phase, the chromosomes will be evenly separated into two daughter cells, and then the mitosis comes to an end. Compared to normal tissues that could carefully control the production and release of growth-promoting signals during the cell cycle, cancer cells could maintain sustainable proliferative signals and avoid programs that may arrest cell cycle progression⁷. We found the five genes, including cyclin-dependent kinase inhibitor 2A (CDKN2A), cyclinD1 (CCND1), cyclin-dependent kinase 4 (CDK4), cyclinE1 (CCNE1), and retinoblastoma 1 (RB1), are frequently aberrant in LUAD patients, and they orchestrate a pathway together which could modulate cell cycle progression. CDKN2A, which encodes the CDK inhibitor p16^{INK4a} and p14^{ARF} protein, is located upstream of the cell cycle progression pathway and could be found loss-of-function in multiple human tumors, including NSCLC⁸. As a tumor suppressor gene, CDKN2A functions as an inhibitor of CDK4, which is a driver for mitosis. CDK4-CCND1 complexes, along with CDK2-CCNE1 complexes, could phosphorylate and inactivate the RB1 to release the E2F transcription factor, thus inducing the cell to complete G1-to-S transition⁹⁻¹². Among all the mitotic steps mentioned above, the transition from G1 to S is the most crucial for cell cycle progression. Once the cell has entered the S phase, it is bound to get through S, G2, and M phases and generate two daughter cells¹³. As we can see, the cell cycle progression pathway is well organized by various regulators, errors at any step may disrupt the homeostasis of cell proliferation, apoptosis, and death, thus contributing to tumor promotion and progression.

In this article, we systematically analyzed the differences in somatic mutations, gene expression, immune cell infiltration, microRNAs (miRNAs), and long non-coding RNAs (lncRNAs) between the cell cycle progression pathway altered and non-altered patients with LUAD. This study aims to enhance our understanding of the function of the cell cycle progression, which could shed new lights on the development of new drugs targeting this pathway.

Materials And Methods

2.1 TCGA data collection

The TCGA data of this study were downloaded from The Cancer Genome Atlas (TCGA) data portal (<https://portal.gdc.cancer.gov/>) and UCSC Xena Browser (<http://xena.ucsc.edu/>), including simple nucleotide variation data, RNA-seq gene expression data, survival data, clinical data, and miRNA expression data. After matching the RNA-seq gene expression data, simple nucleotide variation data, survival data, miRNA data with clinical data, 510 gene and lncRNA expression data, 440 miRNA expression data, 506 simple nucleotide variation data, and 497 survival data were included in this study for further analysis in R software (version 4.0.3).

2.2 Clinical data of the patients and survival analysis

Patients who harbored any mutant cell cycle progression pathway-related genes (CDKN2A, CCND1, CDK4, CCNE1, and RB1) were defined as the mutation group, and the others were defined as the wild group. Clinical data including age, sex, race, anatomic location of the lesion, smoking history, and TNM stage were collected for analysis by t-test. Next, Kaplan-Meier survival analysis was performed between the two groups.

2.3 Somatic mutations and mutational load

The simple nucleotide variation data were divided into two groups according to the CDKN2A-CDK/cyclin-RB1 pathway status, and then they were performed by the R maftools package separately. P-value <0.05 was considered as significantly different.

2.4 Differentially expressed genes (DEGs) and enrichment analysis

Based on the RNA-seq gene expression data (count format), DEGs were analyzed by the R edgeR package. The absolute value of $\text{Log}_2\text{FoldChange}$ ($|\text{Log FC}| > 1$) and p-value <0.05 were considered as significantly different. Subsequently, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed using R org.Hs.eg.db and ClusterProfiler package. Gene set enrichment analysis (GSEA) and Gene set variation analysis (GSVA) were also performed for further analysis.

2.5 Differentially expressed miRNAs and lncRNAs

Differentially expressed miRNAs were analyzed by the R limma package. $|\text{Log FC}| > 0.5$ and $p\text{-value} < 0.05$ were considered as significantly different. Differentially expressed lncRNAs were analyzed by the R edgeR package. $|\text{Log FC}| > 1$ and $p\text{-value} < 0.05$ were considered as significantly different. After matching the differentially expressed miRNAs with DEGs and differentially expressed lncRNAs using miRWalk (<http://mirwalk.umm.uni-heidelberg.de/>) and miRcode (<http://www.mircode.org/>), we made a DEGs-differentially expressed miRNAs-differentially expressed lncRNAs network in Cytoscape (version 3.8.2).

2.6 Characteristics of immune cell infiltration

The profiling of 22 immune cell types was analyzed using gene expression data (FPKM format) and the leucocyte signature matrix (LM22) in CIBERSORT (<https://cibersort.stanford.edu/>) [permutations set =1000]. LM22, composed of 547 genes, could distinguish 22 human hematopoietic cells, including various subtypes of T cells, B cells, plasma cells, NK cells, and myeloid cells¹⁴. Besides, we calculated the immune and stromal scores by the R estimate package. $P\text{-value} < 0.05$ was considered as significantly different.

2.7 Construction of predictive model by LASSO - binary logistic regression analysis

Based on the 656 DEGs, we used LASSO (Least absolute shrinkage and selection operator) and binary logistic regression analysis to build a computational model to predict if a LUAD patient harbored any cell cycle progression pathway-related gene mutation. The LASSO regression analysis is a penalized method to select data with high dimensions and reduce the impact of overfitting^{15,16}. Ten-fold cross-validation was adopted using the R glmnet package to determine the optimal parameter λ and corresponding genes. Subsequently, we conducted binary logistic regression using the candidate genes and build the predictive model. Based on the gene expression data, each sample obtained a predictive value according to this model. Then we exploited the receiver operating characteristic curve (ROC) to determine the cutoff value for the established model and calculate the predictive accuracy.

Results

3.1 Clinical data of the patients and survival analysis

The mutational rates of RB1, CDKN2A, CCNE1, CDK4 and CCND1 in TCGA samples are 7.6%, 4.9%, 1.2%, 0.6% and 0.4%. After dividing the patients into mutation ($n=57$) and wild ($n=453$) groups, we found that more than half of the patients in the mutation group harbored RB1 mutation (68.4%, the common variants in lung cancer were C706F, G748K, and R661W), followed by CDKN2A (43.9%, the common variants in LUAD were R80Q and H83Y), CCNE1 (10.5%), CDK4 (5.3%, the common variants in LUAD were R24L and R24C) and CCND1 (3.5%) [Figure 1A]. Among the 57 patients in the mutation group, 8 of them had more than one mutant gene in the pathway. The clinical data analysis showed that the differences between the two groups in age, sex, anatomical location of the lesion, smoking history, and TNM stage of the disease were not statistically significant [Table 1]. Then we performed the Kaplan-Meier survival analysis between the two groups, which indicated that the mutation group had a higher survival rate most

of the time [Figure 1B]. But as the number of patients at risk in this group was quite few after 42 months, the curve of the mutation group suffered a sharp decline after that, resulting in a worse prognosis than the wild group. Thus, we think this survival result needs to be further verified in larger groups.

3.2 Differences in somatic mutations and mutational load

Different mutational patterns were detected between the mutation and wild groups [Figure 2A&B, Supplementary Figure 1]. Among the top 20 mutant genes, most of them, including tumor protein 53 (TP53, 74% VS 44%), Xin actin binding repeat containing 2 (XIRP2, 37% VS 22%), zinc finger protein 536 (ZNF536, 33% VS 18%), piccolo presynaptic cytomatrix protein (PCLO, 32% VS 16%), and protocadherin related 15 (PCDH15, 33% VS 17%) were significantly higher in the mutation group. Only 2 [Kirsten Rat Sarcoma Viral Proto-Oncogene (KRAS) and Ankyrin 2 (ANK2)] had higher mutational rates in the wild group. After calculating the mutational load between the two groups, we found the mutation group had a higher mutational load ($P < 0.01$) [Figure 3A], which indicated that the mutation group is more likely to benefit from immunotherapies. Then we analyzed the expression patterns of 15 immune checkpoints and 20 co-stimulators¹⁷ between the two groups [Figure 3B&C]. Among co-stimulators, TNF superfamily member 13 (TNFSF13) showed a significant difference and was higher in the wild groups. As for immune checkpoints, lymphocyte activating gene 3 (LAG3) was significantly higher in the mutation group.

