Research Paper

KIAA0101 as a new diagnostic and prognostic marker, and its correlation with gene regulatory networks and immune infiltrates in lung adenocarcinoma

Sheng Hu[1], Weibiao Zeng[1], Wenxiong Zhang[1], Jianjun Xu[1], Dongliang Yu[1], Jinhua Peng[1], Yiping Wei[1]

[1]Department of Thoracic Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China

Correspondence to: Yiping Wei; **email:** weiyip2000@hotmail.com, https://orcid.org/0000-0002-5087-7113
Keywords: *KIAA0101/PCLAF*, lung adenocarcinoma, LUAD, biomarker, survival analysis
Received: June 25, 2020 **Accepted:** September 22, 2020 **Published:** November 20, 2020

ABSTRACT

Proliferating cell nuclear antigen binding factor (encoded by *KIAA0101/ PCLAF*) regulates DNA synthesis and cell cycle progression; however, whether the level of *KIAA0101* mRNA in lung adenocarcinoma is related to prognosis and tumor immune infiltration is unknown. In patients with lung adenocarcinoma, the differential expression of *KIAA0101* was analyzed using the Oncomine, GEPIA, and Ualcan databases. The prognosis of patients with different *KIAA0101* expression levels was evaluated using databases such as Prognostan and GEPIA. Tumor immune infiltration associated with *KIAA0101* was analyzed using TISIDB. Linkedmics was used to perform gene set enrichment analysis of *KIAA0101*. *KIAA0101* expression in lung adenocarcinoma tissues was higher than that in normal lung tissues. Patients with lung adenocarcinoma with low *KIAA0101* expression had a better prognosis than those with high *KIAA0101* expression. We constructed the gene regulatory network of *KIAA0101* in lung adenocarcinoma. *KIAA0101* appeared to play an important role in the regulation of tumor immune infiltration and targeted therapy in lung adenocarcinoma. Thus, *KIAA0101* mRNA levels correlated with the diagnosis, prognosis, immune infiltration, and targeted therapy in lung adenocarcinoma. These results provide new directions to develop diagnostic criteria, prognostic evaluation, immunotherapy, and targeted therapy for lung adenocarcinoma.

INTRODUCTION

Lung cancer remains one of the most common cancers in men worldwide [1, 2]. In the United States, between 2008 and 2014, the 5-year relative survival rate for lung cancer was only 19% [2]. Lung cancer includes multiple subtypes, and the proportion of lung adenocarcinoma (LUAD) has increased in recent years. Despite significant improvements in chemotherapy and molecular targeted therapy, the survival rate of LUAD is still unsatisfactory. Tumor recurrence and metastasis are major problems in the clinical treatment of LUAD [3]. Drug resistance to common gene targeted drugs and therapeutic effect differences between different patients can occur [4]. Pabla, Conroy [5] found that many genes represented by *KIAA0101* are immune related genes of adenocarcinoma, and *KIAA0101* and other genes might affect the presentation of tumor immune antigens. However, the specific relationship between *KIAA0101* and immune escape is not clear. In this study, we found that the expression of *KIAA0101* correlated negatively with a proportion of tumor infiltrating lymphocytes, and the expression of *KIAA0101* was correlated negatively with MHC molecule expression. Therefore, we carried out a study on the relationship between the expression of *KIAA0101* and immune escape in lung adenocarcinoma, and further studied the effect of *KIAA0101* expression on immunotherapy of lung cancer. Our study showed that overexpression of *KIAA0101* promotes immune escape in lung

adenocarcinoma. Therefore, searching for novel molecular biomarkers and improving immunotherapy of tumors in the diagnosis and treatment of LUAD are important [6].

KIAA0101 is a proliferating cell nuclear antigen (PCNA) binding factor discovered by Yu et al. in 2001 [7] using yeast two-hybrid experiments. It has different functions to other binding factors of PCNA, such as p21 and p57, in that KIAA0101 does not inhibit DNA replication and cell cycle process. PCNA-binding proteins act as regulators of DNA repair during DNA replication. Following DNA damage, they also facilitate the bypass of replication-fork-blocking lesions. KIAA0101 also acts as a regulator of centrosome number. However, the role of KIAA0101 in lung adenocarcinoma has not been determined. The protein encoded by *KIAA0101* contains eight domain chains (Supplementary Table 1) [8]. We noted that *KIAA0101* expression in lung adenocarcinoma was confirmed using an antibody in The Human Protein Atlas (Supplementary Figure 1) [9]. The expression of *KIAA0101* in the nucleus was significantly increased and slightly enhanced in the cytoplasm. The expression of *KIAA0101* increased in a variety of cancer tissues and cells, especially in liver cancer [10], breast cancer [11, 12], gastric cancer [13], and other cancers.

The present study aimed to explore the relationship between *KIAA0101* expression and the prognosis and immune infiltration of LUAD. We also searched for *KIAA0101*-related gene regulatory networks. *KIAA0101* expression was analyzed via the Oncomine, GEPIA, and Ualcan databases. The prognosis of LUAD related to *KIAA0101* was evaluated using prognostan and other databases. Analysis of tumor immune infiltration related to *KIAA0101* was analyzed using TISIDB. Gene set enrichment analysis was performed using Linkedmics. We hoped to provide new biomarkers related to *KIAA0101* for future targeted therapy of LUAD. In addition, we assessed whether we could predict the survival of patients with LUAD based on the expression of *KIAA0101*. The discovery of the gene regulatory network of *KIAA0101* will open new avenues for further studies on the epigenetics of *KIAA0101*. The study of immune infiltration related to *KIAA0101* will provide new directions for immunotherapy of lung adenocarcinoma.

RESULTS

Differential expression of *KIAA0101* in lung adenocarcinoma

Figure 1A shows a heat map of multiple genes that were overexpressed in LUAD. The top 23 ranked differentially expressed genes were obtained based on the log2 median centered ratio. After a genome-wide analysis, we identified these 23 genes, including *KIAA0101*, which were overexpressed in LUAD (Figure 1A, red text genes). We used Oncomine to analyze the expression distribution of *KIAA0101* mRNA in different tumors (Figure 1B). Oncomine collected these datasets from public repositories such as Gene Expression Omnibus (GEO) and Array Express. The results showed that when comparing cancer samples with normal samples, *KIAA010*1 was overexpressed in 17 LUAD datasets, but was not under-expressed in any of the LUAD datasets (Figure 1B). The thresholds were a P value less than 0.001, fold change greater than two, and gene rank in the top 10%. All results were statistically significant. The results were analyzed statistically using Student's t test. The mRNA expression level of *KIAA0101* was significantly higher in LUAD than in normal tissues. Figure 1C1–1C8 show the data from Oncomine, GEPIA, and Ualcan comparisons of *KIAA0101* mRNA levels between eight different gene sets of adenocarcinoma and normal tissues. All the results showed that the expression level of *KIAA0101* in LUAD tissues was significantly higher than that in normal lung tissues, as assessed using Student's t test. Further analysis of various clinical features of LUAD in the TCGA showed that *KIAA0101* was highly transcribed. In subgroup analysis based on stage, race, sex, age, smoking habit, and metastasis in the UALCAN database, the transcription level of *KIAA0101* in patients with LUAD was significantly higher than that in healthy people (Supplementary Figure 14A–14F). Therefore, the expression of *KIAA0101* could be used as a potential diagnostic marker for LUAD [14].

KIAA0101 mRNA levels predict prognosis in patients with lung adenocarcinoma

Our next analysis showed that the expression of *KIAA0101* had a significant effect on the prognosis of LUAD (Table 1). Figure 2A1–2A6 show the overall survival curve of five datasets (Jacob-00182-CANDF, MICHIGAN-LC, Jacob-00182-MSK, GSE13213, GSE31210) and the relapse free survival curve of GSE31210 datasets. (Figure 2A1–2A6) show six survival curves representing the five different data sets in Table 1 (from the PrognoScan[15] database). We could observe the blue curves (low expression) were higher than the red curves (high expression) in the Kaplan-Meier plots (Figure 2A1–2A6). In addition, we could see from Table 1 that all six Hazard Ratios (HR) were greater than 1.4, the HR [95% confidence interval (CI)-low CI-up] range was greater than one. The Cox p-values were less than 0.05, indicating statistical significance. The results showed that

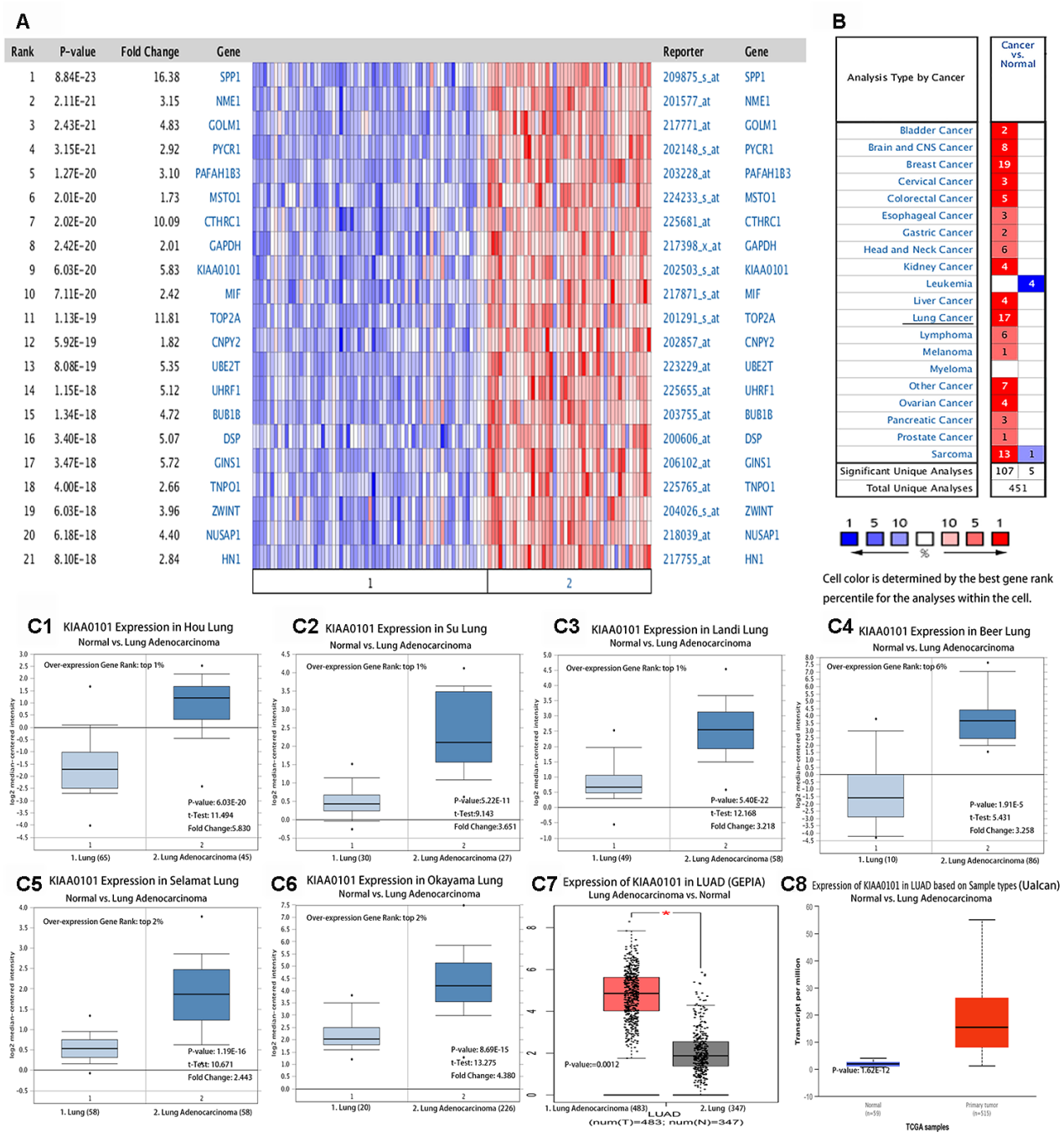


Figure 1. (**A**) Multigene view of lung adenocarcinoma Heat Maps Comparison of all Genes in the study of Hou lung [50]. The expression of *KIAA0101* in lung adenocarcinoma was higher than that in the normal control group. (**B**) Summary view of *KIAA0101*. The transcription level of *KIAA0101* in different types of cancer. Parameter setting: gene: *KIAA0101*, threshold (P-value): 0.001, threshold (fold change):2, threshold (gene rank): 10%, data type: DNA and mRNA. Note: The color is standardized by the Z-score to describe the relative value in the row. They cannot be used to compare values between rows. Among them, Red signifies gene overexpression or copy gain in the analyses represented by that cell in the table; blue represents the gene's underexpression or copy loss in those analyses. Datasets comprised samples represented as microarray data measuring either mRNA expression on primary tumors, cell lines, or xenografts. (**C**) Transcription of *KIAA0101* in lung adenocarcinoma (from Oncomine, GEPIA, and Ualcan). mRNA expression levels of *KIAA0101* were significantly higher in lung adenocarcinoma than in normal tissue. (**C1**–**C6**) Shown are the fold change, associated p-values, and overexpression Gene Rank, based on Oncomine 4.5 analysis. Box plot showing *KIAA0101* mRNA levels in, respectively, the Hou Lung, Su Lung, Landi Lung, Beer Lung, Selamat Lung and Okayama Lung datasets [50]. (**C7**) Expression of *KIAA0101* in LUAD based on GEPIA analysis; the p-value was 0.0012. (**C8**) Shows the expression of *KIAA0101* in LUAD based on Ualcan analysis; the p-value was 1.62E-12.

Table 1. Survival analysis of *KIAA0101* mRNA in lung adenocarcinoma patients (From PrognoScan).

	DATASET	PROBE ID	ENDPOINT	Number	Ln (HR high /HR low)	COX P-VALUE	Ln (HR)	HR [95% CI-low CI-up]
A1	Jacob-00182-CANDF	202503_s_at	Overall Survival	82	1.03	0.037184	0.57	1.76 [1.03 - 2.99]
A2	MICHIGAN-LC	D14657_at	Overall Survival	86	0.96	0.044546	0.43	1.54 [1.01 - 2.35]
A3	Jacob-00182-MSK	202503_s_at	Overall Survival	104	1.12	0.001309	0.7	2.01 [1.31 - 3.07]
A4	GSE13213	A_23_P117852	Overall Survival	117	1.24	0.004089	0.36	1.44 [1.12 - 1.84]
A5	GSE31210	202503_s_at	Overall Survival	204	1.93	0.000001	1.29	3.65 [2.16 - 6.17]
A6	GSE31210	202503_s_at	Relapse Free Survival	204	1.5	0	1	2.73 [1.90 - 3.91]

low expression of *KIAA0101* was associated with a good prognosis in patients with LUAD. Figure 2B1–2B6 show the overall survival curves of patients using data from the GEPIA [16], LinkedOmics [17], Ualcan [14], TISIDB [18], OncoLnc [19], and TCGA portal [20] databases. We found that patients with LUAD with low *KIAA0101* expression had a better prognosis than those with high *KIAA0101* expression in six the overall survival (OS) curves (P-value < 0.01). Figure 2C1 and 2C2 show the disease free survival curves (DFS) of patients using data from the GEPIA database. Figure 2D1 and 2D2 show the progression free survival curves (PFS) of patients using data from Kaplan Meier-plotter. We found that patients with LUAD with low *KIAA0101* expression had a better prognosis than those with high *KIAA0101* expression in four survival curves (P-value < 0.01). These curves reflect the fact that high expression of *KIAA0101* is associated negatively with the prognosis of patients with LUAD.