3.3 DEGs and functional enrichment analysis

Among the 58,387 genes analyzed, 302 were upregulated and 354 were downregulated in the mutation group [Figure 4A]. Among the DEGs, vestigial like family member 2 (VGLL2, logFC=6.41, $P < 0.0001$), glucagon (GCG, logFC=5.64, $P < 0.0001$), and sclerostin (SOST, logFC=4.04, $P < 0.0001$) were the most significantly upregulated genes; defensin alpha 5 (DEFA5, logFC= -8.11, $P < 0.0001$), proline rich protein BstNI subfamily 4 (PRB4, logFC= -6.87, $P < 0.0001$), and sperm associated antigen 11B (SPAG11B, logFC= -6.66, $P < 0.0001$) were the most significantly downregulated genes.

To explore the biological functions affected by DEGs, we performed the GO and KEGG analysis. In GO analysis, the top five significantly different pathways were cornification, antimicrobial humoral response, neurotransmitter transport, negative regulation of peptidase activity, and negative regulation of endopeptidase activity [Figure 4B-D]. KEGG analysis showed the DEGs were related to fat digestion and absorption, drug metabolism – cytochrome P450, and retinol metabolism [Figure 4E]. So we can see that these DEGs may make a difference in the metabolism items in our body.

Apart from GO and KEGG analysis, we further investigated the influence of cell cycle progression pathway status using GSEA and GSVA analysis. Processes like cell cycle, DNA replication, and homologous recombination were significantly different between the two groups by GSEA analysis [Figure 5A-C]. GSVA analysis showed biological processes such as DNA replication, E2F targets, mitotic spindle, and G2M checkpoints were significantly different between the two groups [Figure 5D-I], which was consistent with the GSEA results, suggesting that the different gene expression patterns have a significant impact on the cell cycle-related events.

3.4 Differentially expressed miRNAs and lncRNAs

Both miRNAs and lncRNAs could act as important modulators of the expression and function of genes. Here, we analyzed the differentially expressed miRNAs and lncRNAs between the two groups [Supplementary Figure 2A&B]. The results showed that 31 miRNAs were upregulated in the mutation group, the most significantly different ones of which were miR-3689e (logFC=2.30, P<0.01), miR-3689f (logFC=1.95, P<0.01), and miR-1197 (logFC=1.70, P<0.001). 5 miRNAs were downregulated in the mutation group and the most significantly different ones were miR-203b-3p (logFC=-0.86, P<0.001), miR-508-3p (logFC=-0.74, P<0.05), and miR-135b-5p (logFC=-0.73, P<0.01). For lncRNAs, LINC01305 (logFC=3.21, P<0.0001) was the most significantly different one among 31 upregulated lncRNAs in the mutation group; AF003626.1 (logFC=-4.63, P<0.0001) was the most significantly different one among 64 downregulated lncRNAs.

After matching DEGs and differentially expressed lncRNAs with differentially expressed miRNAs, we built a ceRNA network of DEGs-differentially expressed miRNAs- differentially expressed lncRNAs (Here, only genes negatively regulated by miRNAs were included in this network) [Figure 6]. For example, miR-509-3p could regulate ZNF713 and the other genes. We found that miR-3131 and miR-185-3p could regulate more than 40 genes, which indicated that the two miRNAs and the genes they regulate might play a significant role in the differences between the mutation and wild group. As for lncRNAs, RP11-230G5.2 handled the most miRNAs.

3.5 Characteristics of immune cell infiltration

Tumor immune cell infiltration refers to the migration of immune cells from the peripheral blood to the tumor tissue, where they exert their function¹⁸. Here, different immune cell infiltration patterns were detected between the two groups [Figure 7A, Supplementary Figure 3] using CIBERSORT. Among all the significantly different cell types, we found that the proportion of naïve B cells (P<0.05) and M1 macrophages (P<0.01) were higher in the mutation group, while CD4⁺ memory resting T cells (P<0.01), monocytes (P<0.001), activated mast cells (P<0.05), resting (P<0.05) and activated (P<0.01) dendritic cells are higher in the wild group, indicating that the mutation group may have a cold tumor microenvironment. Then we further explore the tumor purity between the mutation and wild groups. The results suggested that the immune and stromal scores were higher in the wild group, but we failed to find any significant difference between the two groups [Figure 7B&C].

3.6 Construction of predictive model by LASSO - binary logistic regression analysis

In LASSO analysis, the optimal parameter λ was set as lambda.1se criteria, which means one standard deviation from the minimum binomial deviance [Figure 8A&B]. Seven genes including natriuretic peptide A (NPPA), free fatty acid receptor 2 (FFAR2), mucolipin TRP cation channel 3 (MCOLN3), zinc finger protein 713 (ZNF713), F-box protein 43 (FBXO43), TATA-box binding protein associated factor 7 like (TAF7L), and synaptic vesicle glycoprotein 2B (SV2B) were employed to build the predictive model. Then we conducted the binary logistic regression analysis to get the coefficients of every factor, and the

predictive model is as follows: $\text{score} = -0.8434 \times \text{NPPA} + 0.4617 \times \text{FFAR2} + 0.6178 \times \text{MCOLN3} + 0.9924 \times \text{ZNF713} + 0.8930 \times \text{FBXO43} + 0.5772 \times \text{TAF7L} + 1.3533 \times \text{SV2B}$ [Figure 8C]. ROC demonstrated the cutoff value of this model was 3.008, and AUC (area under the curve) is 0.777 [Figure 8D]. When the score is equal to or higher than 3.008, the test sample is more likely to harbor cell cycle progression pathway-related gene mutation. Otherwise, it is more likely to carry no cell cycle progression pathway-related gene mutation. After testing this model in TCGA data, we found the predictive accuracy is 0.880.

Discussion

Lung cancer is the leading cause of cancer-related death worldwide, among which LUAD has been the most prevalent histopathological type since 1985¹⁹. In our study, we explored the differences between the CDKN2A-CDK/cyclin-RB1 cell cycle progression pathway altered/non-altered patients with LUAD in somatic mutations, gene expression, immune cell infiltration, and various biological processes. Compared to the wild group, the mutation group had a higher mutational load and mutational rates of multiple genes such as TP53. By analyzing the RNA-seq gene expression data, we found that 302 protein-coding genes were upregulated in the mutation group, such as VGLL2, GCG, and cysteine-rich secretory protein 3 (CRISP3), and 354 were downregulated like DEFA5. Through enrichment analysis, the DEGs were closely related to metabolism items and cell cycle-related events. The same as DEGs, miRNAs, and lncRNAs, which play an important role in epigenetic modification, differed between the two groups. Subsequently, we built a ceRNA network of DEGs-differentially expressed miRNAs- differentially expressed lncRNAs to demonstrate their relationships. After performing immune infiltration analysis, we found that 7 out of 22 cell types were significantly different between the two groups. The proportion of CD4⁺ memory resting T cells ($P < 0.01$), monocytes ($P < 0.001$), activated mast cells ($P < 0.05$), resting ($P < 0.05$), and activated ($P < 0.01$) dendritic cells were higher in the wild group, indicating that the mutation group may have a cold tumor microenvironment.

Most RB1-mutant tumors also harbor TP53 co-alterations²⁰. In our study, the mutational rate of TP53 between mutation and the wild group is 74% VS 44%. As a well-known tumor suppressor gene, TP53 could monitor the abnormalities within the cell. Once the cell suffers excessive DNA damage or the status is not optimal for mitosis, TP53 will suspend the progression of the cell cycle in time until the conditions have been normalized. Besides, when a cell encounters irreversible impairment, TP53 could trigger apoptosis to maintain homeostasis⁷. Mutant TP53 not only lost the abilities of surveillance but also create a circumstance that favors immune evasion and tumor progression by disrupting the innate immune signaling²¹. Given the important roles RB1 and TP53 play in the cell cycle progression and the high mutational rates in LUAD patients, both of them may be potential targets for drug development.

The mutation group had a higher mutational load ($P < 0.01$), indicating that these patients are likely to produce more neo-antigens that could be recognized by T cells²². Therefore, the mutation group may acquire a better clinical outcome to immunotherapies compared to the wild group. Aside from mutational

load, LAG3, an immune checkpoint molecule, is also higher in the mutation group. Immune checkpoints are crucial for maintaining autoimmune tolerance but could be harnessed by cancer cells to evade immune surveillance. The advent of anti-programmed death 1 (PD-1) and programmed death-ligand 1 (PD-L1) drugs have shown great efficacy in lung cancer patients. However, less than 25 percent of patients are expected to benefit from them²³. LAG3 could be detected on tumor-infiltrating lymphocytes (TILs), B cells, natural killer cells, and Tregs²⁴. With nearly 20% sequence identical to CD4, it could bind with MHC-II, thus negatively regulating the function of NK cells, CD4⁺ and CD8⁺ T cells, and Tregs²⁵. In NSCLC patients, LAG3 is usually associated with PD-1 expression and poor prognosis²⁶. It is reported that LAG3 may have a synergistic effect with PD-1/PD-L1. Therefore, LAG3 may be a new biomarker for NSCLC patients and provide new strategies to improve the effects of anti-PD-1 and PD-L1 therapies. Patients with altered cell cycle progression pathways are likely to benefit from it.

As our results show, the mutation group had a higher expression level of CRISP3. CRISP3 comprises two domains: a CAP domain common to the entire superfamily and a clade-specific CRISP domain²⁷. The CRISP domain is shown to regulate certain types of ion channels. Accumulating evidence suggests that CRISP3 may be involved in the oncogenesis of oral squamous cell carcinoma²⁸ and progression of prostate cancer²⁷. Highly enriched in male reproductive tract, CRISP3 protein is found to play an important role in the dissemination of cancer cells in mouse models of prostate cancer, and it is considered a potential target for treatment²⁷. In breast cancer patients, CRISP3 is negatively related to prolonged overall survival (OS) and disease-free survival (DFS)²⁹. Analogous to prostate and breast cancer, CRISP3 may also play a pro-tumor role in LUAD, but its exact effect is still unknown.

Compared to the wild group, DEFA5 is significantly downregulated in the mutation group. Defensins are small cysteine-rich cationic polypeptides that are secreted by specific leukocytes and epithelial cells³⁰. Recognized as antimicrobial agents, defensins also play an important anti-tumor role³¹. Via binding with B cell-specific Moloney murine leukemia virus integration site 1 (BMI1), DEFA5 could exert anti-tumor effects by inhibiting the cell mitosis. BMI1 serves as a transcription inhibitor of CDKN2A³², which is a tumor suppressor gene and could arrest the cell cycle progression. Therefore, by antagonizing BMI1 expression, DEFA5 could indirectly call a halt to the aberrant cell division. In an experiment of gastric cancer cells, DEFA5 overexpression dramatically increased the number of G1-phase cells but significantly decreased G2/M-phase cells³⁰, indicating that DEFA5 overexpression could result in cell cycle arrest at the G1 phase. Given the role of DEFA5 in cell cycle arrest, it may be a novel target for developing new drugs, but further exploration is needed to clarify its effect in LUAD.

miR-3131, which was upregulated in the mutation group, regulated more than 40 genes in the ceRNA network. Important as it is, its function is barely reported. Wang CQ et al. founded that the overexpression of miR-3131 is positively correlated with cell viability of hepatocellular cancer, and miR-3131 could inhibit apoptosis by upregulating BCL-2 and downregulating caspase-3³³. So miRNA-3131 may also exert pro-tumor effects in LUAD, but its role still needs further verification. MiR-185-3p, also upregulated in the mutation group, is proven a tumor suppressor in colorectal cancer (CRC)³⁴, the overexpression of which

could enhance the chemosensitivity of CRC patients. Therefore, it is considered a potential target for the treatment of 5-FU-insensitive CRC³⁵.

The tumor microenvironment, which is composed of tumor cells, immune cells, and stromal cells, is closely correlated with the response to immunotherapies and drug resistance³⁶. We found that the wild group had a higher proportion of monocytes and dendritic cells through immune cell infiltration analysis. Monocytes are highly heterogeneous in the complex tumor microenvironment. Based on the expression level of Ly6C, monocytes can be divided into Ly6C^{hi} classical monocytes, and Ly6C^{lo} non-classical monocytes, both of which could play pivotal roles in pro- and anti-tumor effects in different conditions³⁷. For example, non-classical monocytes could release immunosuppressive cytokines, such as IL-10, which suppresses the function of cytotoxic T lymphocytes in tumors³⁸. They are also capable of recruiting NK cells to the pre-metastatic niche³⁹, thus killing the cancer cells. This NK cell-recruiting function has been confirmed in patients with early-stage lung cancer⁴⁰. Since monocytes function like a double-edged sword, further research may look at in which condition it can exert anti-tumor effects in LUAD patients.

As the most important antigen-presenting cells (APCs) in our body, dendritic cells (DCs) play a significant role in both innate and adaptive immunity. By presenting the antigen-MHC complex to the T cells, DCs could activate the T cell-mediated immune response to kill cancer cells⁴¹. What's more, by secreting high levels of type- γ interferons and chemokines such as CXCL9 and CXCL10, DCs could facilitate the recruitment of effector T cells and NK cells into tumors and maintain the cytotoxic functions of effector cells^{42,43}. Sipuleucel-T, a DC-based vaccine, has shown clinical benefits and has been approved by the US Food and Drug Administration (FDA) for treating prostate cancer⁴⁴. As the stunning anti-tumor effects DCs could elicit, it can be employed for killing cancer cells and may greatly improve the prognosis of LUAD patients.

Since sustaining proliferative signals is a hallmark of cancer⁷, inhibition of the dysregulated cell division is a promising strategy for cancer therapies. The administration of CDK4/6 inhibitors could prevent the phosphorylation of RB1 by CDK-cyclin complexes, calling a halt to the cell cycle progression. Up to date, three generations of CDK inhibitors have been developed for cancer treatment. As the third-generation selective CDK inhibitors, Palbociclib, ribociclib, and abemaciclib have been approved by the US FDA to treat breast cancer⁴⁵. Despite the clinical benefit in specific breast cancer patients, the results of clinical trials accessing the effect of single-agent CDK inhibitors in NSCLC are frustrating⁴⁶. To fully evaluate the application of CDK inhibitors in NSCLC, further investigation in this field may target combination with other anti-cancer therapies and the discovery of predictive biomarkers.

There are also some limitations of our study. Besides the status of the cell cycle progression pathway, variants such as comorbidities could also contribute to differences between the two groups, but these factors were neglected in this article. What's more, the number of patients in the mutation group is smaller than the wild group, and we cannot verify our results in Gene Expression Omnibus (GEO) datasets due to the lack of gene mutation information, so further investigation is needed in larger groups.

Conclusion

In conclusion, we analyzed the differences in somatic mutations, gene expression, various biological pathways, and immune cell infiltration between the CDKN2A-CDK/cyclin-RB1 cell cycle progression pathway altered/ non-altered patients with LUAD. Patients in the two groups have the different somatic mutation and gene expression patterns, profiling of immune cell infiltration, and biological pathways that may play an important role in oncogenesis and tumor metastasis. We hope our study could improve our understanding of the function of the cell cycle progression pathway, thus contributing to the development of new therapies and precision medicine in the future.

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Competing interests: The authors declare no competing interests in this work.

Acknowledgments: Not applicable.

Funding: This research received no external funding.

Author Contributions: Conceptualization was contributed by Cheng Zhan, Yi Zhang, and Hong Fan; Data collection and curation were contributed by Guangyao Shan and Guoshu Bi; Data analysis and interpretation were contributed by Guangyao Shan, Guoshu Bi, Yunyi Bian, Besskaya Valeria, and Guangyu Yao; Draft of the manuscript was contributed by Guangyao Shan, Huan Zhang, and Dejun Zeng; Critical revision of the manuscript was contributed by Cheng Zhan, Yi Zhang, and Hong Fan. Final approval of manuscript and submission were contributed by all authors.

Availability of data and material: Data adopted in this study are available in TCGA(<http://portal.gdc.cancer.gov/>) and UCSC Xena Browser (<http://xena.ucsc.edu/>).