Gene set enrichment analysis of *KIAA0101* functional networks

Gene set enrichment analysis (GSEA) of KIAA0101 functional networks

(using LinkedOmics [17]). Table 2 shows the top five most significant of the seven GSEA results. All the results were statistically significant; the normalized enrichment scores (NES) were all greater than one; the P-values were all less than 0.05, and the false discovery rates (FDR) were all less than 0.05. Figure 3A shows a bar graph of the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Supplementary Figure 2A–2P display the enrichment plot of pathways in the enrichment results from Figure 3A, with the top 8 and the bottom 8 according to the normalized enrichment score. We observed that the top five pathways of the KEGG [21] enrichment analysis were the cell cycle (Figure 3B), ribosome (Supplementary Figure 3), proteasome (Supplementary Figure 4), spliceosome (Supplementary Figure 5), and DNA replication (Supplementary Figure 6). The gene ontology (GO) results for the gene sets (Cellular Component) included condensed chromosome, chromosomal region, mitochondrial protein complex, ribosome, and spindle. The GO results for the gene sets (Biological Process) included chromosome segregation, DNA replication, cell cycle checkpoint, double-strand break repair, and spindle organization. The GO results for the gene sets (Molecular Function) included structural constituent of ribosome, catalytic activity, acting on DNA, single-stranded DNA binding, helicase activity and catalytic activity, and acting on RNA. Kinase targets included CDK1 (cyclin-dependent kinase 1), PLK1 (polo-like kinase 1), AURKB (aurora kinase B), CDK2 (cyclin-dependent kinase 2), and ATR (ATR serine/threonine kinase). MicroRNA targets included MIR-484, MIR-512-3P, MIR-19A, MIR-19B, MIR-219, and MIR-326. Transcription factor targets included E2F Q6, E2F Q4, E2F4DP1 01, E2F1 Q6, and E2F 02. Supplementary Tables 2–8 show the details of all Leading Edge Genes specifically contained in each gene set.

Pearson correlated genes, miRNAs and lncRNAs of *KIAA0101*

The differentially expressed genes correlate highly with KIAA0101 in lung adenocarcinoma (LinkedOmics)

Figure 4A shows a Volcano plot of the results of the Pearson correlation test of *KIAA0101* in LUAD. The red and green dots in Figure 4A represent the differentially expressed genes associated with *KIAA0101*. Figure 4B and 4C show heat maps of the top 50 genes with strong positive and negative correlations with *KIAA0101* in LUAD. Figure 4D and 4E show the network views summarizing the predicted association network of proteins strongly related to *KIAA0101*. Supplementary Figure 7A–7P show scatter diagrams of *KIAA0101* and the first 16 genes that are highly related to *KIAA0101*. From these figures, we observed that the P-value was less than 1E-118; the FDR (BH) was less than 1E-113; and the Pearson correlation coefficient was > 0.8 (Figure 7A–7H), the Pearson correlation coefficient was <-0.6 (Figure 7I–7P). All the genes in the map had a strong positive correlation with *KIAA0101*.

miRNAs associated with KIAA0101 in patients with LUAD and their prognostic significance

We performed a miRNA Pearson correlation test and a COXPH test of *KIAA0101* (Figure 5). From the Volcano plot of *KIAA0101*-related miRNAs (Figure 5A), we found that the miRNAs hsa-mir-421, hsa-mir-133a-1, hsa-mir-128-2, hsa-mir-101-2, and hsa-mir-664 correlated strongly with *KIAA0101*. We identified the positive and negative miRNAs of *KIAA0101* in LUAD (top 50) from the miRNA heat plots of *KIAA0101*

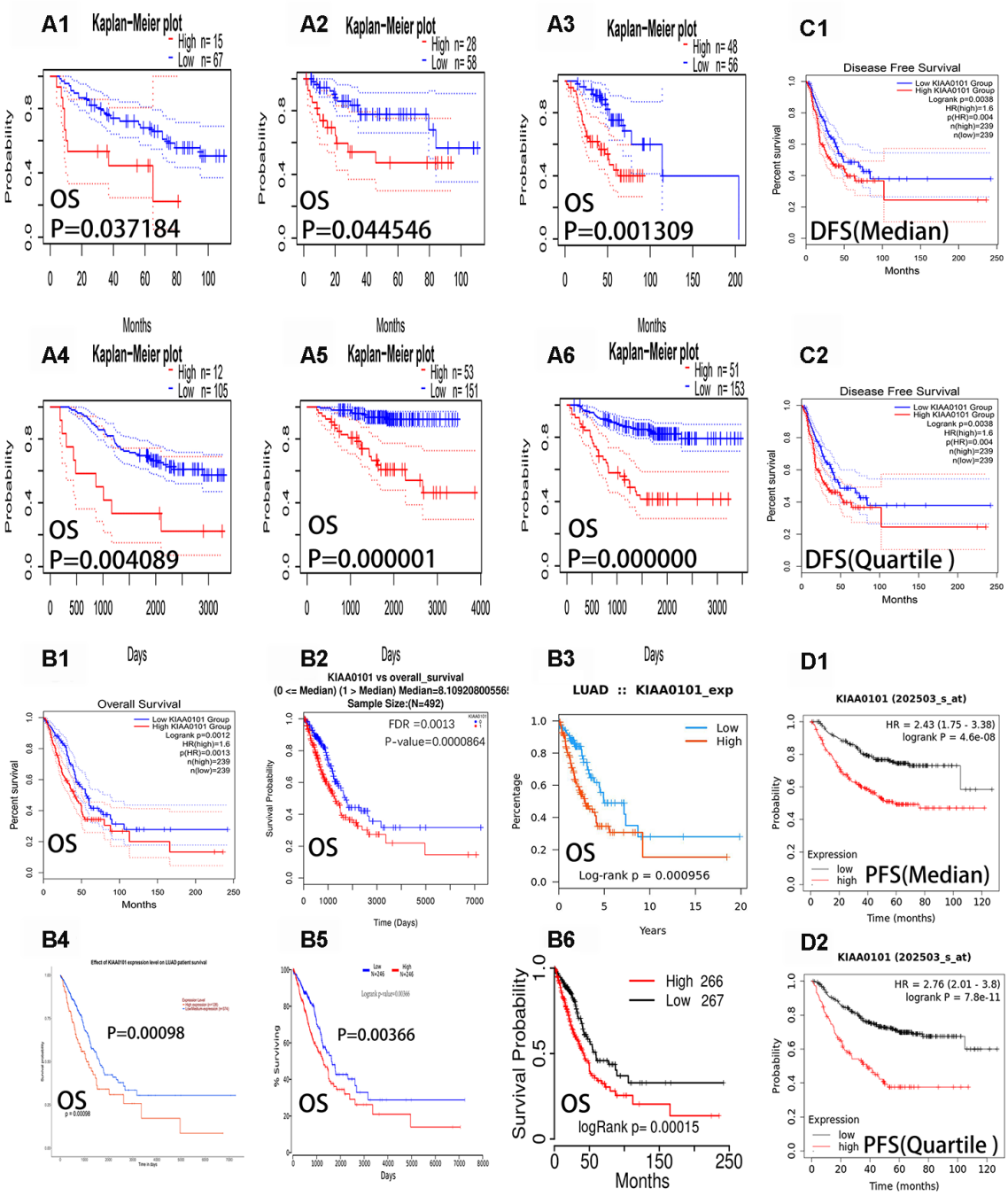


Figure 2. Overall survival curves, progression-free survival curves and disease-free survival curves of *KIAA0101* in lung adenocarcinoma. The blue curves represent patients with lung adenocarcinoma with low expression of *KIAA0101*, and the red curves represent patients with lung adenocarcinoma with high expression of *KIAA0101*. (**A1**–**A6**) Six survival curves representing the six different data sets in Table 1 (from PrognoScan databases), respectively. (**B1**–**B6**) The six overall survival curves from the GEPIA, Linkedmics, Ualcan, TISIDB, Oncolnc, and TCGA portal databases, respectively. (**C1**–**C2**) Disease free survival curves (DFS) of *KIAA0101* from the GEPIA database. (**D1**–**D2**) Progression free survival curves (PFS) of *KIAA0101* from Kaplan Meier-plotter.

Table 2. Gene Ontology (GO) (Cellular Component), GO (Biological Process), GO (Molecular Function), Kyoto Encyclopedia of Genes and Genomes (KEGG pathway), Kinase, miRNA, and transcription factor-target networks of *KIAA0101* in lung adenocarcinoma (LinkedOmics).

Enriched Category	Geneset	LeadingEdgeNum	NES	P Value	FDR
GO (Cellular Component)	condensed chromosome	66	2.6425	0	0
	chromosomal region	94	2.5825	0	0
	mitochondrial protein complex	152	2.3976	0	0
	ribosome	143	2.3858	0	0
	spindle	70	2.3837	0	0
GO (Biological_Process)	chromosome segregation	97	2.6361	0	0
	DNA replication	96	2.4830	0	0
	cell cycle checkpoint	60	2.3720	0	0
	double-strand break repair	61	2.3053	0	0
	spindle organization	44	2.2998	0	0
GO (Molecular_Function)	structural constituent of ribosome	109	2.3582	0	0
	catalytic activity, acting on DNA	69	2.2380	0	0
	single-stranded DNA binding	43	2.1594	0	0
	helicase activity	43	2.0895	0	0
	catalytic activity, acting on RNA	122	2.0460	0	0
KEGG pathway	Cell cycle	48	2.4639	0	0
	Ribosome	100	2.2904	0	0
	Proteasome	40	2.2761	0	0
	Spliceosome	63	2.2601	0	0
	DNA replication	27	2.1887	0	0
Kinase Target	Kinase_CDK1	74	2.3361	0	0
	Kinase_PLK1	32	2.3154	0	0
	Kinase_AURKB	35	2.1660	0	0
	Kinase_CDK2	73	2.1657	0	0
	Kinase_ATR	20	2.0601	0	0
miRNA Target	GAGCCTG,MIR-484	40	-–1.7728	0	0.025589
	CAGCACT,MIR-512-3P	53	–1.6953	0	0.029684
	TTTGCAC,MIR-19A,MIR-19B	148	–1.7059	0	0.034546
	GACAATC,MIR-219	60	–1.7191	0	0.039237
	CCCAGAG,MIR-326	58	–1.6503	0	0.039481
Transcription Factor Target	V$E2F_Q6	87	2.2583	0	0
	V$E2F_Q4	86	2.2538	0	0
	V$E2F4DP1_01	94	2.2226	0	0
	V$E2F1_Q6	96	2.2207	0	0
	V$E2F_02	93	2.1958	0	0

Abbreviations: Leading EdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA); NES, Normalized Enrichment Score. It is generally considered that the absolute value of NES ≥ 1.0, NOM p-val ≤ 0.05, FDR q-val ≤ 0.25 are significant gene sets. The annotation was found in the Molecular Signatures Database (MSigDB) for transcription factors (TF).

(Figure 5B, 5C). Supplementary Figure 8 shows the scatter plots of positive (Supplementary Figure 8A1–8A10) and negative (Supplementary Figure 8B1–8B10) correlations between the expression of *KIAA0101* and miRNAs. Their P-values were less than 1E-14, and their FDRs (BH) were less than 1E -12, thus the results were statistically significant. We constructed a Volcano plot of *KIAA0101* correlated miRNAs with significance for patient survival in lung adenocarcinoma (Figure 5D). One hundred miRNAs that correlated positively and negatively with the prognosis of patients with LUAD were identified from Figure 5E and 5F. Figure 5G1–5G12 show the *KIAA0101*-related miRNAs with low expression that were associated with good prognosis. Figure 5G13–5G18 show the *KIAA0101*-related miRNAs with high expression that were associated with good prognosis. The P values were less than 0.001 and the COXPH

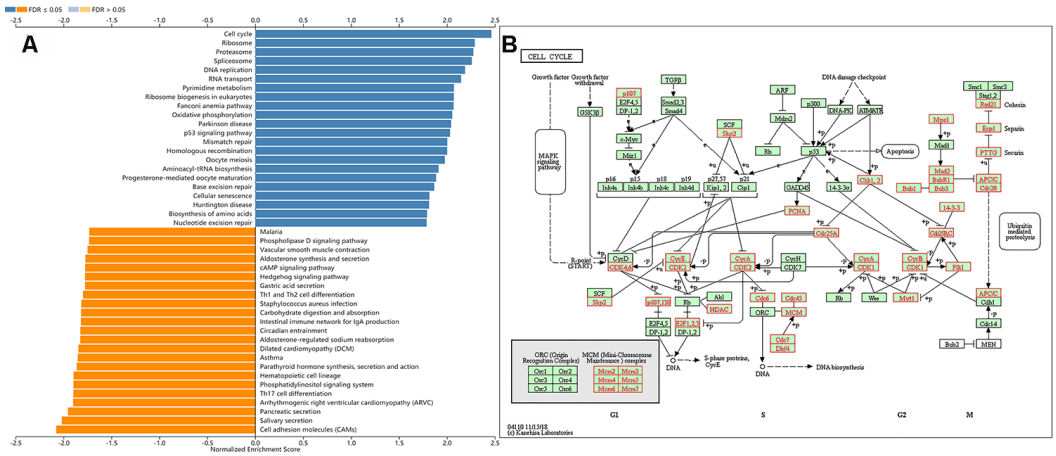


Figure 3. Gene set enrichment analysis of the *KIAA0101* via the KEGG pathway. (**A**) Bar chart for gene set enrichment analysis of the *KIAA0101* via the KEGG pathway. (**B**) KEGG pathway annotations of the cell cycle pathway (hsa04110). Red denotes leading edge genes; green denotes the remaining genes.

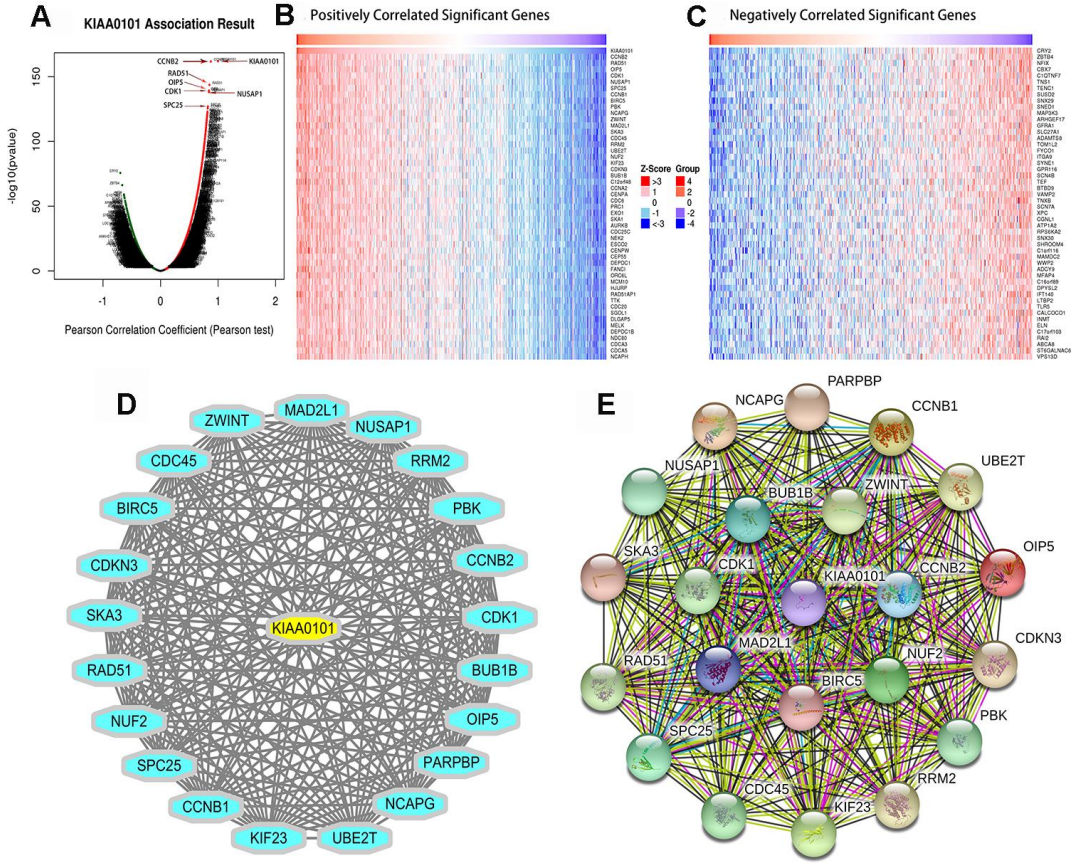


Figure 4. *KIAA0101* highly correlated genes. (**A**) Pearson test was used to analyze the correlation between *KIAA0101* and differentially expressed genes in LUAD. (**B**, **C**) Heat maps showing the genes that correlated positively and negatively with *KIAA0101* in LUAD (top 50). Red indicates positively correlated genes, green indicates negatively correlated genes. (**D**, **E**) The network view summarizes the predicted association network of proteins that are strongly related to the protein product of *KIAA0101*. Network nodes are proteins. The edge represents the functional association of the prediction. A red line indicates the existence of fusion evidence, a green line represents neighborhood evidence, a blue line represents co-occurrence evidence, a purple line represents experimental evidence, a yellow line represents text mining evidence, and a light blue line represents database evidence.

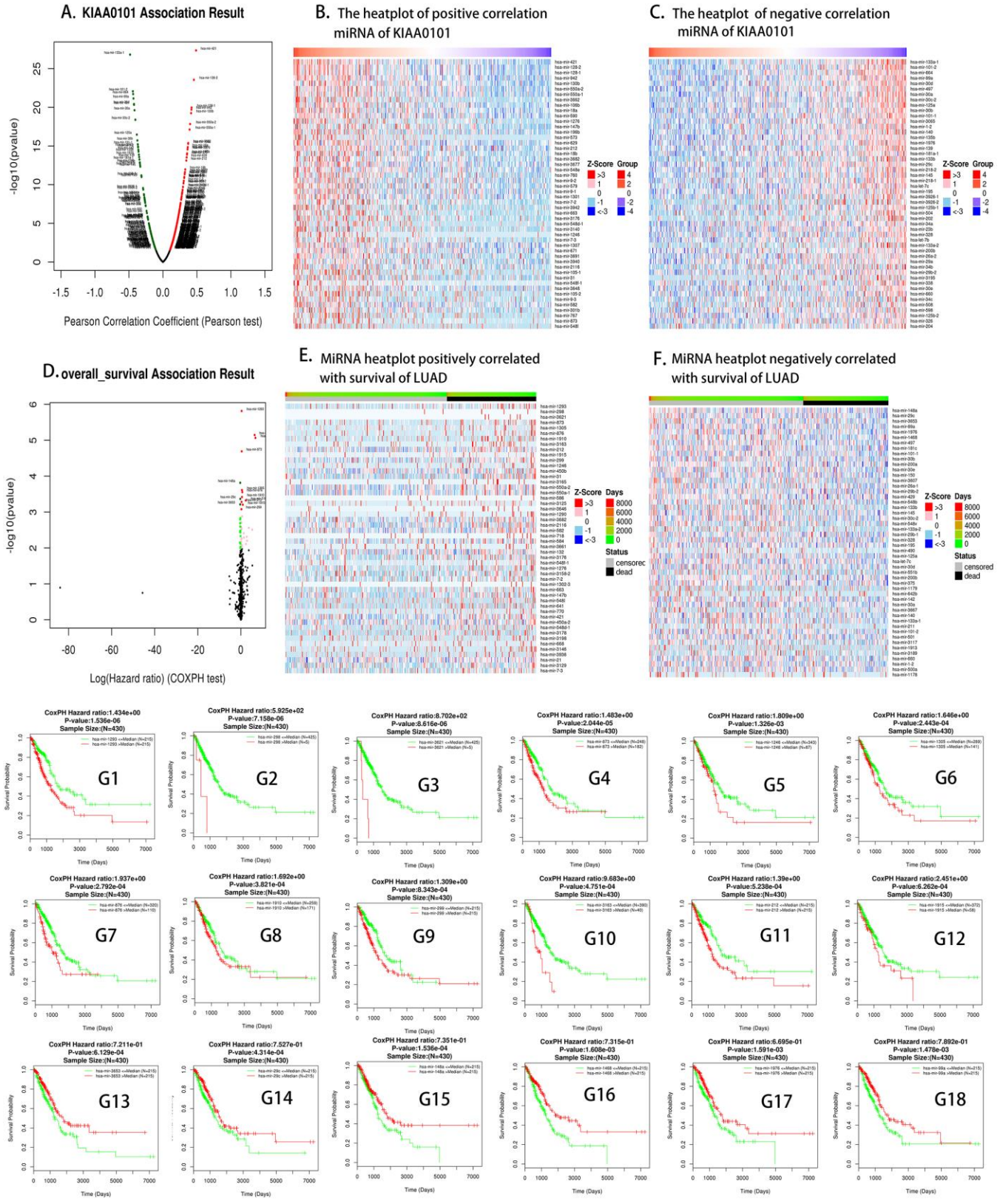


Figure 5. *KIAA0101* highly correlated miRNAs. (**A**) Volcano plot of *KIAA0101* related miRNAs; (**B**, **C**) positively and negatively correlated significant miRNA heat plots of *KIAA0101*. (**D**) The miRNA volcano map related to the overall survival of LUAD; (**E**, **F**) miRNAs positively and negatively related to the overall survival of LUAD, respectively; (**G1**–**G18**) the survival curves of *KIAA0101* related miRNAs; green represents low expression of the corresponding miRNA, while red represents high expression of the corresponding miRNA.

hazard ratios were greater than one, indicating that the results were statistically significant.

Long noncoding RNAs (lncRNAs) that correlated with KIAA0101 and were significant for patient survival

Figure 6 shows scatter plots of the lncRNAs that correlated positively (Figure 6A1–6A8) and negatively (Figure 6B1–6B8) with *KIAA0101* expression. The results were statistically significant. Figure 6C1–6C8 show the OS curves of lncRNAs related to *KIAA*0101 expression, in which the Cox P-value of all lncRNAs were less than or equal to 0.001 and the log-rank P-value of all lncRNAs were less than or equal to 0.002. Figure 6C1–6C8 show the *KIAA0101*-related lncRNAs with low expression that were associated with better prognosis of patients with LUAD. Figure 6D1–6D8 show the *KIAA0101*-related lncRNAs with low expression that were associated with better prognosis of patients with LUAD.

The relationship between immune infiltration and the expression of *KIAA0101* in lung adenocarcinoma

Figure 7A–7F show heat maps of the correlation between the expression of *KIAA0101* in multiple cancers and different tumor infiltrating lymphocytes (TILs), immunoinhibitors, immunostimulators, MHC molecules, chemokines, and receptors. Figure 7G1–7G18 show scatter diagrams of tumor infiltrating lymphocytes that correlated negatively with *KIAA0101* expression in LUAD. Figure 7H1–7H10 show scatter diagrams of TILs that correlated positively with *KIAA0101* expression in LUAD. When the expression of *KIAA0101* increased, the number of activated CD4 T cells increased. The numbers of Th1 cell, Th17 cells, and follicular helper cell were negatively related to the expression of *KIAA0101*, and Th2 cell numbers were positively related to the expression of *KIAA0101* ($P < 0.05$). *KIAA0101* expression correlated positively with the numbers of effector memory CD8 T cells, natural killer cells, regulatory T cells, activated B cells, immature B cells, plasmacytoid dendritic cells, macrophages, eosinophil, mast cells, and neutrophils (all $P < 0.05$). Supplementary Figures 9–13 show scatter plots of *KIAA0101* expression related to immunoinhibitors, immunostimulators, MHC molecules, chemokines, and chemokine receptors, respectively. Most chemokine receptors correlated negatively with the expression of *KIAA0101* (Supplementary Figure 13). The majority of MHC molecules correlated negatively with the expression of *KIAA0101* (Supplementary Figure 11).

We found that the metabolism of seven drugs upregulated the expression of *KIAA0101* (Table 3). They included Acetaminophen [22], Estradiol [23],

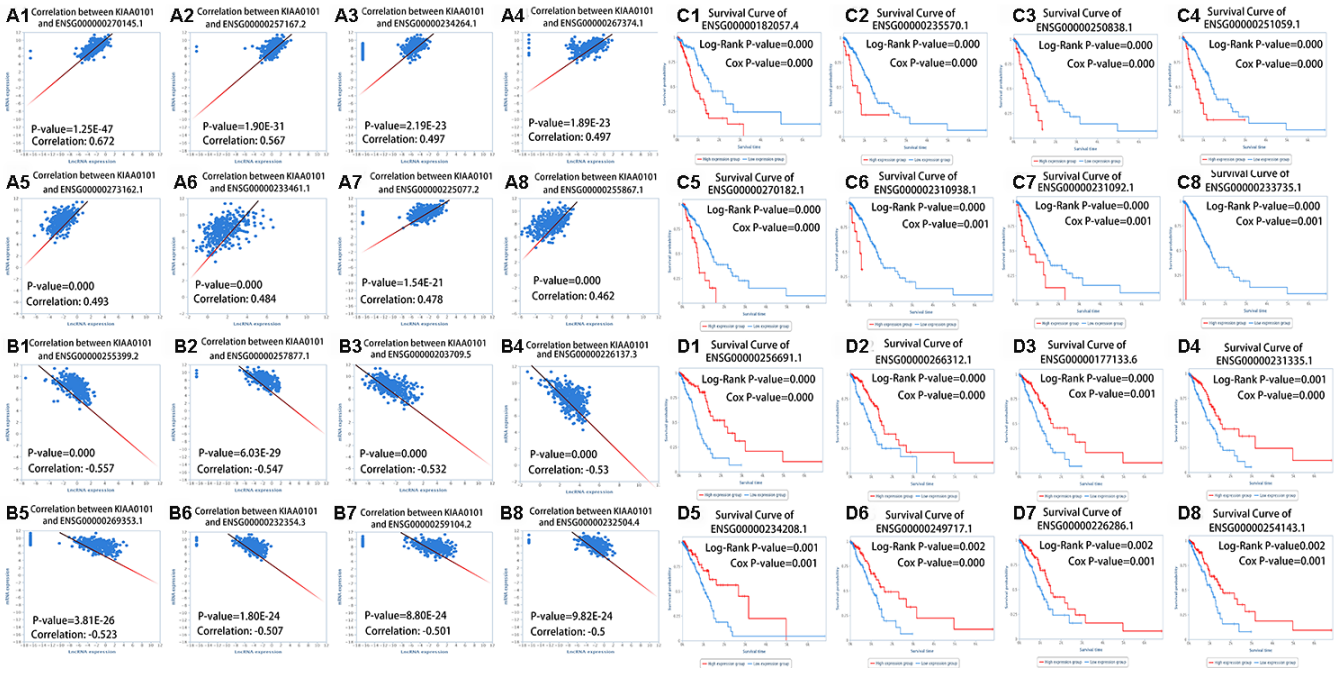


Figure 6. *KIAA0101* highly correlated lncRNAs. (**A1**–**A8**) Scatter plots of eight kinds of lncRNAs that correlated positively with *KIAA0101* expression. (**B1**–**B8**) Scatter plots of eight kinds of lncRNAs that correlated negatively with *KIAA0101* expression. (**C1**–**C8**) Survival curves of eight kinds of lncRNAs; patients with low expression of lncRNAs have a higher survival rate. (**D1**–**D8**) Survival curves of eight kinds of lncRNAs; patients with high expression of lncRNAs have a higher survival rate.

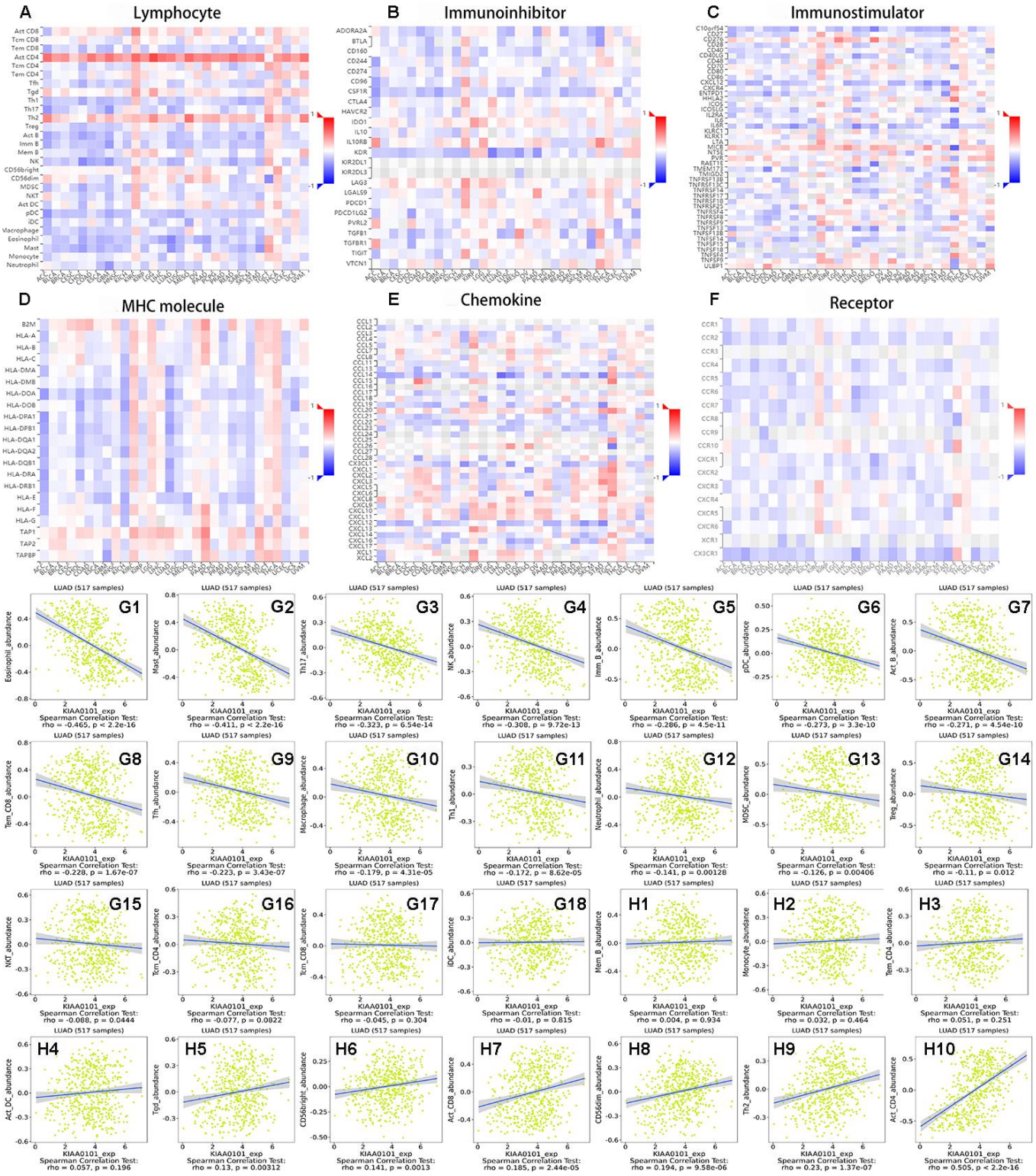


Figure 7. The relationship between immune infiltration and expression of *KIAA0101* in lung adenocarcinoma. (**A**–**F**) Heat maps of *KIAA0101* expression and lymphocytes, immunoinhibitors, immunostimulators, MHC molecules, chemokines, and receptors in different cancers. (**G1**–**G18**) are scatter plots of the negative correlation between *KIAA0101* expression and lymphocytes in the treatment of lung adenocarcinoma. (**H1**–**H10**) are scatter plots of the positive correlation between *KIAA0101* expression and lymphocytes in the treatment of lung adenocarcinoma.

Table 3. The Pharmaco-transcriptomics of *KIAA0101* (Up or downregulation of genes caused by the metabolism of pharmaceutical compounds).