References

- 1 Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, doi:10.3322/caac.21660 (2021).
- 2 Oser, M. G., Niederst, M. J., Sequist, L. V. & Engelman, J. A. Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. *The Lancet Oncology* **16**, e165-e172, doi:10.1016/s1470-2045(14)71180-5 (2015).
- 3 Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107-1120, doi:10.1016/j.cell.2012.08.029 (2012).

- 4 Miller, K. D. *et al.* Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin***69**, 363-385, doi:10.3322/caac.21565 (2019).
- 5 Sudhindra, A., Ochoa, R. & Santos, E. S. Biomarkers, prediction, and prognosis in non-small-cell lung cancer: a platform for personalized treatment. *Clin Lung Cancer***12**, 360-368, doi:10.1016/j.clcc.2011.02.003 (2011).
- 6 Icard, P., Fournel, L., Wu, Z., Alifano, M. & Lincet, H. Interconnection between Metabolism and Cell Cycle in Cancer. *Trends Biochem Sci***44**, 490-501, doi:10.1016/j.tibs.2018.12.007 (2019).
- 7 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell***144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- 8 Jeong, E. H. *et al.* Anti-tumor effect of CDK inhibitors on CDKN2A-defective squamous cell lung cancer cells. *Cell Oncol (Dordr)***41**, 663-675, doi:10.1007/s13402-018-0404-6 (2018).
- 9 Ohtani, N., Yamakoshi, K., Takahashi, A. & Hara, E. The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression. *J Med Invest***51**, 146-153 (2004).
- 10 Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature***511**, 543-550, doi:10.1038/nature13385 (2014).
- 11 Knudsen, E. S., Pruitt, S. C., Hershberger, P. A., Witkiewicz, A. K. & Goodrich, D. W. Cell Cycle and Beyond: Exploiting New RB1 Controlled Mechanisms for Cancer Therapy. *Trends Cancer***5**, 308-324, doi:10.1016/j.trecan.2019.03.005 (2019).
- 12 Helsten, T. *et al.* Cell-Cycle Gene Alterations in 4,864 Tumors Analyzed by Next-Generation Sequencing: Implications for Targeted Therapeutics. *Mol Cancer Ther***15**, 1682-1690, doi:10.1158/1535-7163.MCT-16-0071 (2016).
- 13 Wiman, K. G. & Zhivotovsky, B. Understanding cell cycle and cell death regulation provides novel weapons against human diseases. *J Intern Med***281**, 483-495, doi:10.1111/joim.12609 (2017).
- 14 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods***12**, 453-457, doi:10.1038/nmeth.3337 (2015).
- 15 Vasquez, M. M. *et al.* Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. *BMC Med Res Methodol***16**, 154, doi:10.1186/s12874-016-0254-8 (2016).
- 16 Tibshirani, R. Regression shrinkage selection via the LASSO. *Journal of the Royal Statistical Society Series B (Statistical Methodology)***73**, 273-282 (2011).

- 17 Bi, G. *et al.* Identification and validation of tumor environment phenotypes in lung adenocarcinoma by integrative genome-scale analysis. *Cancer Immunol Immunother***69**, 1293-1305, doi:10.1007/s00262-020-02546-3 (2020).
- 18 Sui, Q. *et al.* Genetic and microenvironmental differences in non-smoking lung adenocarcinoma patients compared with smoking patients. *Transl Lung Cancer Res***9**, 1407-1421, doi:10.21037/tlcr-20-276 (2020).
- 19 Lu, T. *et al.* Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *Cancer Manag Res***11**, 943-953, doi:10.2147/CMAR.S187317 (2019).
- 20 Skoulidis, F. & Heymach, J. V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat Rev Cancer***19**, 495-509, doi:10.1038/s41568-019-0179-8 (2019).
- 21 Ghosh, M. *et al.* Mutant p53 suppresses innate immune signaling to promote tumorigenesis. *Cancer Cell*, doi:10.1016/j.ccell.2021.01.003 (2021).
- 22 Jardim, D. L., Goodman, A., de Melo Gagliato, D. & Kurzrock, R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer Cell***39**, 154-173, doi:10.1016/j.ccell.2020.10.001 (2021).
- 23 He, Y. *et al.* LAG3 Protein Expression in Non-Small Cell Lung Cancer and Its Relationship with PD-1/PD-L1 and Tumor-Infiltrating Lymphocytes. *J Thorac Onco***12**, 814-823, doi:10.1016/j.jtho.2017.01.019 (2017).
- 24 Okamura, T., Fujio, K., Sumitomo, S. & Yamamoto, K. Roles of LAG3 and EGR2 in regulatory T cells. *Ann Rheum Dis***71 Suppl 2**, i96-100, doi:10.1136/annrheumdis-2011-200588 (2012).
- 25 He, Y. *et al.* Lymphocyte-activation gene-3, an important immune checkpoint in cancer. *Cancer Sci***107**, 1193-1197, doi:10.1111/cas.12986 (2016).
- 26 Dal Bello, M. G., Alama, A., Coco, S., Vanni, I. & Grossi, F. Understanding the checkpoint blockade in lung cancer immunotherapy. *Drug Discov Today***22**, 1266-1273, doi:10.1016/j.drudis.2017.05.016 (2017).
- 27 Volpert, M. *et al.* CRISP3 expression drives prostate cancer invasion and progression. *Endocr Relat Cancer***27**, 415-430, doi:10.1530/ERC-20-0092 (2020).
- 28 Ko, W.-C. *et al.* Copy number changes of CRISP3 in oral squamous cell carcinoma. *Oncology letters***3**, 75-81 (2012).
- 29 Wang, Y. *et al.* Low expression of CRISP3 predicts a favorable prognosis in patients with mammary carcinoma. *J Cell Physiol***234**, 13629-13638, doi:10.1002/jcp.28043 (2019).
- 30 Wu, Z., Ding, Z., Cheng, B. & Cui, Z. The inhibitory effect of human DEFA5 in growth of gastric cancer by targeting BMI1. *Cancer Sci***112**, 1075-1083, doi:10.1111/cas.14827 (2021).

- 31 Jin, G. & Weinberg, A. Human antimicrobial peptides and cancer. *Semin Cell Dev Biol***88**, 156-162, doi:10.1016/j.semcdb.2018.04.006 (2019).
- 32 Siddique, H. R. & Saleem, M. Role of BMI1, a stem cell factor, in cancer recurrence and chemoresistance: preclinical and clinical evidences. *Stem Cells***30**, 372-378, doi:10.1002/stem.1035 (2012).
- 33 Wang, C. *et al.* An indel polymorphism within pre-miR3131 confers risk for hepatocellular carcinoma. *Carcinogenesis***38**, 168-176, doi:10.1093/carcin/bgw206 (2017).
- 34 Lu, Z. J. *et al.* MicroRNA-185 suppresses growth and invasion of colon cancer cells through inhibition of the hypoxia-inducible factor-2alpha pathway in vitro and in vivo. *Mol Med Rep***10**, 2401-2408, doi:10.3892/mmr.2014.2562 (2014).
- 35 Zhou, C., Kong, W., Ju, T., Xie, Q. & Zhai, L. MiR-185-3p mimic promotes the chemosensitivity of CRC cells via AQP5. *Cancer Biol Ther***21**, 790-798, doi:10.1080/15384047.2020.1761238 (2020).
- 36 Wu, T. & Dai, Y. Tumor microenvironment and therapeutic response. *Cancer Lett***387**, 61-68, doi:10.1016/j.canlet.2016.01.043 (2017).
- 37 Jeong, J., Suh, Y. & Jung, K. Context Drives Diversification of Monocytes and Neutrophils in Orchestrating the Tumor Microenvironment. *Front Immunol***10**, 1817, doi:10.3389/fimmu.2019.01817 (2019).
- 38 Jung, K. *et al.* Ly6Clo monocytes drive immunosuppression and confer resistance to anti-VEGFR2 cancer therapy. *J Clin Invest***127**, 3039-3051, doi:10.1172/JCI93182 (2017).
- 39 Plebanek, M. P. *et al.* Pre-metastatic cancer exosomes induce immune surveillance by patrolling monocytes at the metastatic niche. *Nat Commun***8**, 1319, doi:10.1038/s41467-017-01433-3 (2017).
- 40 Meylan, E., Tschopp, J. & Karin, M. Intracellular pattern recognition receptors in the host response. *Nature***442**, 39-44, doi:10.1038/nature04946 (2006).
- 41 Mildner, A. & Jung, S. Development and function of dendritic cell subsets. *Immunity***40**, 642-656, doi:10.1016/j.immuni.2014.04.016 (2014).
- 42 Mitchell, D., Chintala, S. & Dey, M. Plasmacytoid dendritic cell in immunity and cancer. *J Neuroimmunol***322**, 63-73, doi:10.1016/j.jneuroim.2018.06.012 (2018).
- 43 Spranger, S., Dai, D., Horton, B. & Gajewski, T. F. Tumor-Residing Batf3 Dendritic Cells Are Required for Effector T Cell Trafficking and Adoptive T Cell Therapy. *Cancer Cell***31**, 711-723 e714, doi:10.1016/j.ccell.2017.04.003 (2017).