DRUG	CHANGE	INTERACTION	REFERENCES (PMID)
Acetaminophen	upregulated	Acetaminophen results in increased expression of *KIAA0101* mRNA	25704631
Afimoxifene	downregulated	Afimoxifene results in decreased expression of *KIAA0101* mRNA	16514628
Estradiol	downregulated	Estradiol results in decreased expression of *KIAA0101* mRNA	23019147
Estradiol	upregulated	Estradiol results in increased expression of *KIAA0101* mRNA	16474171
Custirsen	downregulated	CC-8490 results in decreased expression of *KIAA0101* mRNA	15604281
Copper	downregulated	Binding with copper resulted in decreased expression of *KIAA0101* mRNA	20971185
Cyclosporine	downregulated	Cyclosporine results in decreased expression of *KIAA0101* mRNA	20106945
Dasatinib	downregulated	Dasatinib results in decreased expression of *KIAA0101* mRNA	20579391
Dronabinol	downregulated	Dronabinol results in decreased expression of *KIAA0101* mRNA	18454173
Calcitriol	downregulated	Calcitriol results in decreased expression of *KIAA0101* mRNA	21592394
Fluorouracil	downregulated	Fluorouracil results in decreased expression of *KIAA0101* mRNA	17039268
Fulvestrant	downregulated	fulvestrant results in decreased expression of *KIAA0101* mRNA	16514628
Genistein	downregulated	Genistein results in decreased expression of *KIAA0101* mRNA	19371625
Genistein	upregulated	Genistein results in increased expression of *KIAA0101* mRNA	16474171
Lucanthone	downregulated	Lucanthone results in decreased expression of *KIAA0101* mRNA	21148553
Metamfetamine	downregulated	Methamphetamine results in decreased expression of *KIAA0101* mRNA	25290377
Methotrexate	downregulated	Methotrexate results in decreased expression of *KIAA0101* mRNA	24657277
Palbociclib	downregulated	Palbociclib results in decreased expression of *KIAA0101* mRNA	22869556
Phenethyl Isothiocyanate	downregulated	Phenethyl isothiocyanate results in decreased expression of *KIAA0101* mRNA	26678675
Piroxicam	downregulated	Piroxicam results in decreased expression of *KIAA0101* mRNA	21858171
Progesterone	downregulated	Progesterone results in decreased expression of *KIAA0101* mRNA	21795739
Progesterone	upregulated	Progesterone results in increased expression of *KIAA0101* mRNA	18070364
Raloxifene	downregulated	Raloxifene Hydrochloride results in decreased expression of *KIAA0101* mRNA	16514628
Resveratrol	upregulated	Resveratrol results in increased expression of *KIAA0101* mRNA	19167446
Rifampicin	upregulated	Rifampin results in increased expression of *KIAA0101* mRNA	24552687
Silicon dioxide	downregulated	Silicon Dioxide analog results in decreased *expression* of KIAA0101 mRNA	25895662
Tamoxifen	downregulated	Tamoxifen results in decreased expression of *KIAA0101* mRNA	15604281
Testosterone	downregulated	Testosterone results in decreased expression of *KIAA0101* mRNA	21592394
Valproic acid	upregulated	Valproic Acid results in increased expression of *KIAA0101* mRNA	23179753
Vitamin E	downregulated	Vitamin E results in decreased expression of *KIAA0101* mRNA	19244175
Zoledronic acid	downregulated	Zoledronic acid results in decreased expression of *KIAA0101* mRNA	24714768

Genistein [23], Progesterone [24], Resveratrol [25], Rifampicin [26], and Valproic acid [27]. The metabolism of 24 drugs downregulated the expression of *KIAA0101*. They included amoxicifene, estradiol, yolk, copper, and cyclosporine (Table 3). The metabolism of Estradiol [23, 28], Genistein [23, 29], and Progesterone [24, 30] could both upregulate and downregulate the expression of *KIAA0101* (Table 3).

DISCUSSION AND CONCLUSIONS

Early detection of LUAD is a challenge for clinicians. Low dose computed tomography (LDCT) has the advantages of simplicity and high sensitivity. Compared with the traditional chest film and tumor markers, the false negative rate of LDCT was reduced [31]. However, more accurate and timely biomarkers are needed to support a diagnosis of LUAD. It has been reported that *KIAA0101* is differentially expressed in ovarian cancer [32], adrenal cancer [33], renal cell carcinoma [10], hepatocellular carcinoma [10], and breast cancer [12]. However, there are few studies about *KIAA0101* in LUAD [34]. The results of the present study showed that *KIAA0101* was significantly differentially expressed in all lung adenocarcinomas (Figure 1). This is very interesting; however, it requires verification in further comparative studies.

We further explored the mechanism by which *KIAA0101* can act as a marker of LUAD. Hosokawa et al. [35] found that if *KIAA0101* was knocked down using small interfering RNA (siRNA), the DNA replication rate was greatly reduced in pancreatic cancer cell lines. The exogenous overexpression of *KIAA0101* enhanced the growth of cancer cells, suggesting the carcinogenicity of *KIAA0101*. They also found that

inhibition of the *KIAA0101-PCNA* interaction significantly inhibited the growth of cancer cells. Lv et al. [11] found that knockdown of *KIAA0101* could promote the formation of p53 / SP1 complexes in breast cancer, and inhibited the proliferation and cell cycle process of breast cancer cells. Many studies have shown that *KIAA0101* is involved in the regulation of DNA replication and repair, cell cycle progression, and cell proliferation, and can reduce the apoptosis induced by UV [35–37]. Our results are consistent with these previous reports. KEGG enrichment analysis of *KIAA0101* identified the cell cycle, ribosome, proteasome, spliceosome, and DNA replication pathways as enriched. In LUAD, *KIAA0101* was related to kinase networks including CDK1, PLK1, AURKB, CDK2, and ATR. These kinases regulate cell cycle stability, chromosome separation, and the cell cycle during mitosis and meiosis [38, 39]. In fact, ATR is one of the key kinase regulators for genomic stability; it directly phosphorylates a variety of important substrates, including p53 protein and cyclin [40]. ATR kinase inhibitors can reduce the growth of tumor cells [41]. In addition, differentially expressed genes highly related to *KIAA0101* in lung adenocarcinoma were associated with the cell cycle, DNA replication and modification, mitosis, and meiosis. Thus, *KIAA0101* might regulate cell cycle progression and DNA replication and repair via the above kinases in LUAD [42]. Cell cycle progression and DNA replication occur mainly in the nucleus, which might also be the main reason for the markedly increased expression of *KIAA0101* in the nucleus and slightly enhanced expression in the cytoplasm.

Based on the above results (Figure 2), we found that the prognosis of patients with LUAD could be predicted by the expression of *KIAA0101*. We found that some drugs [43] could reduce the expression of *KIAA0101* (Table 3). The decrease in gene expression by these drugs [22] might be transient or concentration-dependent, and we will further evaluate the effect of different concentrations and times on *KIAA0101* expression in the future. Some drugs that upregulated the expression of *KIAA0101* might be potential carcinogens of LUAD. We also identified miRNAs and lncRNAs that are related to the expression of *KIAA0101* (Figures 5 and 6). The expression of these miRNAs and lncRNAs was related to the survival of patients with lung adenocarcinoma. Thus, these miRNAs and lncRNAs could also be used as potential biomarkers of lung adenocarcinoma [17, 44]. These miRNAs and lncRNAs could be further studied to improve the epigenetics of *KIAA0101* in lung adenocarcinoma.

The interaction between tumor cells and their surrounding matrix not only affects the occurrence and development of the disease, but also is closely related to the prognosis of patients [45]. Evidence suggests that TILs infiltrate tumor tissue and play a role in the disease by regulating anti-tumor immunity [46]. With the increase in *KIAA0101* expression, the numbers of Th1 cells, Th17 cells, and follicular helper cells decreased, and the number of Th2 cells increased. Type 1 T helper cells are positive immunoregulatory cells, and type 2 T helper cells are negative immunoregulatory cells. Thus, variation in *KIAA0101* expression leads to a shift of the Th1/Th2 ratio. The immunosuppressive state will seriously affect the anti-tumor immunity of the body, and eventually lead to the occurrence and development of tumors [47]. With the increase in *KIAA0101* expression, the number of effector memory CD8 T cells, natural killer cells, regulatory T cells, activated B cells, immature B cells, plasmacytoid dendritic cells, macrophages, eosinophil, mast cells, and neutrophils will decrease. This finding confirms the view that increased expression of *KIAA0101* reduces survival in patients with LUAD. Chemokine receptors are seven transmembrane G protein-coupled receptors. Their main function is to receive chemokine signals and further conduct signals, which plays an important positive role in promoting tumor peripheral immune infiltration [48]. The main function of MHC molecules is to participate in antigen presentation. MHC molecules play an important role in the differentiation and maturation of T cells, and play an important positive role in tumor peripheral immune infiltration [49]. When the expression of *KIAA0101* increased, the levels of most chemokine receptors and MHC molecules in patients decreased (Supplementary Figures 11, 12), which would further reduce the tumor peripheral immune function. This finding again confirmed that an increase of *KIAA0101* expression reduces the survival rate of patients with LUAD. Therefore, our research will contribute to the improvement of immunotherapy of LUAD and the development of new immunotherapy targets.

In conclusion, we speculated that *KIAA0101* is a new LUAD marker. The prognosis of patients with LUAD with high expression of *KIAA0101* was poor. *KIAA0101* is related to the cell cycle, ribosome, protocol, splicing, and DNA replication. The results showing the immune cell associations of *KIAA0101* could lead to new methods of immunotherapy of LUAD. The study of drug transcription and metabolism associated with *KIAA0101* might provide alternative drugs for the targeted treatment of LUAD. This study will provide new directions for the diagnosis, prognostic evaluation, immunotherapy, and targeted therapy of LUAD.

We must acknowledge the potential limitations of our analysis. Further experimental verification is needed. In

addition, our study is limited to lung adenocarcinoma, and there was no analysis in other types of lung cancer. We hope that our next study will assess the association of *KIAA0101* with other types of lung cancer besides adenocarcinoma, to further compare the specific role of *KIAA0101* expression in the diagnosis and treatment of different types of lung cancer.

MATERIALS AND METHODS

Oncomine, GEPIA, and Ualcan analysis

All genes in the study of Hou Lung were analyzed and compared using Oncomine [50]. Overexpression of the genes in lung adenocarcinoma was compared with that in normal control tissue (log2 median-centered intensity). We used Oncomine to analyze the expression of *KIAA0101* in different tumor studies. In the Oncomine dataset, the naming convention lists the primary author's last name first, followed by the tissue type used in the study. We then used GEPIA [16] to analyze the expression of *KIAA0101* in a variety of tumors compared with that in normal tissue (Figure 1C). *KIAA0101* transcription in subgroups of patients with lung adenocarcinoma, stratified based on sex, age, and other criteria, were analyzed using the Ualcan database [14]. We also analyzed the differential expression of *KIAA0101* in LUAD tissues and normal lung tissues using Ualcan.

PrognoScan, OncoLnc, and TCGA portal analysis

The prognosis of patients with LUAD with different expression levels of *KIAA0101* was evaluated using PrognoScan [15]. PrognoScan is a new database for the meta-analysis of the prognostic value of genes. We used GEPIA [16], LinkedOmics [17], Ualcan [14], TISIDB [18], OncoLnc [19], and TCGA portals [20] to construct the OS curves for patients with LUAD. We used GEPIA [16] to construct the DFS curves for patients with LUAD. We used Kaplan Meier-plotter to construct the PFS curves for patients with LUAD. The survival curves between patients with low expression of *KIAA0101* and patients with high expression of *KIAA0101* in lung adenocarcinoma were compared.

LinkedOmics and TANRIC analysis

We conducted Gene Set Enrichment Analysis (GSEA) of KEGG pathway data [51], gene ontology cellular component, gene ontology biological process, gene ontology molecular function, a kinase target network, an miRNA target network, and a transcription factor target network of *KIAA0101*. We performed an miRNA Pearson correlation test and COXPH test of *KIAA0101* using Linkedomics [17]. We found the differentially expressed genes related to KIAA0101 in LUAD. We analyzed lncRNAs associated with *KIAA0101* using TANRIC [44] and assessed whether the level of lncRNA expression associated with *KIAA0101* had different impacts on the survival of patients.

TISIDB analysis and DrugBank analysis

We used TISIDB [18] to analyze the relationship between the expression of *KIAA0101* and immune infiltration in LUAD. Transcriptome and clinical data of 30 cancer types from the Cancer Genome Atlas (TCGA) were assessed [18]. High throughput screening data were then used to detect the resistance and sensitivity of tumor cells to T-cell-mediated killing. We used Drugbank to analyze the pharmacotranscription of *KIAA0101* [52]. The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information.

Abbreviations

KIAA0101/PCLAF*:* proliferating cell nuclear antigen binding factor; PCNA: proliferating cell nuclear antigen; LUAD: lung adenocarcinoma; GEO: Gene Expression Omnibus.

AUTHOR CONTRIBUTIONS

Sheng Hu and Yiping Wei designed and performed the research; analyzed and interpreted the data; and drafted the manuscript. Weibiao Zeng, Wenxiong Zhang, Jianjun Xu, Dongliang Yu and Jinhua Peng conceived the study and participated in the research design and data interpretation. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We hereby thank the Department of Thoracic Surgery of the Second Affiliated Hospital of Nanchang University for their strong support for this study.

CONFLICTS OF INTEREST

All the data in this paper are public, and the quoted data have indicated the source. There is no conflicts of interest in this paper.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (81560345 and 81860379), the Preeminence Youth Fund of Jiangxi Province (20162BCB23058), the China Postdoctoral

Science Foundation Grant (2017M610401), and the Science and Technology Planning Project at the Department of Science and Technology of Jiangxi Province, China (20151BBG70165 and 20171BAB 205075).

REFERENCES

1. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. CA Cancer J Clin. 2016; 66:115–32. https://doi.org/10.3322/caac.21338 PMID:26808342

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019; 69:7–34. https://doi.org/10.3322/caac.21551 PMID:30620402

3. Govindan R, Page N, Morgensztern D, Read W, Tierney R, Vlahiotis A, Spitznagel EL, Piccirillo J. Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. J Clin Oncol. 2006; 24:4539–44. https://doi.org/10.1200/JCO.2005.04.4859 PMID:17008692

4. Ruiz-Cordero R, Devine WP. Targeted therapy and checkpoint immunotherapy in lung cancer. Surg Pathol Clin. 2020; 13:17–33. https://doi.org/10.1016/j.path.2019.11.002 PMID:32005431

5. Pabla S, Conroy JM, Nesline MK, Glenn ST, Papanicolau-Sengos A, Burgher B, Hagen J, Giamo V, Andreas J, Lenzo FL, Yirong W, Dy GK, Yau E, et al. Proliferative potential and resistance to immune checkpoint blockade in lung cancer patients. J Immunother Cancer. 2019; 7:27. https://doi.org/10.1186/s40425-019-0506-3 PMID:30709424

6. Sun CC, Li SJ, Yuan ZP, Li DJ. MicroRNA-346 facilitates cell growth and metastasis, and suppresses cell apoptosis in human non-small cell lung cancer by regulation of XPC/ERK/Snail/E-cadherin pathway. Aging (Albany NY). 2016; 8:2509–24. https://doi.org/10.18632/aging.101080 PMID:27777383

7. Moldovan GL, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell. 2007; 129:665–79. https://doi.org/10.1016/j.cell.2007.05.003 PMID:17512402

8. Halling-Brown MD, Bulusu KC, Patel M, Tym JE, Al-Lazikani B. canSAR: an integrated cancer public translational research and drug discovery resource. Nucleic Acids Res. 2012; 40:D947–56. https://doi.org/10.1093/nar/gkr881 PMID:22013161

9. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, et al. A subcellular map of the human proteome. Science. 2017; 356:eaal3321. https://doi.org/10.1126/science.aal3321 PMID:28495876

10. Zhang T, Guo J, Gu J, Chen K, Wang Z, Li H, Wang G, Wang J. KIAA0101 is a novel transcriptional target of FoxM1 and is involved in the regulation of hepatocellular carcinoma microvascular invasion by regulating epithelial-mesenchymal transition. J Cancer. 2019; 10:3501–16. https://doi.org/10.7150/jca.29490 PMID:31293655

11. Lv W, Su B, Li Y, Geng C, Chen N. KIAA0101 inhibition suppresses cell proliferation and cell cycle progression by promoting the interaction between p53 and Sp1 in breast cancer. Biochem Biophys Res Commun. 2018; 503:600–06. https://doi.org/10.1016/j.bbrc.2018.06.046 PMID:29902451

12. Kais Z, Barsky SH, Mathsyaraja H, Zha A, Ransburgh DJ, He G, Pilarski RT, Shapiro CL, Huang K, Parvin JD. KIAA0101 interacts with BRCA1 and regulates centrosome number. Mol Cancer Res. 2011; 9:1091–99. https://doi.org/10.1158/1541-7786.MCR-10-0503 PMID:21673012