44 Bol, K. F., Schreibelt, G., Gerritsen, W. R., de Vries, I. J. & Figdor, C. G. Dendritic Cell-Based Immunotherapy: State of the Art and Beyond. *Clin Cancer Res***22**, 1897-1906, doi:10.1158/1078-0432.CCR-15-1399 (2016).

45 Yuan, K. *et al.* Selective inhibition of CDK4/6: A safe and effective strategy for developing anti-cancer drugs. *Acta Pharm Sin B***11**, 30-54, doi:10.1016/j.apsb.2020.05.001 (2021).

46 Qin, A., Reddy, H. G., Weinberg, F. D. & Kalemkerian, G. P. Cyclin-dependent kinase inhibitors for the treatment of lung cancer. *Expert Opin Pharmacother***21**, 941-952, doi:10.1080/14656566.2020.1738385 (2020).

Figures

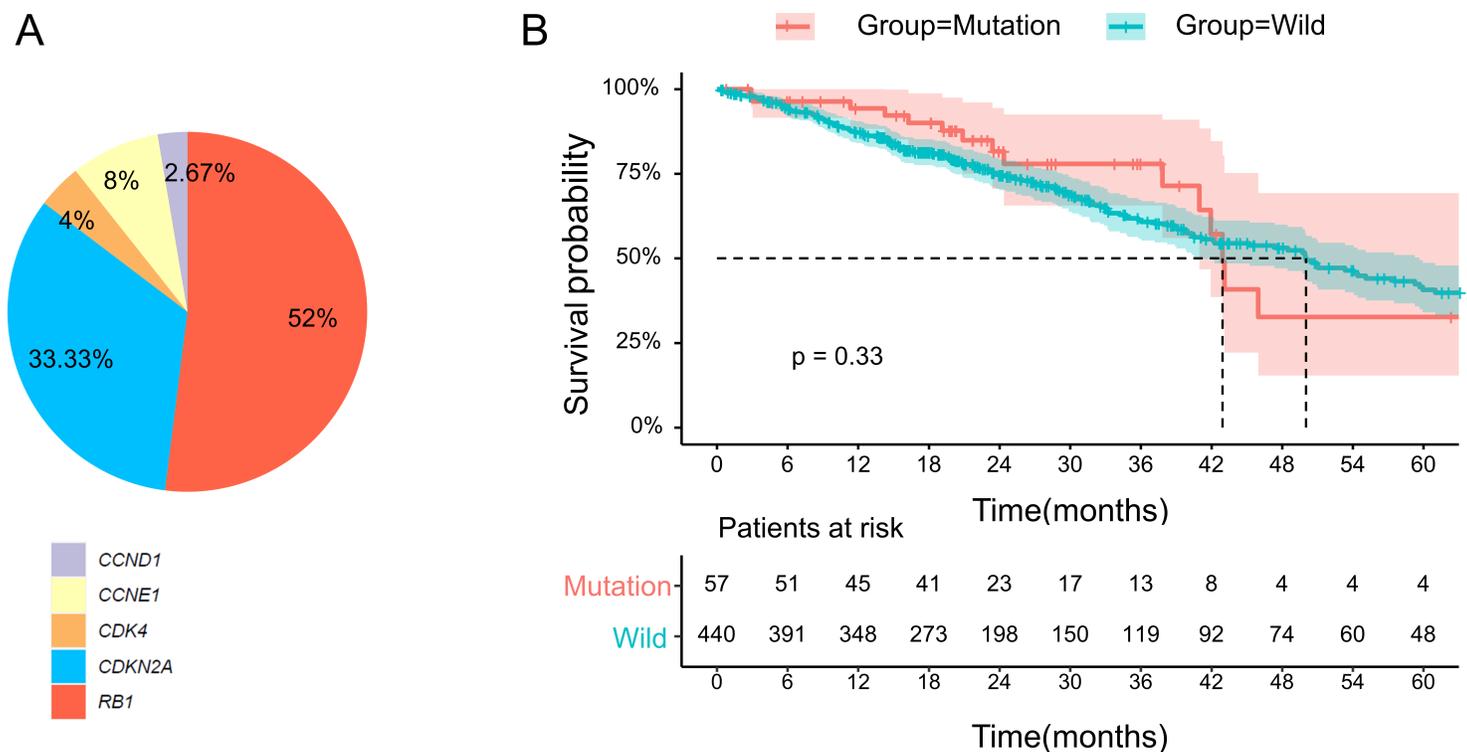


Figure 1

(A) Proportions of cell cycle progression pathway-related genes in the mutation group. (B) Survival curve of the mutation and wild group. $P < 0.05$ was considered as significantly different.

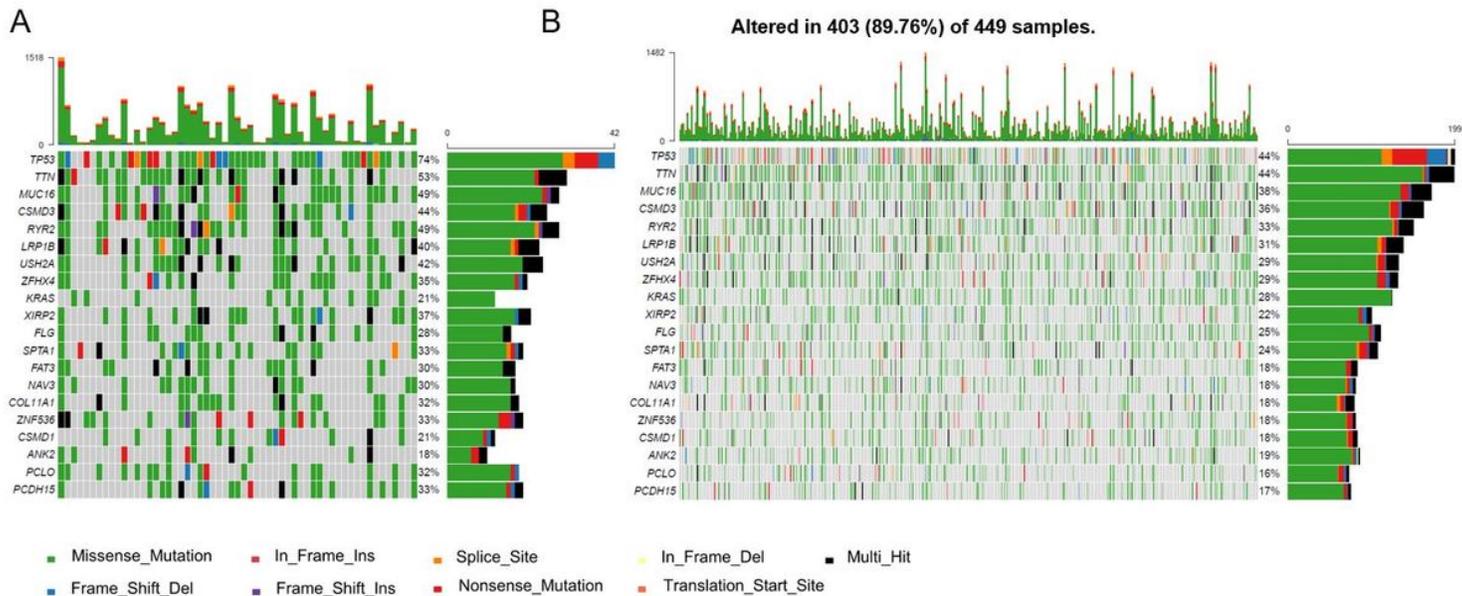


Figure 2

Waterfall maps of the top 20 mutant genes in the mutation (A) and wild (B) group.