13. Zhu K, Diao D, Dang C, Shi L, Wang J, Yan R, Yuan D, Li K. Elevated KIAA0101 expression is a marker of recurrence in human gastric cancer. Cancer Sci. 2013; 104:353–59. https://doi.org/10.1111/cas.12083 PMID:23240630

14. Chandrashekar DS, Bashel B, Balasubramanya SA, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BV, Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia. 2017; 19:649–58. https://doi.org/10.1016/j.neo.2017.05.002 PMID:28732212

15. Mizuno H, Kitada K, Nakai K, Sarai A. PrognoScan: a new database for meta-analysis of the prognostic value of genes. BMC Med Genomics. 2009; 2:18. https://doi.org/10.1186/1755-8794-2-18 PMID:19393097

16. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017; 45:W98–102. https://doi.org/10.1093/nar/gkx247 PMID:28407145

17. Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. 2018; 46:D956–63. https://doi.org/10.1093/nar/gkx1090 PMID:29136207

18. Ru B, Wong CN, Tong Y, Zhong JY, Zhong SS, Wu WC, Chu KC, Wong CY, Lau CY, Chen I, Chan NW, Zhang J. TISIDB: an integrated repository portal for tumor-immune system interactions. Bioinformatics. 2019; 35:4200–02. https://doi.org/10.1093/bioinformatics/btz210 PMID:30903160

19. Anaya J. Guide to using OncoLnc, a new TCGA data portal. Figshare. 2016. https://doi.org/10.6084/m9.figshare.2991640.v1

20. Xu S, Feng Y, Zhao S. Proteins with evolutionarily hypervariable domains are associated with immune response and better survival of basal-like breast cancer patients. Comput Struct Biotechnol J. 2019; 17:430–40. https://doi.org/10.1016/j.csbj.2019.03.008 PMID:30996822

21. KEGG. Curr Opin Plant Biol. 2000; 3:177. https://doi.org/10.1016/S1369-5266(00)80051-5

22. Jiang J, Briedé JJ, Jennen DG, Van Summeren A, Saritas-Brauers K, Schaart G, Kleinjans JC, de Kok TM. Increased mitochondrial ROS formation by acetaminophen in human hepatic cells is associated with gene expression changes suggesting disruption of the mitochondrial electron transport chain. Toxicol Lett. 2015; 234:139–50. https://doi.org/10.1016/j.toxlet.2015.02.012 PMID:25704631

23. Buterin T, Koch C, Naegeli H. Convergent transcriptional profiles induced by endogenous estrogen and distinct xenoestrogens in breast cancer cells. Carcinogenesis. 2006; 27:1567–78. https://doi.org/10.1093/carcin/bgi339 PMID:16474171

24. Wilcox CB, Feddes GO, Willett-Brozick JE, Hsu LC, DeLoia JA, Baysal BE. Coordinate up-regulation of TMEM97 and cholesterol biosynthesis genes in normal ovarian surface epithelial cells treated with progesterone: implications for pathogenesis of ovarian cancer. BMC Cancer. 2007; 7:223. https://doi.org/10.1186/1471-2407-7-223 PMID:18070364

25. Dip R, Lenz S, Gmuender H, Naegeli H. Pleiotropic combinatorial transcriptomes of human breast cancer cells exposed to mixtures of dietary phytoestrogens. Food Chem Toxicol. 2009; 47:787–95. https://doi.org/10.1016/j.fct.2009.01.008 PMID:19167446

26. Takahashi K, Tatsumi N, Fukami T, Yokoi T, Nakajima M. Integrated analysis of rifampicin-induced microRNA and gene expression changes in human hepatocytes. Drug Metab Pharmacokinet. 2014; 29:333–40. https://doi.org/10.2133/dmpk.dmpk-13-rg-114 PMID:24552687

27. Krug AK, Kolde R, Gaspar JA, Rempel E, Balmer NV, Meganathan K, Vojnits K, Baquié M, Waldmann T, Ensenat-Waser R, Jagtap S, Evans RM, Julien S, et al. Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. Arch Toxicol. 2013; 87:123–43. https://doi.org/10.1007/s00204-012-0967-3 PMID:23179753

28. Gertz J, Reddy TE, Varley KE, Garabedian MJ, Myers RM. Genistein and bisphenol a exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. Genome Res. 2012; 22:2153–62. https://doi.org/10.1101/gr.135681.111 PMID:23019147

29. Boehme K, Simon S, Mueller SO. Gene expression profiling in ishikawa cells: a fingerprint for estrogen active compounds. Toxicol Appl Pharmacol. 2009; 236:85–96. https://doi.org/10.1016/j.taap.2009.01.006 PMID:19371625

30. Soloff MS, Jeng YJ, Izban MG, Sinha M, Luxon BA, Stamnes SJ, England SK. Effects of progesterone treatment on expression of genes involved in uterine quiescence. Reprod Sci. 2011; 18:781–97. https://doi.org/10.1177/1933719111398150 PMID:21795739

31. Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, Pio R, Zulueta JJ, Spira A, Massion PP, Mazzone PJ, Montuenga LM. Biomarkers in lung cancer screening: achievements, promises, and challenges. J Thorac Oncol. 2019; 14:343–57. https://doi.org/10.1016/j.jtho.2018.11.023 PMID:30529598

32. Chen H, Xia B, Liu T, Lin M, Lou G. KIAA0101, a target gene of miR-429, enhances migration and chemoresistance of epithelial ovarian cancer cells. Cancer Cell Int. 2016; 16:74. https://doi.org/10.1186/s12935-016-0353-y PMID:27708548

33. Jain M, Zhang L, Patterson EE, Kebebew E. KIAA0101 is overexpressed, and promotes growth and invasion in adrenal cancer. PLoS One. 2011; 6:e26866. https://doi.org/10.1371/journal.pone.0026866 PMID:22096502

34. Kato T, Daigo Y, Aragaki M, Ishikawa K, Sato M, Kaji M. Overexpression of KIAA0101 predicts poor prognosis in primary lung cancer patients. Lung Cancer. 2012; 75:110–18.

https://doi.org/10.1016/j.lungcan.2011.05.024
PMID:21689861

35. Emanuele MJ, Ciccia A, Elia AE, Elledge SJ. Proliferating cell nuclear antigen (PCNA)-associated KIAA0101/PAF15 protein is a cell cycle-regulated anaphase-promoting complex/cyclosome substrate. Proc Natl Acad Sci USA. 2011; 108:9845–50. https://doi.org/10.1073/pnas.1106136108 PMID:21628590

36. Simpson F, Lammerts van Bueren K, Butterfield N, Bennetts JS, Bowles J, Adolphe C, Simms LA, Young J, Walsh MD, Leggett B, Fowles LF, Wicking C. The PCNA-associated factor KIAA0101/p15(PAF) binds the potential tumor suppressor product p33ING1b. Exp Cell Res. 2006; 312:73–85. https://doi.org/10.1016/j.yexcr.2005.09.020 PMID:16288740

37. van Bueren KL, Bennetts JS, Fowles LF, Berkman JL, Simpson F, Wicking C. Murine embryonic expression of the gene for the UV-responsive protein p15(PAF). Gene Expr Patterns. 2007; 7:47–50. https://doi.org/10.1016/j.modgep.2006.05.006 PMID:16815099

38. Fang H, Niu K, Mo D, Zhu Y, Tan Q, Wei D, Li Y, Chen Z, Yang S, Balajee AS, Zhao Y. RecQL4-aurora B kinase axis is essential for cellular proliferation, cell cycle progression, and mitotic integrity. Oncogenesis. 2018; 7:68. https://doi.org/10.1038/s41389-018-0080-4 PMID:30206236

39. Herrera MC, Chymkowitch P, Robertson JM, Eriksson J, Bøe SO, Alseth I, Enserink JM. Corrigendum: Cdk1 gates cell cycle-dependent tRNA synthesis by regulating RNA polymerase III activity. Nucleic Acids Res. 2018; 46:12188–89. https://doi.org/10.1093/nar/gky1102 PMID:30395271

40. Cortez D, Guntuku S, Qin J, Elledge SJ. ATR and ATRIP: partners in checkpoint signaling. Science. 2001; 294:1713–16. https://doi.org/10.1126/science.1065521 PMID:11721054

41. Hafsi H, Dillon MT, Barker HE, Kyula JN, Schick U, Paget JT, Smith HG, Pedersen M, McLaughlin M, Harrington KJ. Combined ATR and DNA-PK inhibition radiosensitizes tumor cells independently of their p53 status. Front Oncol. 2018; 8:245. https://doi.org/10.3389/fonc.2018.00245 PMID:30057890

42. Lin Y, Liang R, Qiu Y, Lv Y, Zhang J, Qin G, Yuan C, Liu Z, Li Y, Zou D, Mao Y. Expression and gene regulation network of RBM8A in hepatocellular carcinoma based on data mining. Aging (Albany NY). 2019; 11:423–47. https://doi.org/10.18632/aging.101749 PMID:30670676

43. Scafoglio C, Ambrosino C, Cicatiello L, Altucci L, Ardovino M, Bontempo P, Medici N, Molinari AM, Nebbioso A, Facchiano A, Calogero RA, Elkon R, Menini N, et al. Comparative gene expression profiling reveals partially overlapping but distinct genomic actions of different antiestrogens in human breast cancer cells. J Cell Biochem. 2006; 98:1163–84. https://doi.org/10.1002/jcb.20820 PMID:16514628

44. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, Weinstein JN, Liang H. TANRIC: an interactive open platform to explore the function of lncRNAs in cancer. Cancer Res. 2015; 75:3728–37. https://doi.org/10.1158/0008-5472.CAN-15-0273 PMID:26208906

45. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. Nat Rev Cancer. 2009; 9:239–52. https://doi.org/10.1038/nrc2618 PMID:19279573

46. Sasada T, Suekane S. Variation of tumor-infiltrating lymphocytes in human cancers: controversy on clinical significance. Immunotherapy. 2011; 3:1235–51. https://doi.org/10.2217/imt.11.106 PMID:21995574

47. Lauerova L, Dusek L, Simickova M, Kocák I, Vagundová M, Zaloudík J, Kovarík J. Malignant melanoma associates with Th1/Th2 imbalance that coincides with disease progression and immunotherapy response. Neoplasma. 2002; 49:159–66. PMID:12098001

48. Liu J, Li Y, Zhu X, Li Q, Liang X, Xie J, Hu S, Peng W, Li C. Increased CX3CL1 mRNA expression level is a positive prognostic factor in patients with lung adenocarcinoma. Oncol Lett. 2019; 17:4877–90. https://doi.org/10.3892/ol.2019.10211 PMID:31186696

49. Holling TM, Schooten E, van Den Elsen PJ. Function and regulation of MHC class II molecules in T-lymphocytes: of mice and men. Hum Immunol. 2004; 65:282–90. https://doi.org/10.1016/j.humimm.2004.01.005 PMID:15120183

50. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D, Chinnaiyan AM. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia. 2007; 9:166–80. https://doi.org/10.1593/neo.07112 PMID:17356713

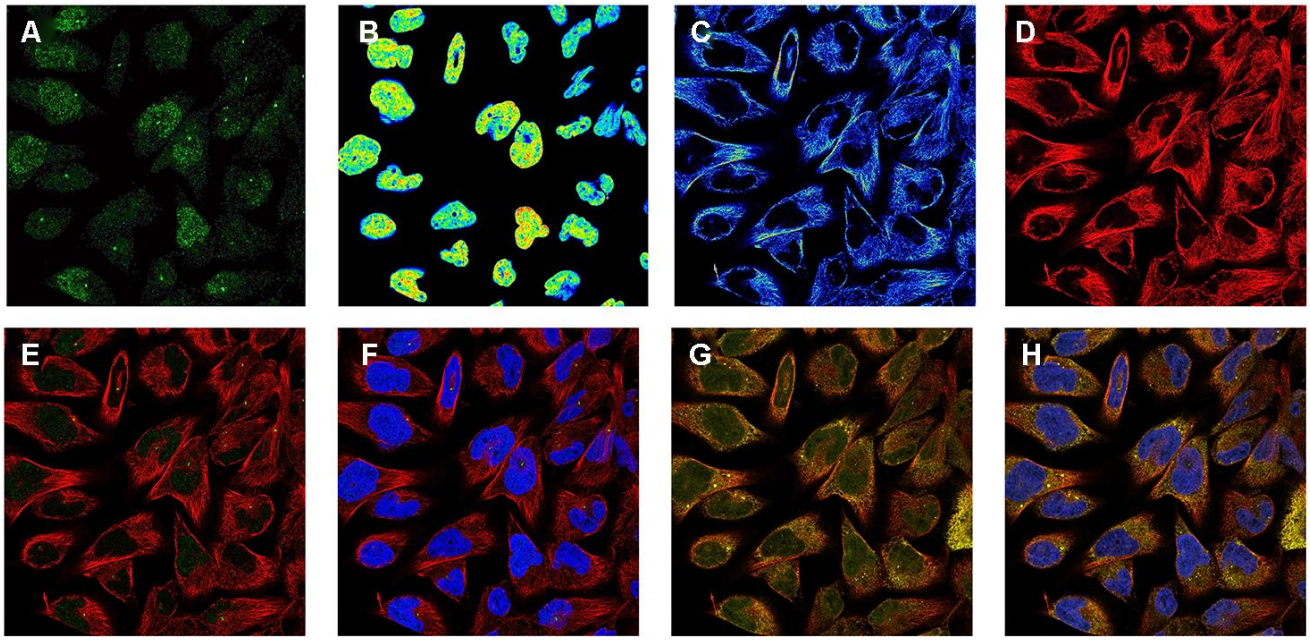
51. Kanehisa M, Sato Y. KEGG mapper for inferring cellular functions from protein sequences. Protein Sci. 2020; 29:28–35.

https://doi.org/10.1002/pro.3711
PMID:31423653

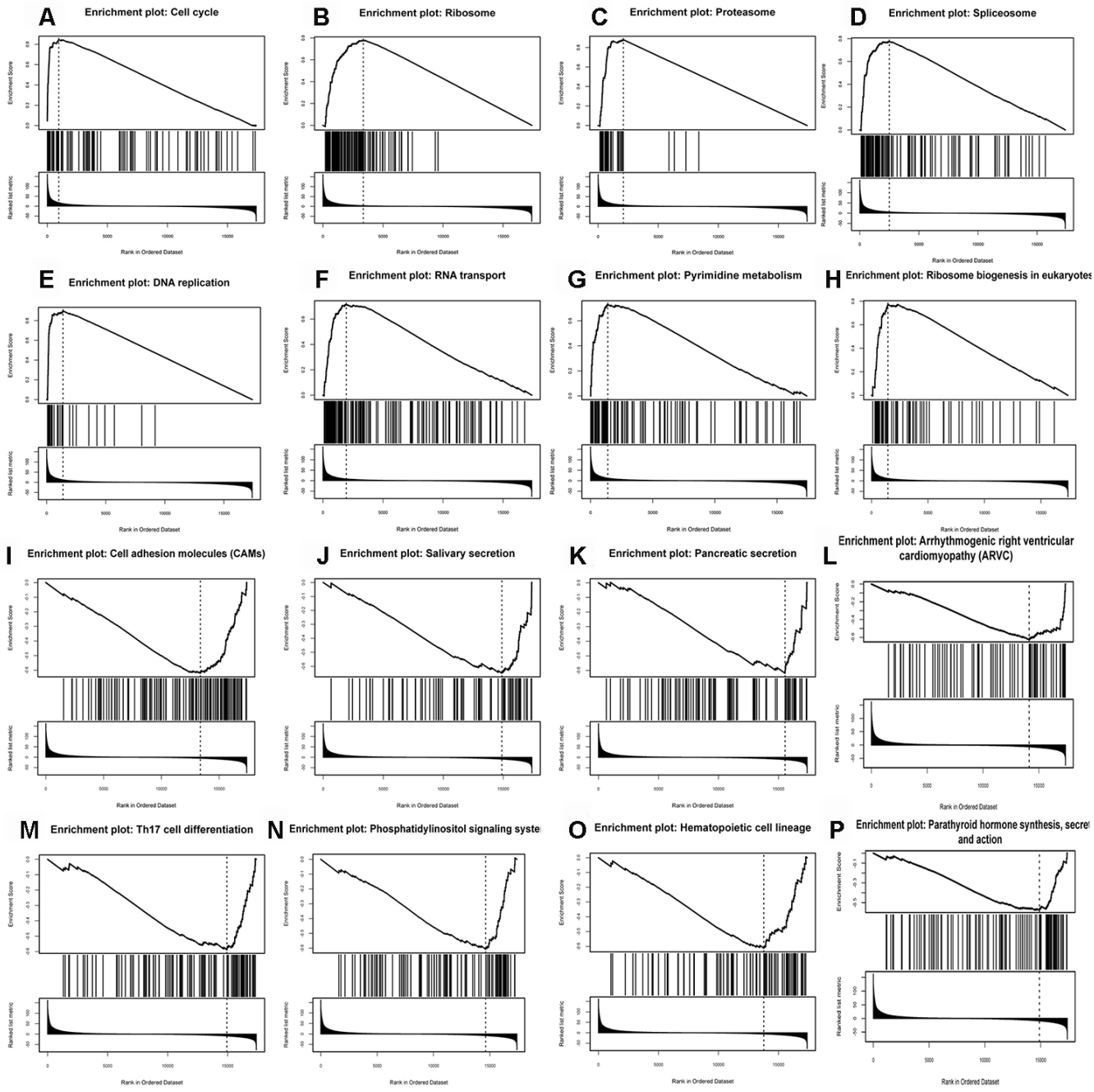
52. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018; 46:D1074–82.
https://doi.org/10.1093/nar/gkx1037
PMID:29126136