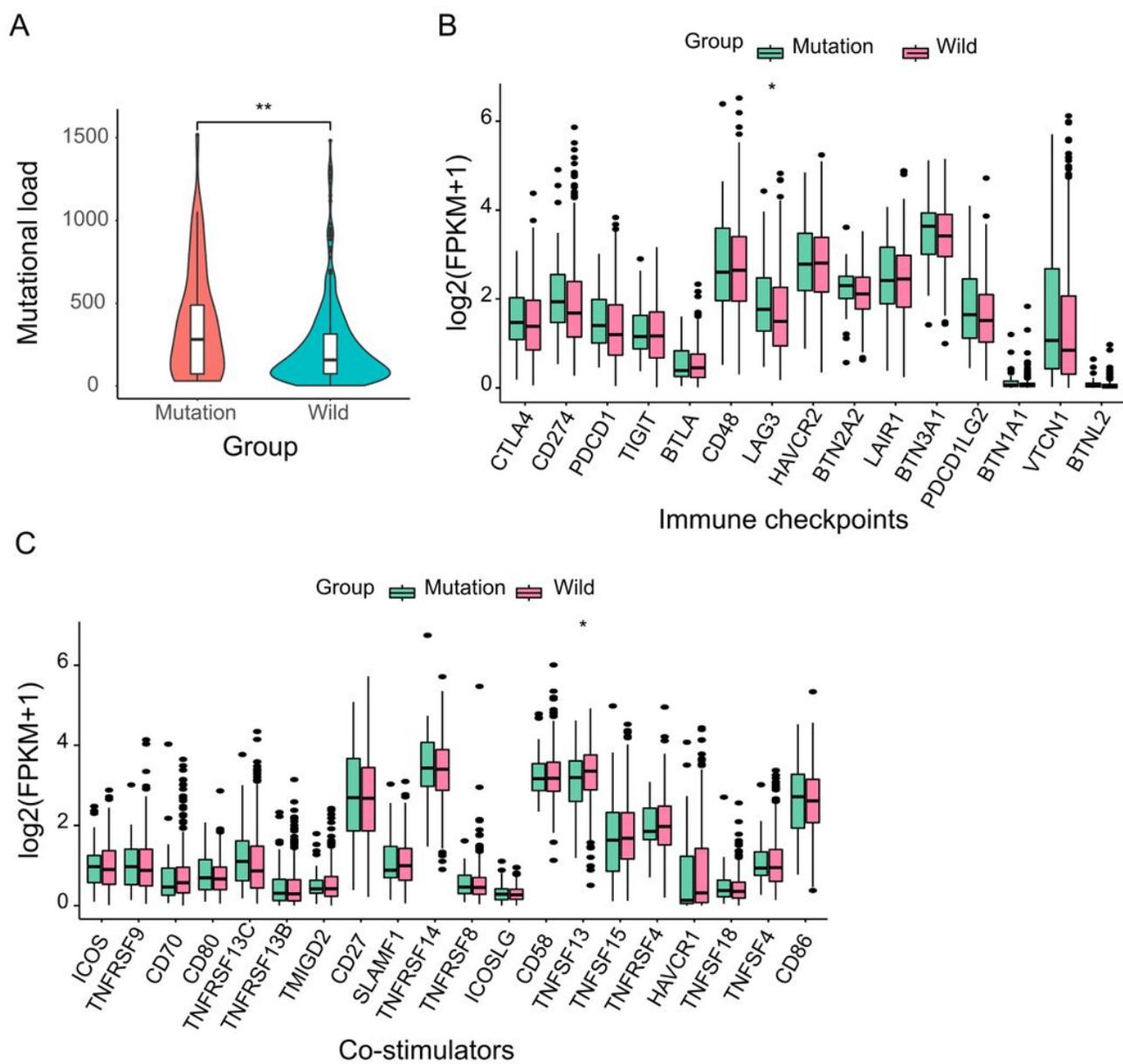


Figure 3

(A) Mutational load between the mutation and wild group. (B) Expression levels of immune checkpoints between the mutation and wild groups. (C) Expression levels of co-stimulators between the mutation and wild groups. $P < 0.05$ was considered as significantly different. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

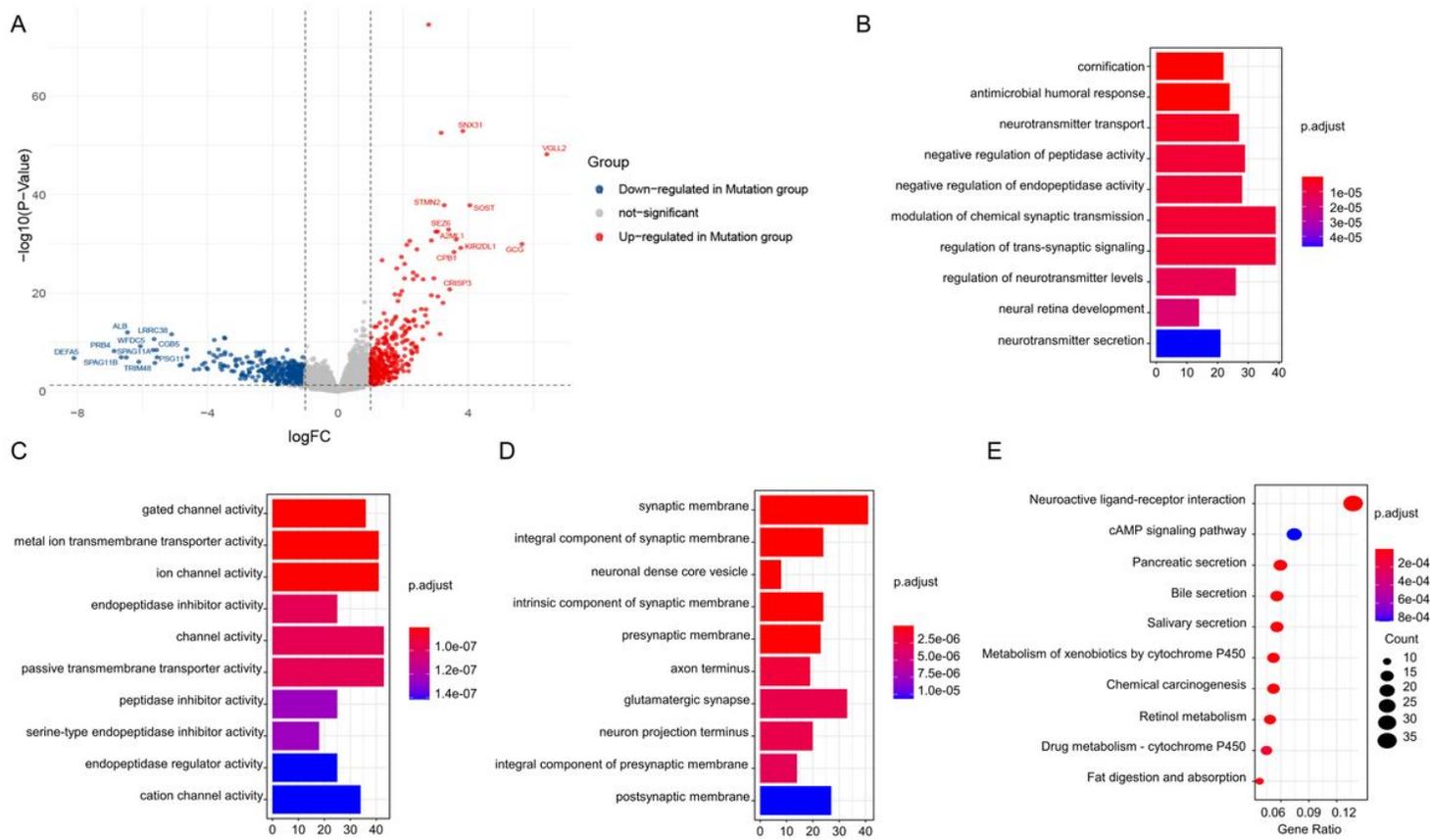


Figure 4

(A) Volcano map of DEGs between the mutation and wild group. Blue dots represented genes downregulated in the mutation group. Grey dots represent genes not significant between the two groups. Red dots represented genes upregulated in the mutation group. $P < 0.05$ was considered as significantly different. (B-D) Barplots of the top ten significantly different Biological Processes (BP), Molecular Function (MF), and Cellular Component (CC) by GO enrichment analysis. (E) Dot plot of the top ten significantly different enrichment pathways by KEGG analysis.