SUPPLEMENTARY MATERIALS

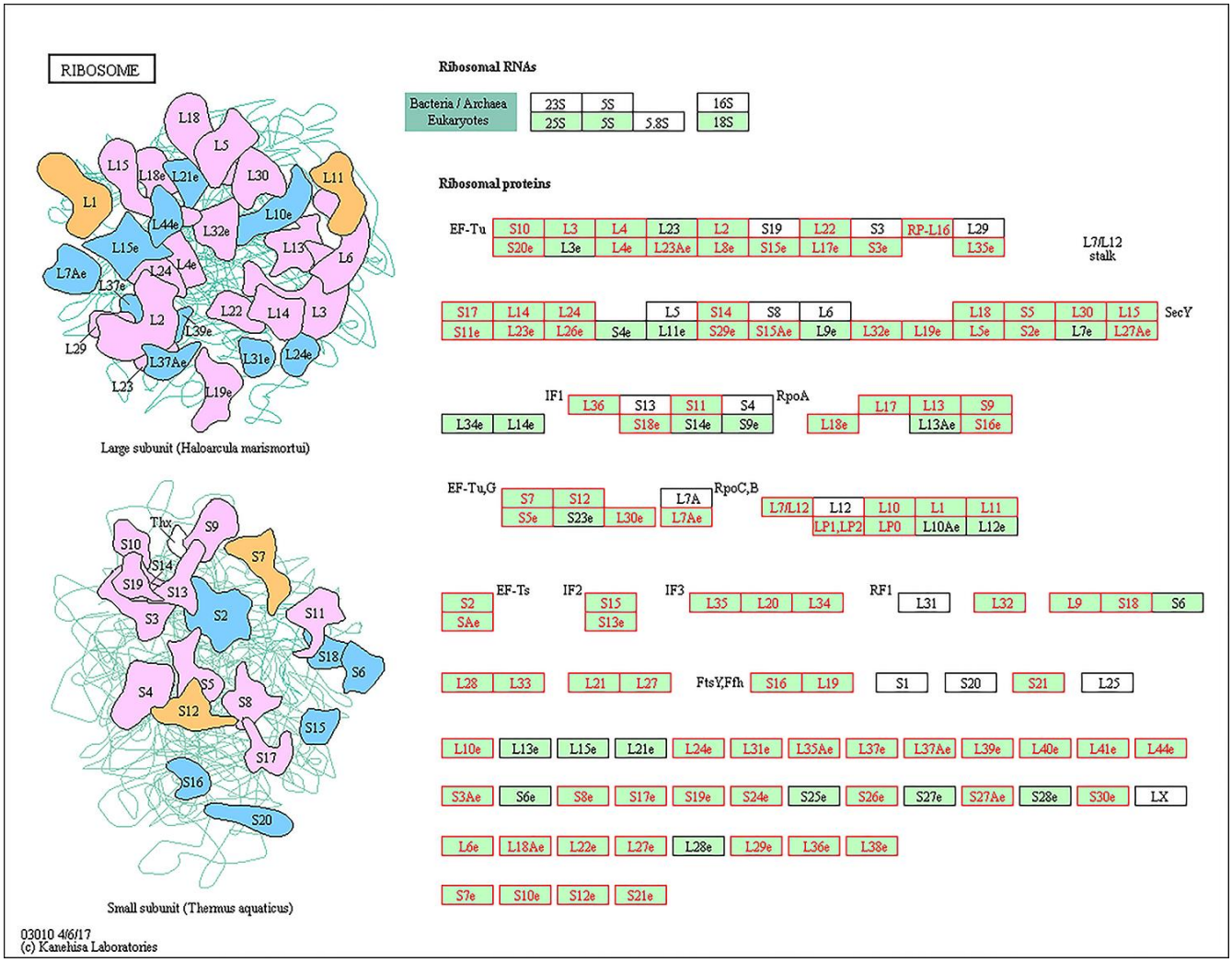
Supplementary Figures



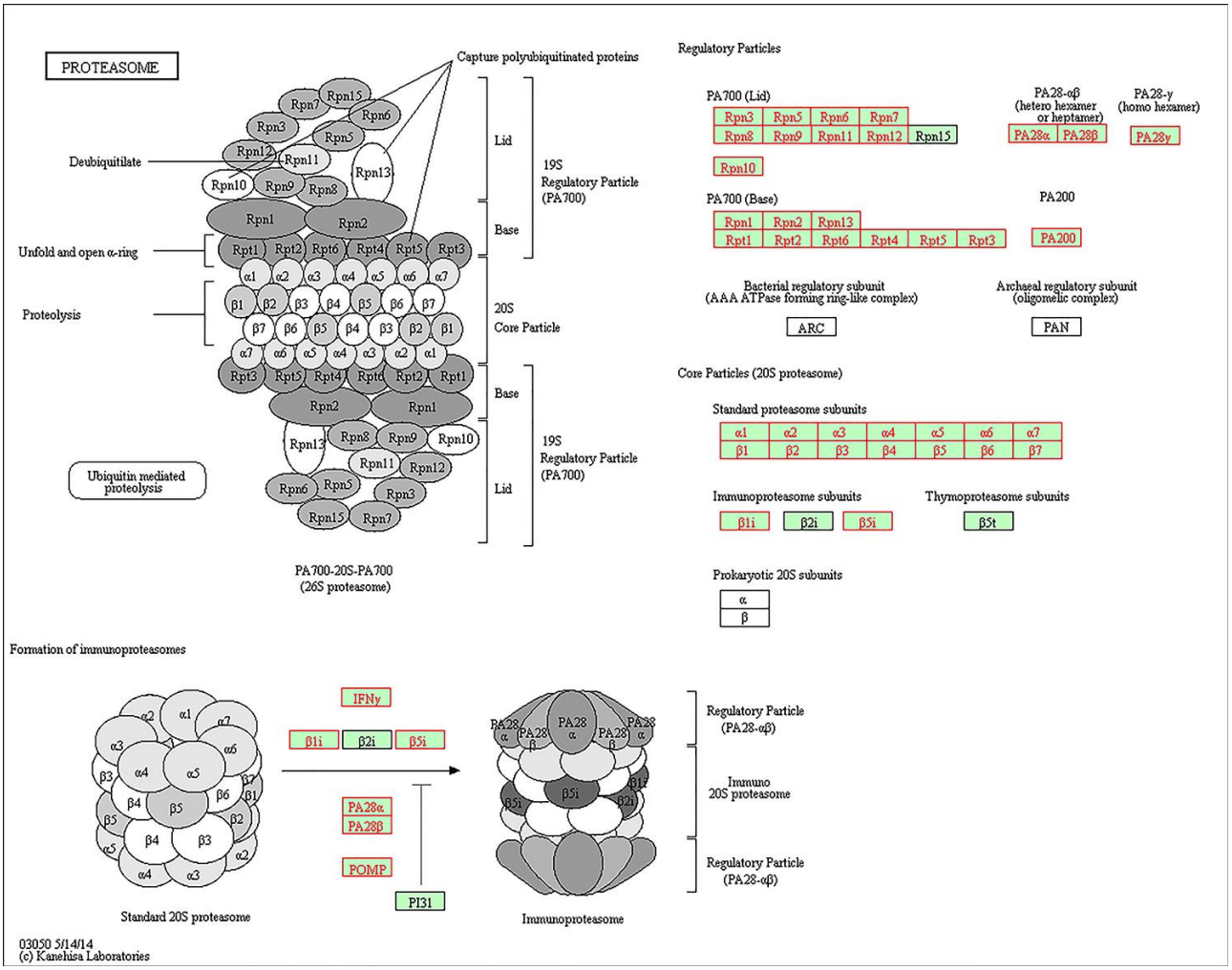
Supplementary Figure 1. Subcellular mapping of the KIAA0101 human proteome in lung adenocarcinoma cells. (**A**–**H**) Represent the subcellular mapping of the KIAA0101 human proteome (RH-30, HPA047929), respectively. (**A**) Antibody. (**B**) Nucleus and intensity. (**C**) Microtubule and intensity. (**D**) Microtubule. (**E**) Antibody and microtubule. (**F**) Antibody, nucleus, and microtubule. (**G**) Microtubule and endoplasmic reticulum. (**H**) Microtubule and endoplasmic reticulum. The cell structure sections were based on representative images of antibody-stained human cancer cell lines. Each sample was stained with an in-house generated HPA-antibody and counterstained with markers for microtubules, the endoplasmic reticulum, and the nucleus.



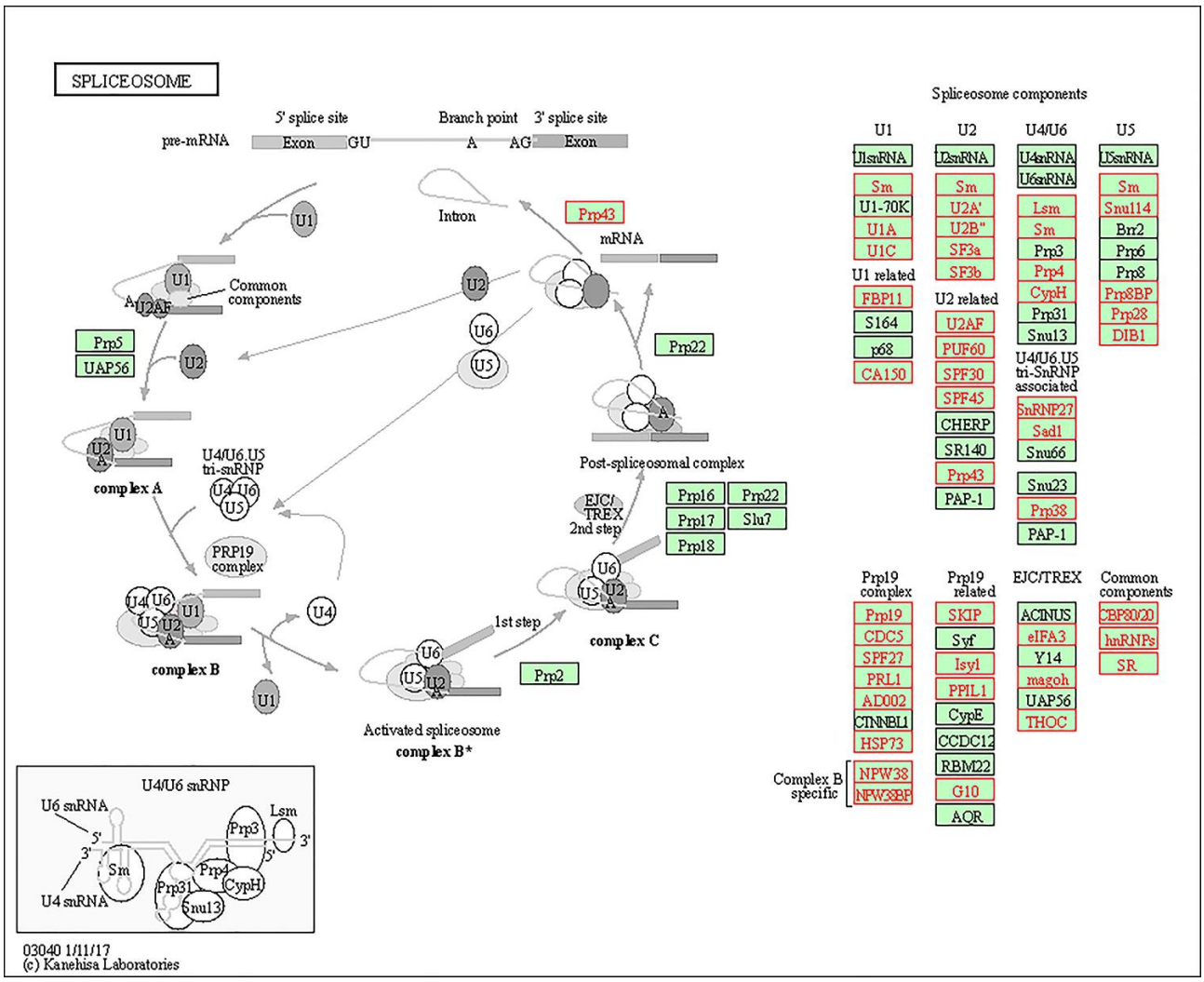
Supplementary Figure 2. Enrichment plot of pathways in the enrichment results of figure 3A with the top 8 and the bottom 8 genes according to the normalized enrichment score. (**A**) hsa04110, Cell cycle; (**B**) hsa03010, Ribosome; (**C**) hsa03050, Proteasome; (**D**) hsa03040, Spliceosome; (**E**) hsa03030, DNA replication; (**F**) hsa03013, RNA transport; (**G**) hsa00240, Pyrimidine metabolism; (**H**) hsa03008, Ribosome biogenesis in eukaryotes; (**I**) hsa04514, Cell adhesion molecules (CAMs); (**J**) hsa04970, Salivary secretion; (**K**) hsa04972, Pancreatic secretion; (**L**) hsa05412, Arrhythmogenic right ventricular cardiomyopathy (ARVC); (**M**) hsa04659, Th17 cell differentiation; (**N**) hsa04070, Phosphatidylinositol signaling system; (**O**) hsa04640, Hematopoietic cell lineage; (**P**) hsa04928, Parathyroid hormone synthesis secretion and action.



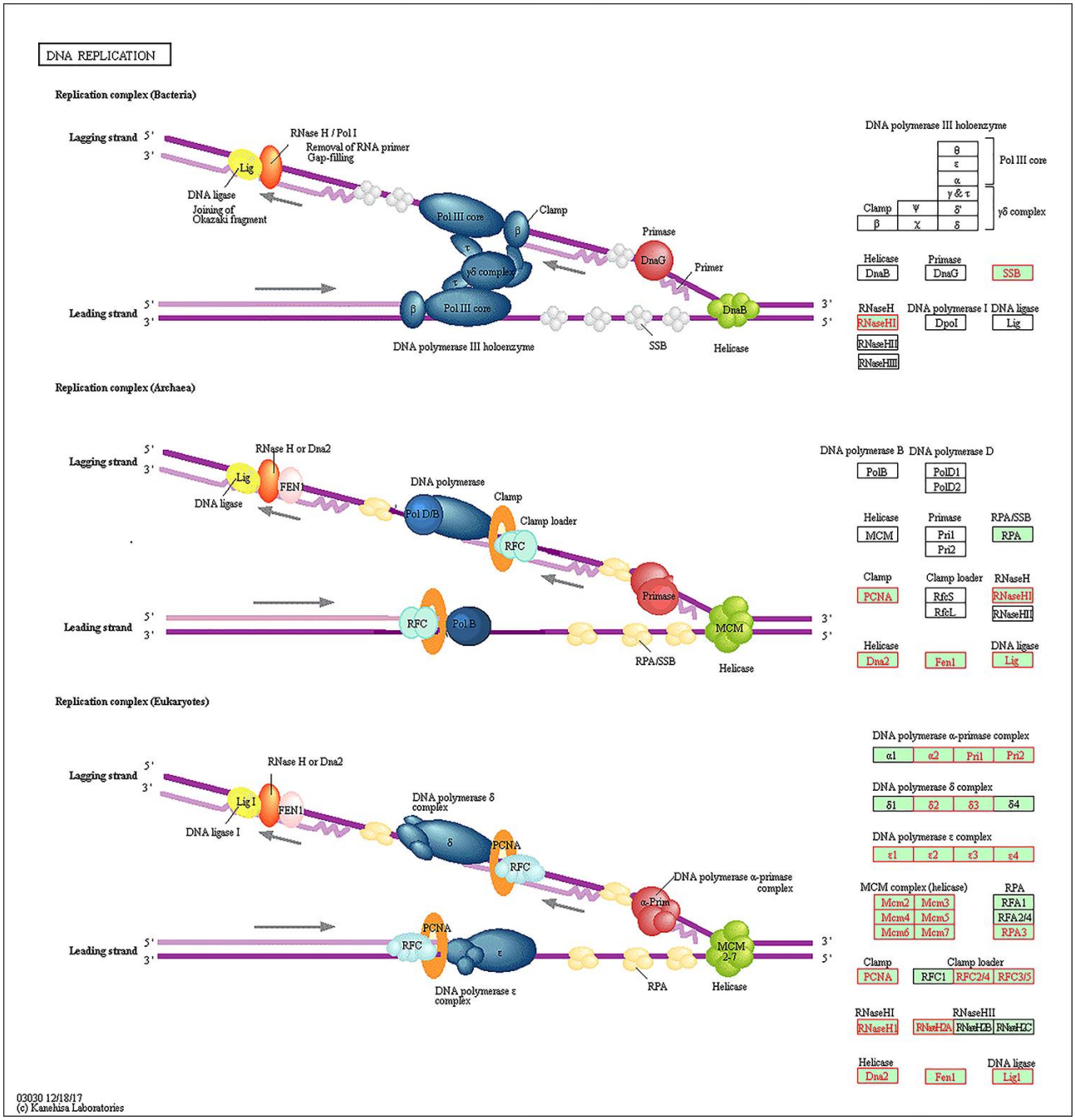
Supplementary Figure 3. KEGG pathway annotations of the ribosome pathway (HSA03010). Red denotes leading edge genes; green denotes the remaining genes.



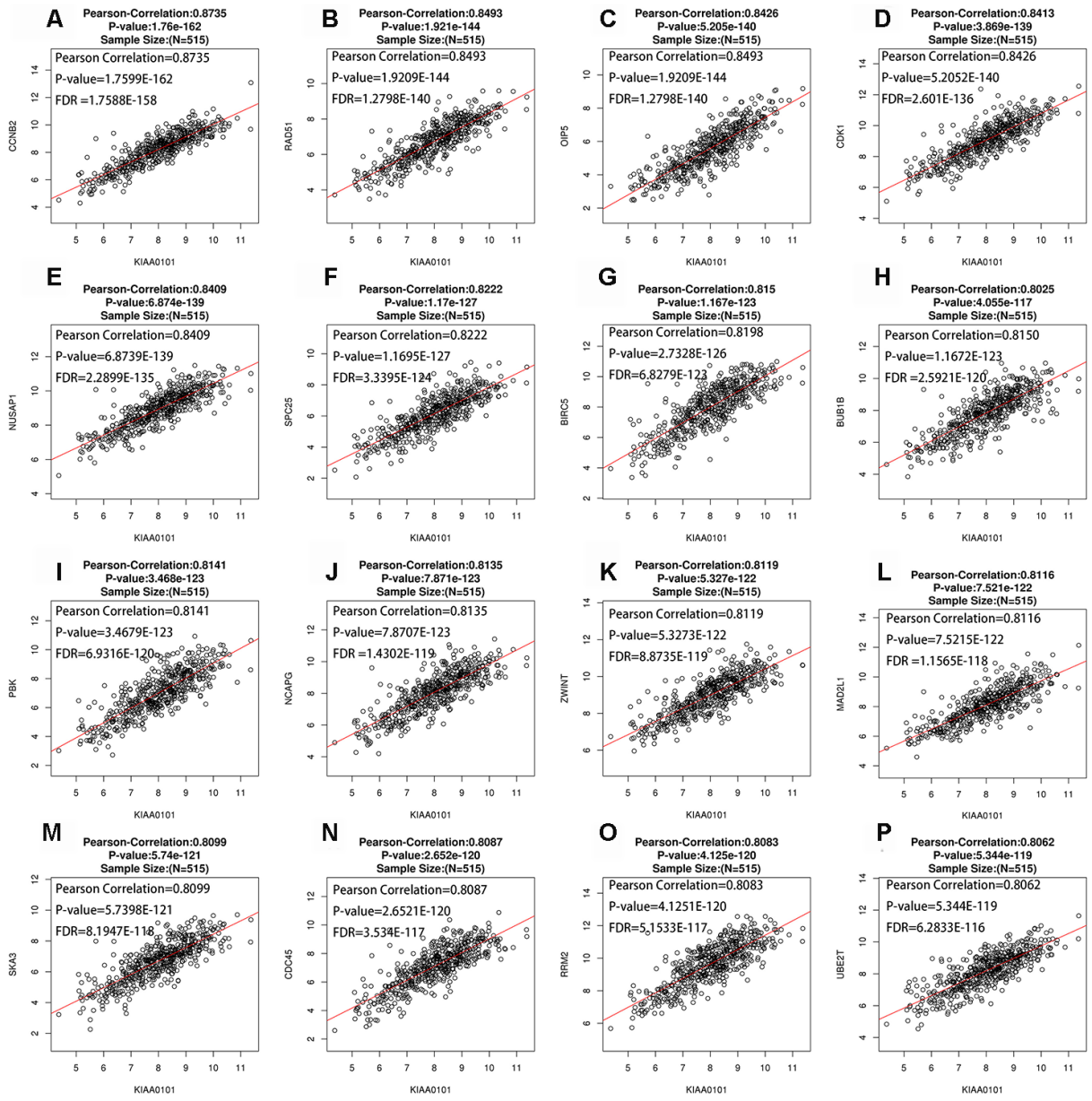
Supplementary Figure 4. KEGG pathway annotations of the Proteasome pathway (hsa03050). Red denotes leading edge genes; green denotes the remaining genes.



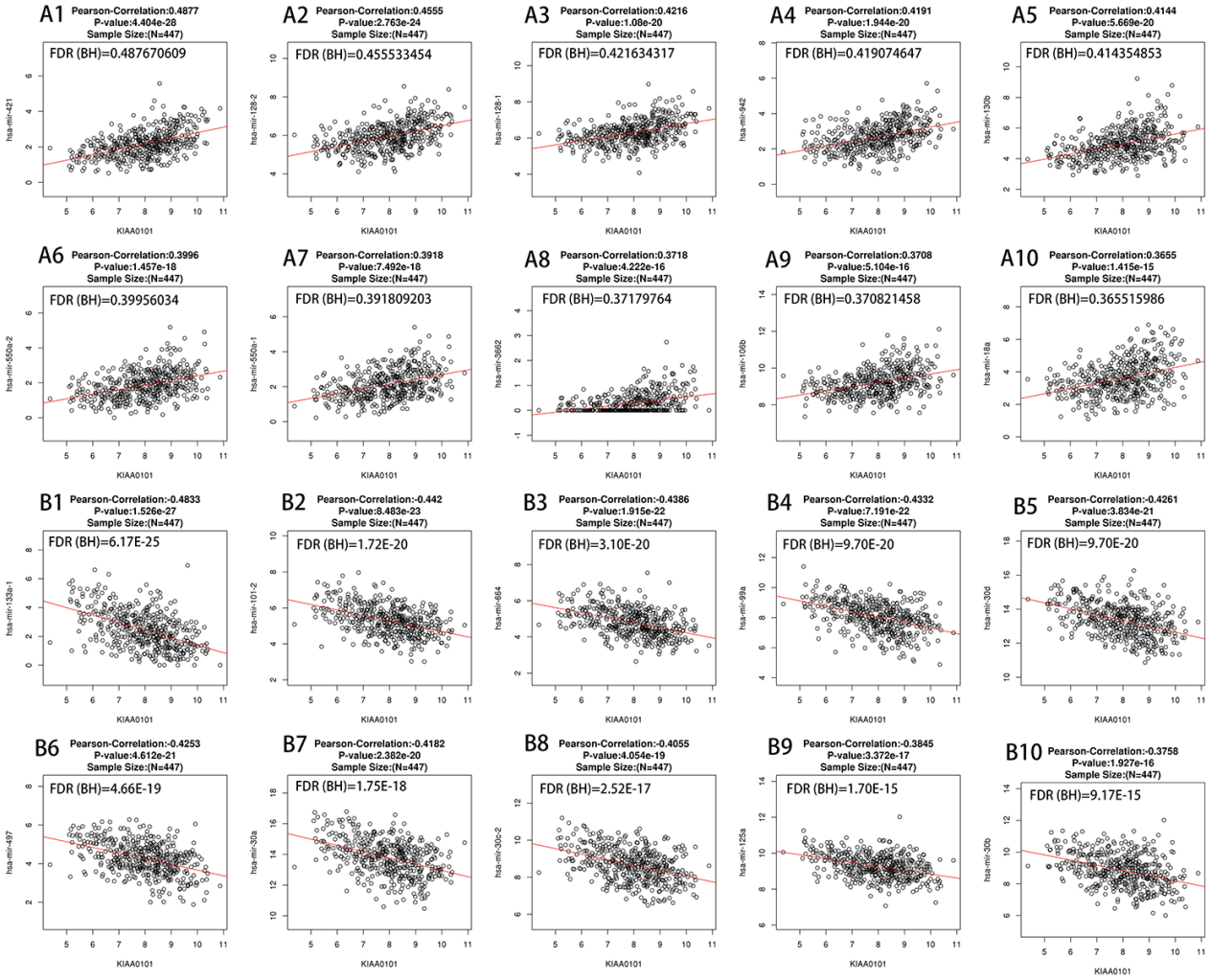
Supplementary Figure 5. KEGG pathway annotations of the Spliceosome pathway (hsa03040). Red denotes leading edge genes; green denotes the remaining genes.



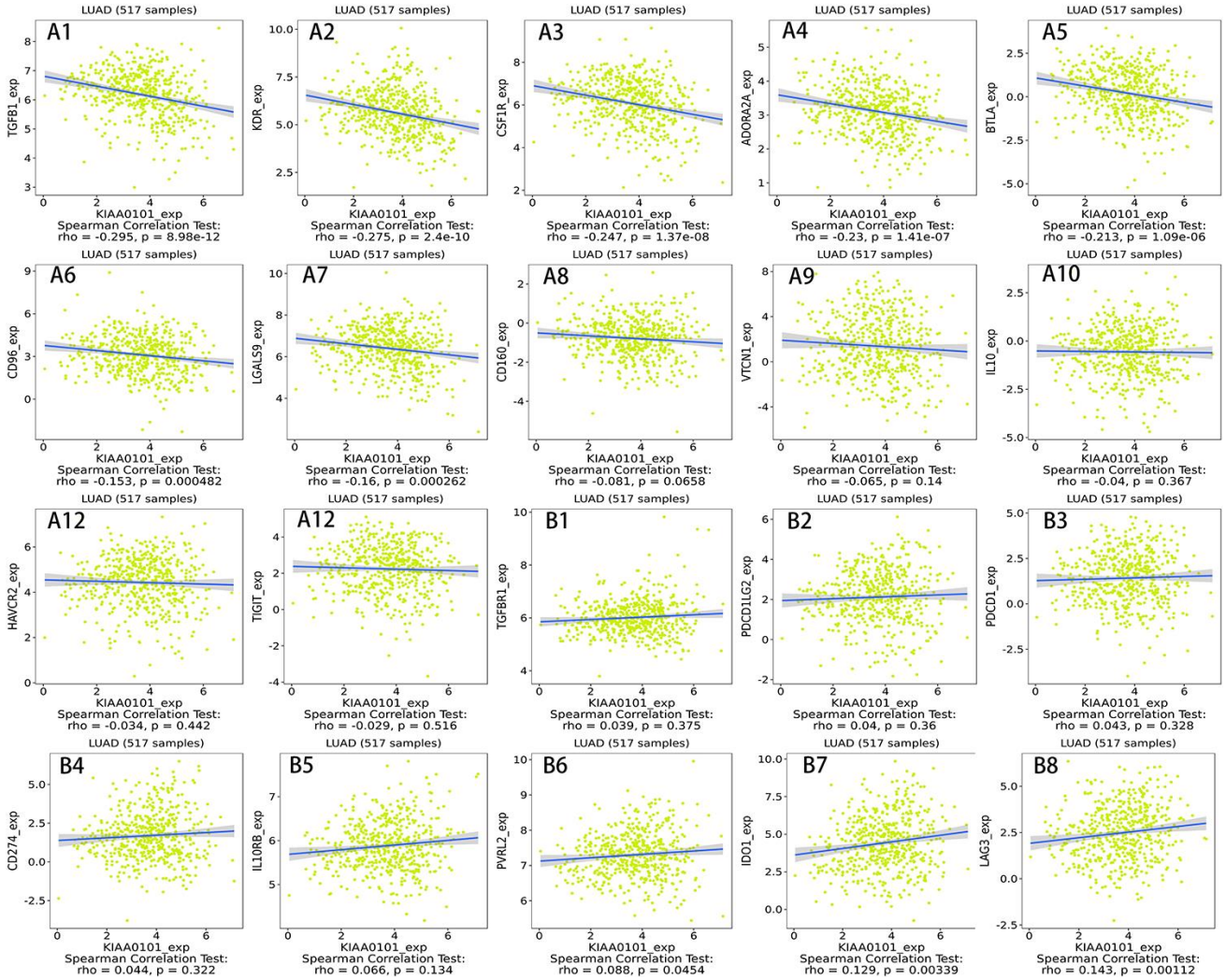
Supplementary Figure 6. KEGG pathway annotations of the DNA replication pathway (hsa03030). Red denotes leading edge genes; green denotes the remaining genes.



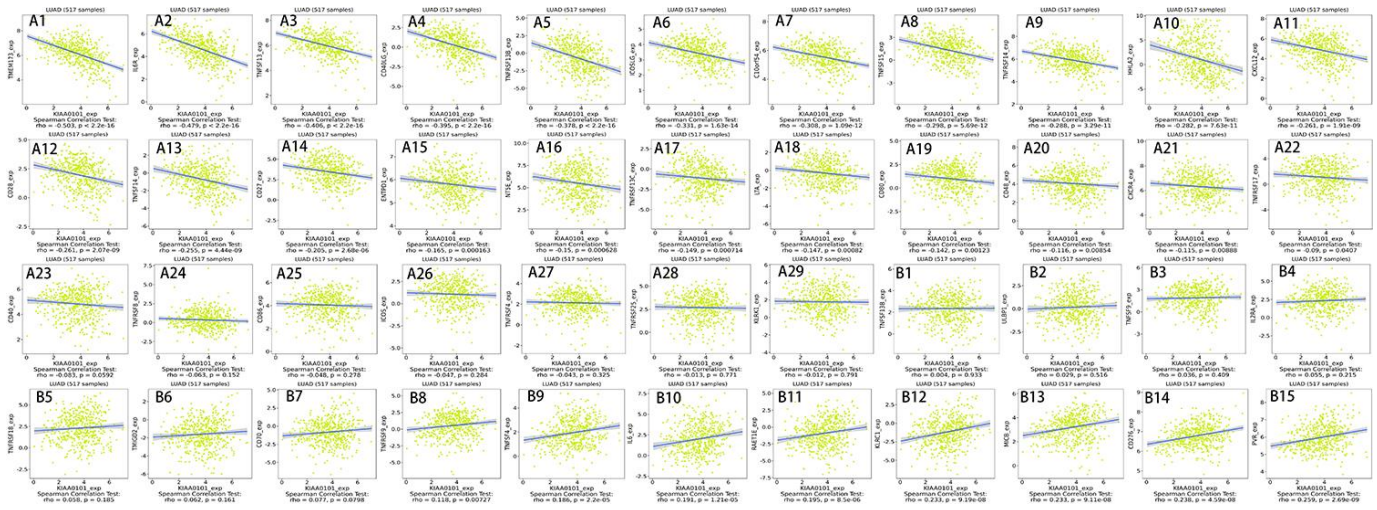
Supplementary Figure 7. The scatter diagram of *KIAA0101* and the first 16 genes highly related to *KIAA0101* (Pearson correlation coefficient > 0.8). (**A**) *CCNB2*; (**B**) *RAD51*; (**C**) *OIP5*; (**D**) *CDK1*; (**E**) *NUSAP1*; (**F**) *SPC25*; (**G**) *CCNB1*; (**H**) *BIRC5*; (**I**) *PBK*; (**J**) *NCAPG*; (**K**) *ZWINT*; (**L**) *MAD2L1*; (**M**) *SKA3*; (**N**) *CDC45*; (**O**) *RRM2*; (**P**) *UBE2T*.