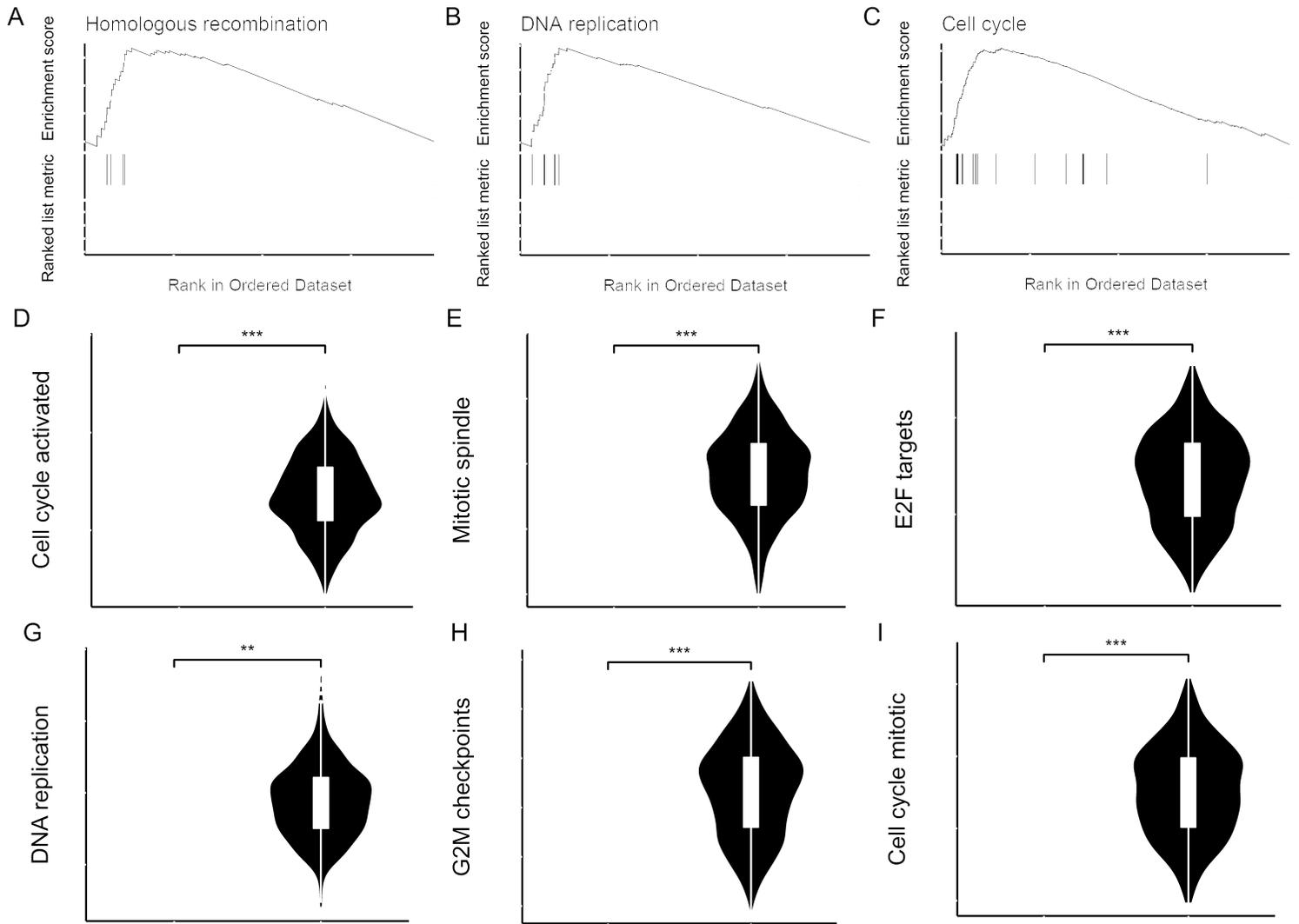


Figure 5

GSEA (A-C) and GSVA (D-I) of the DEGs between the mutation and wild groups.

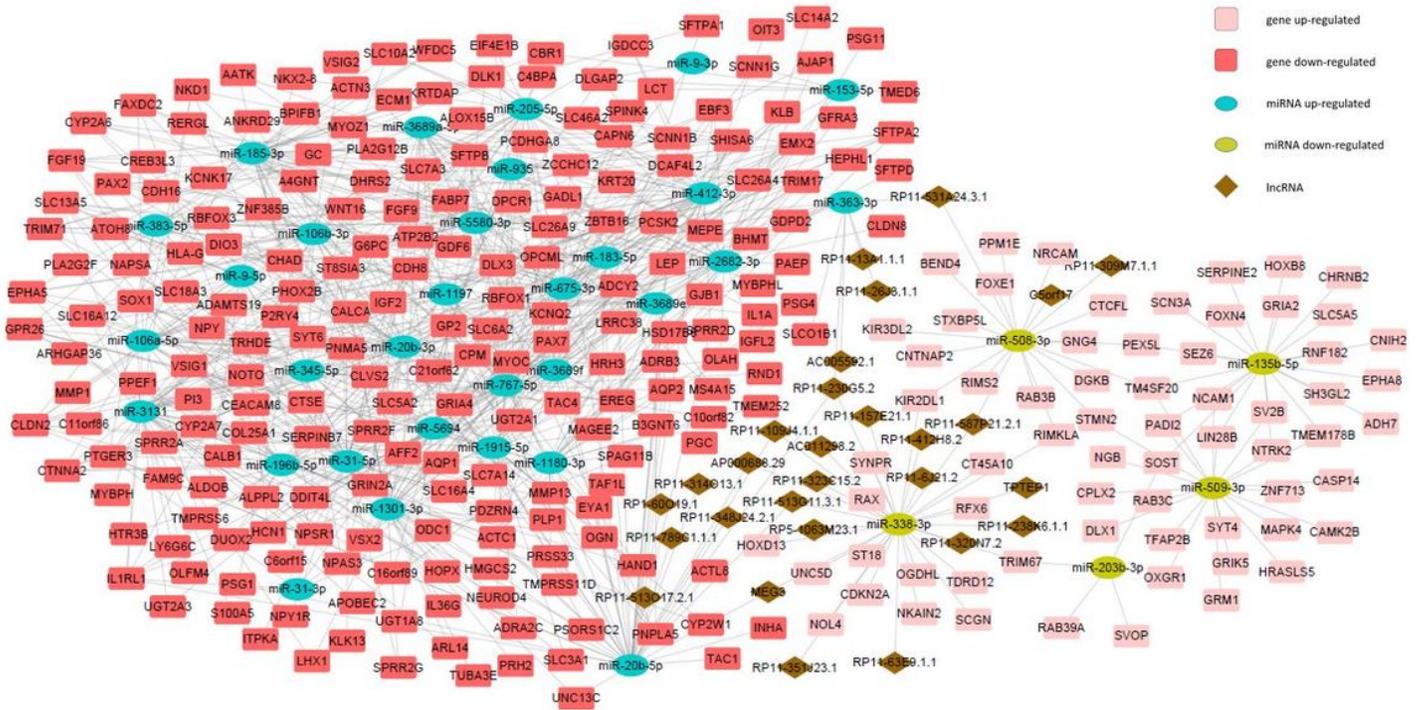


Figure 6

ceRNA network of DEGs - differentially expressed miRNAs - differentially expressed lncRNAs.

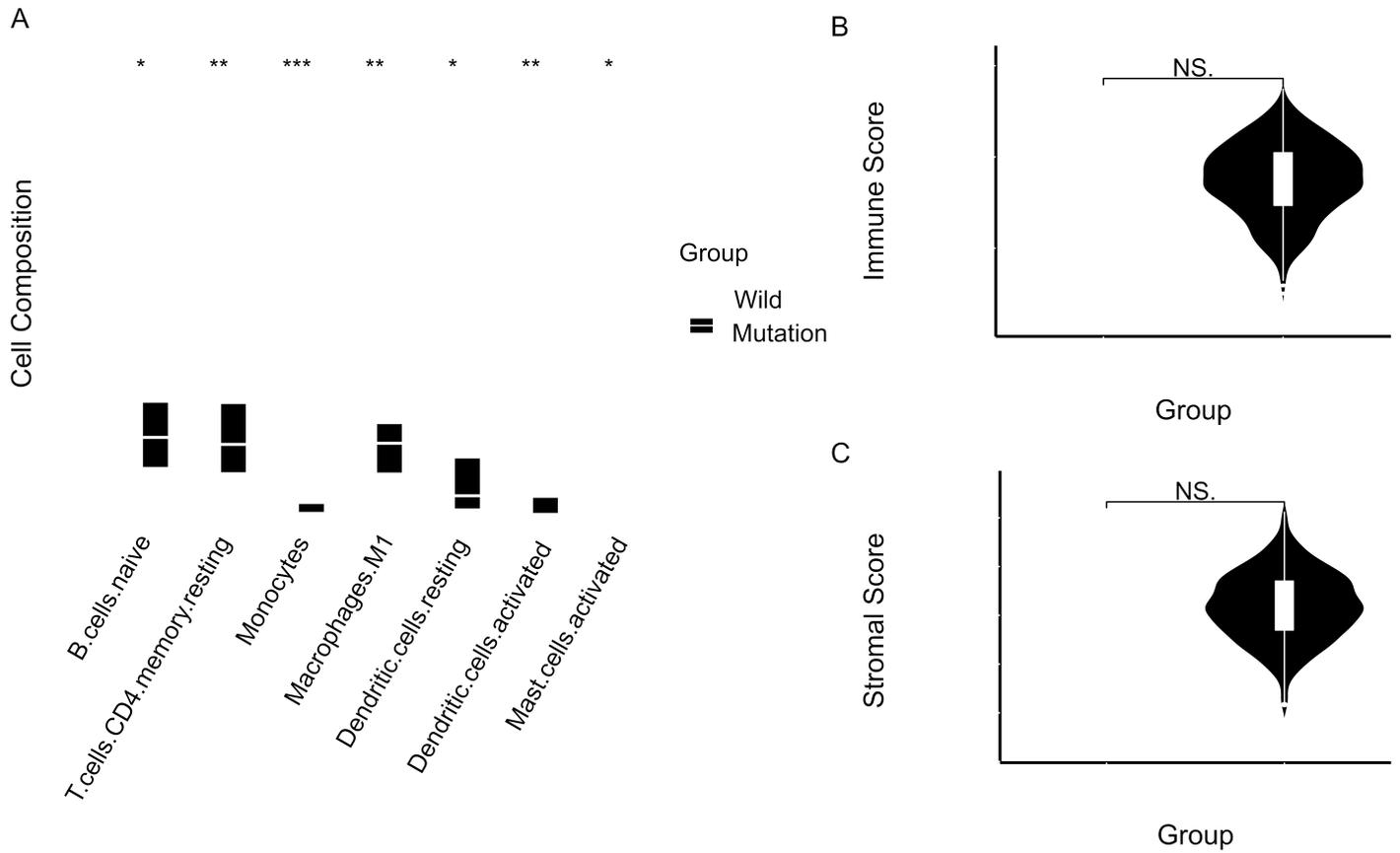


Figure 7

(A) Significantly different immune cell infiltration patterns between the mutation and wild group. $P < 0.05$ was considered as significantly different. (B-C) Stromal and Immune scores between the mutation and wild groups. NS, Not Significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

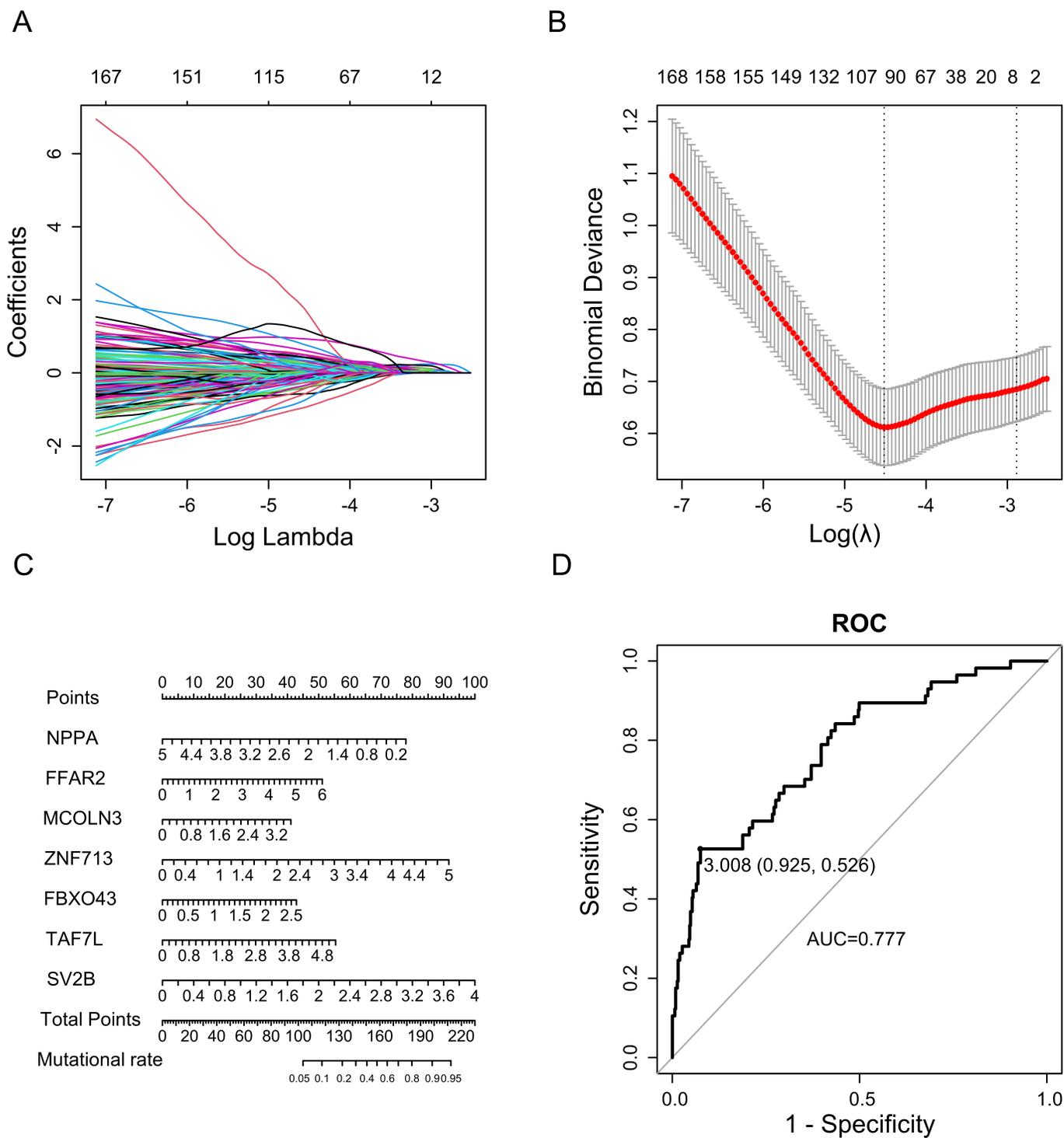


Figure 8

(A) Coefficient profiles of variables in the LASSO regression model. (B) Tenfold cross-validation for turning parameter (λ) selection in the LASSO regression model. The partial likelihood deviance is plotted in $\log(\lambda)$, in which vertical lines are shown at the optimal values by minimum criteria and 1-SE criteria. (C) Nomogram of the predictive model. (D) ROC of the LASSO-binary logistic regression model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)
- [SupplementaryFigure1.tif](#)
- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure3.tif](#)
- [SupplementaryFigure1.tif](#)
- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure3.tif](#)