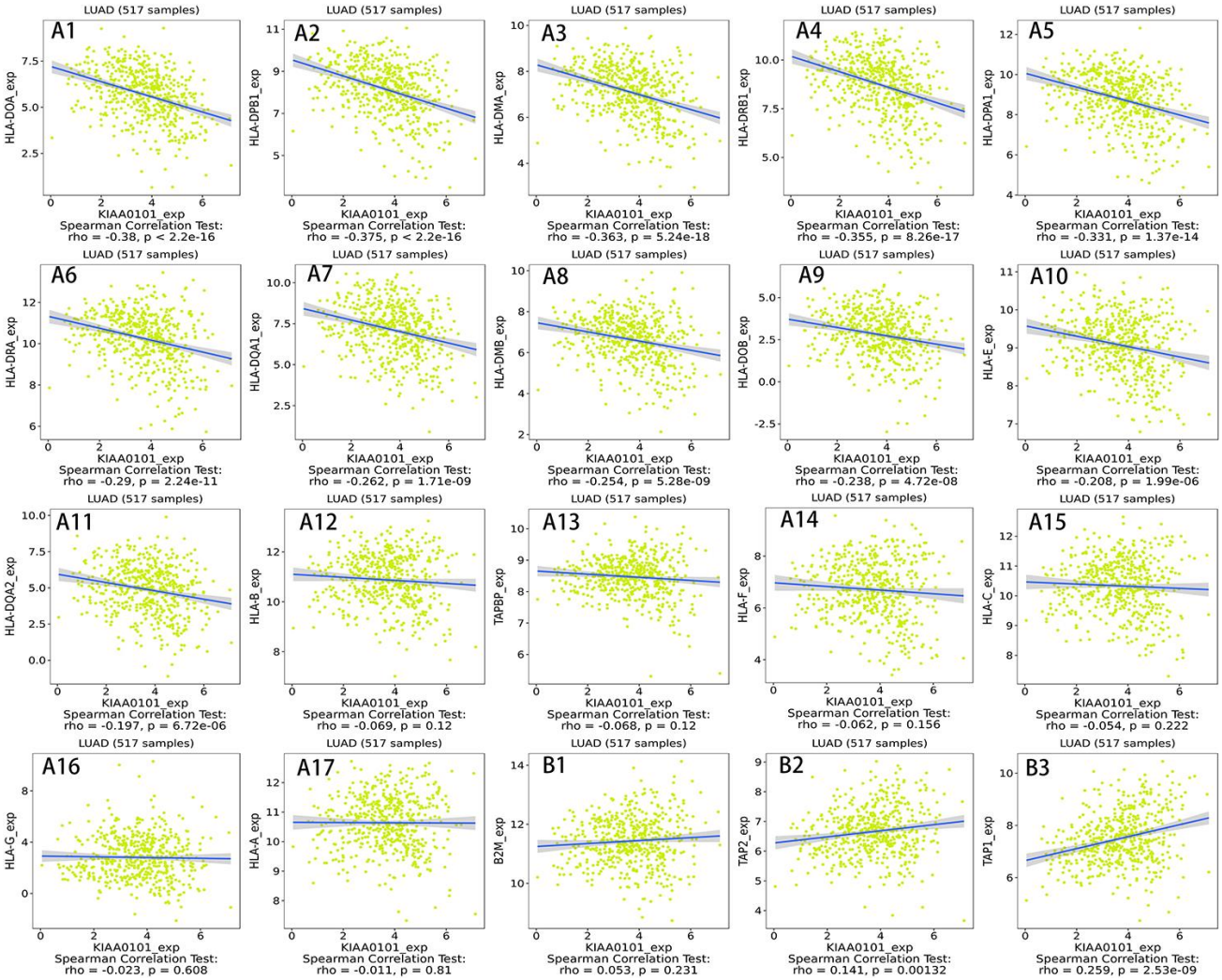
Supplementary Figure 8. Expression of *KIAA0101* and correlation scatter diagram of miRNAs. (**A1**–**A10**) Positive correlations between the expression of *KIAA0101* and miRNAs. (**B1**–**B10**) Negative correlations between the expression of *KIAA0101* and miRNAs.



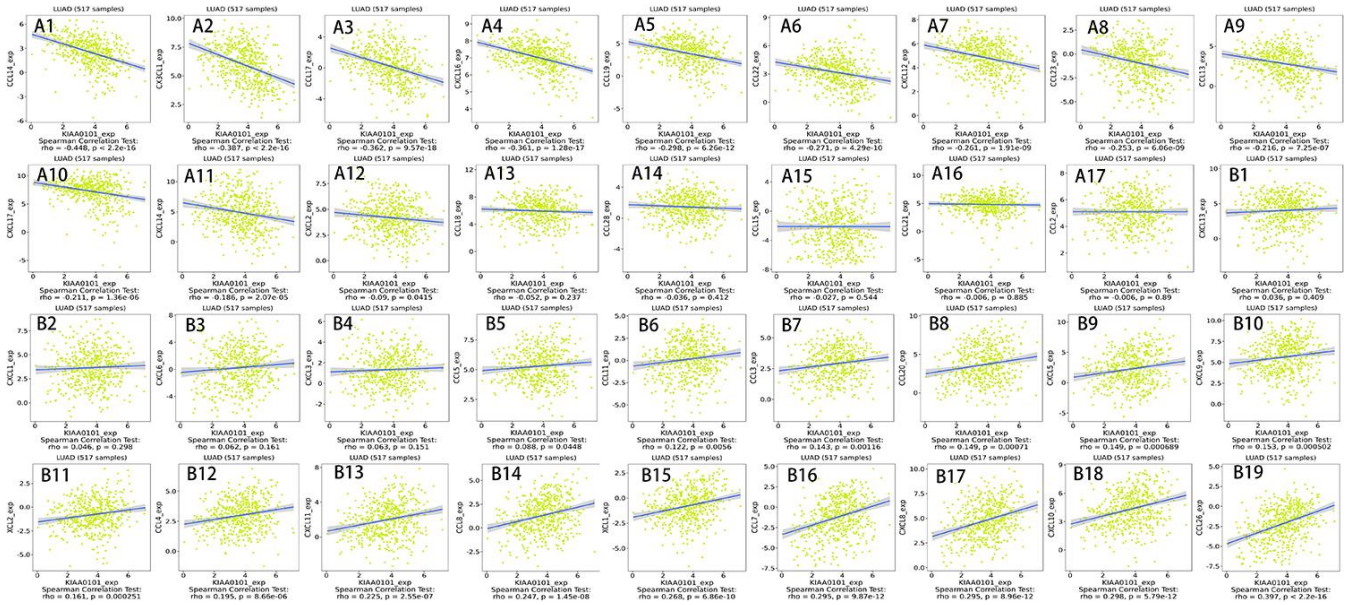
Supplementary Figure 9. Correlations between the expression of *KIAA0101* and immunoinhibitors. (**A1**–**A12**) are scatter plots of the negative correlations between *KIAA0101* expression and immunoinhibitors in the treatment of lung adenocarcinoma. (**B1**–**B8**) are scatter plots of the positive correlations between *KIAA0101* expression and immunoinhibitors in the treatment of lung adenocarcinoma.



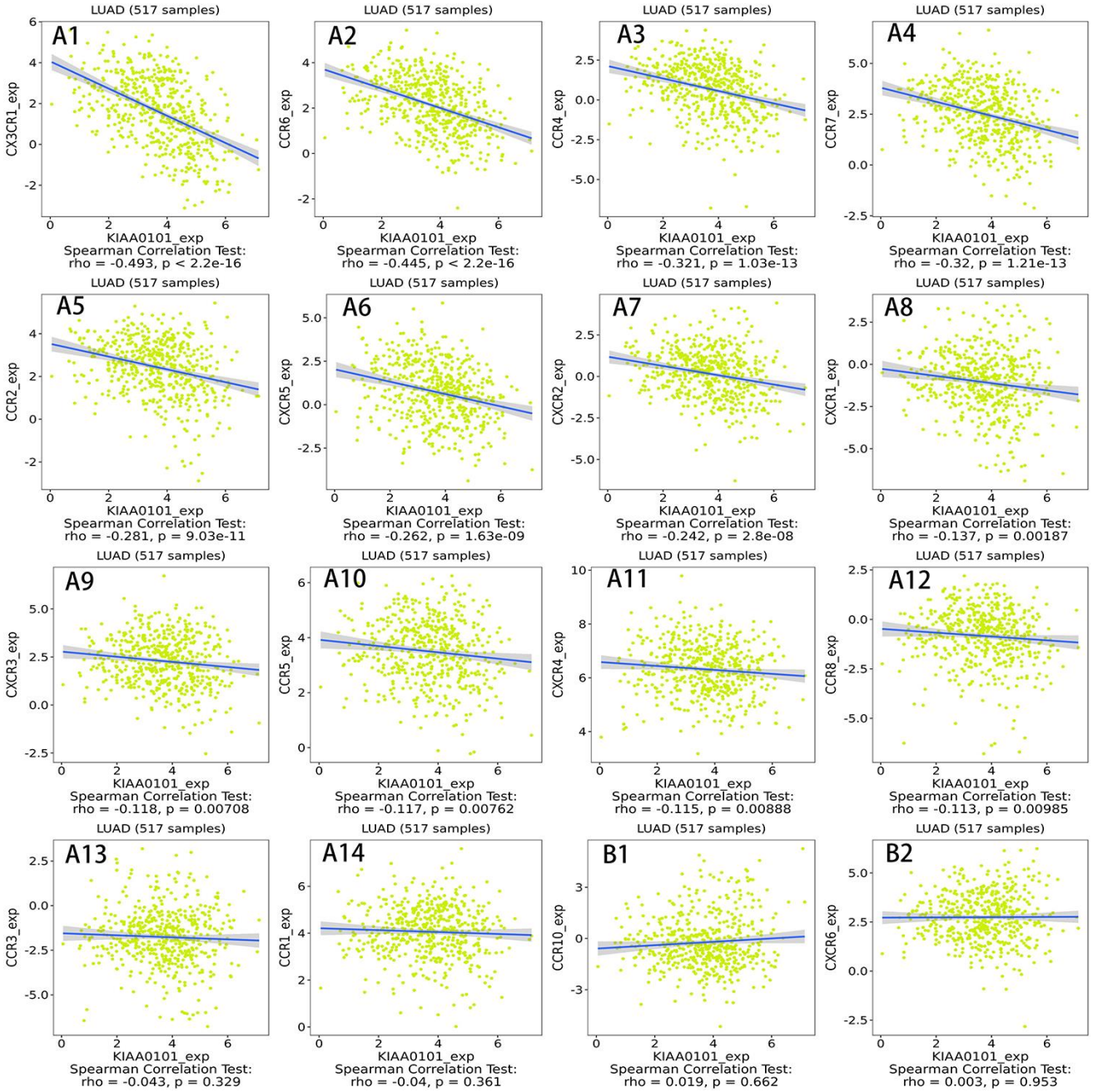
Supplementary Figure 10. The correlation between the expression of *KIAA0101* and immunostimulators. (**A1**–**A29**) are scatter plots of the negative correlations between *KIAA0101* expression and Immunostimulators in the treatment of lung adenocarcinoma. (**B1**–**B15**) are scatter plots of the positive correlations between *KIAA0101* expression and Immunostimulators in the treatment of lung adenocarcinoma.



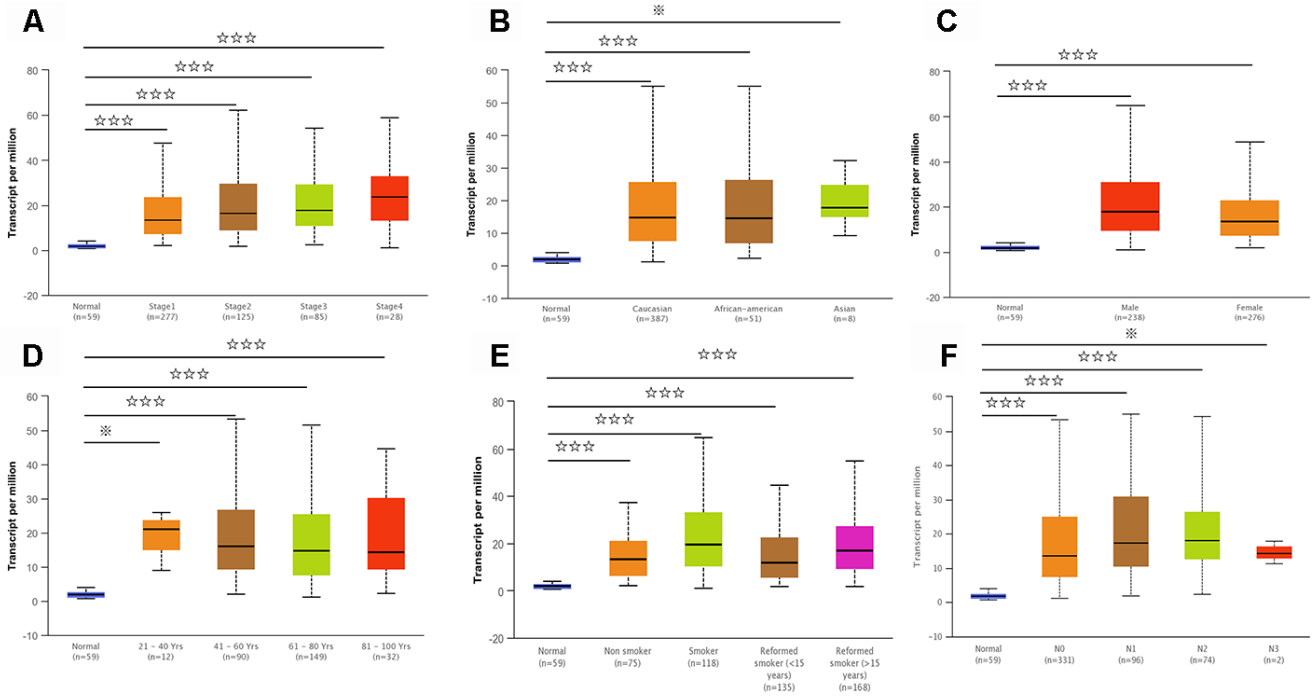
Supplementary Figure 11. The correlation between the expression of *KIAA0101* and MHC molecules. (**A1**–**A17**) are scatter plots of the negative correlations between *KIAA0101* expression and MHC molecules in the treatment of lung adenocarcinoma. (**B1**–**B3**) are scatter plots of the positive correlations between *KIAA0101* expression and MHC molecules in the treatment of lung adenocarcinoma.



Supplementary Figure 12. The correlation between the expression of *KIAA0101* and chemokines. (**A1**–**A17**) are scatter plots of the negative correlations between *KIAA0101* expression and chemokines in the treatment of lung adenocarcinoma. (**B1**–**B19**) are scatter plots of the positive correlations between *KIAA0101* expression and chemokines in the treatment of lung adenocarcinoma.



Supplementary Figure 13. The correlation between the expression of *KIAA0101* and receptors. (**A1**–**A14**) are scatter plots of the negative correlation between *KIAA0101* expression and receptors in the treatment of lung adenocarcinoma. (**B1**–**B2**) are scatter plots of the positive correlation between *KIAA0101* expression and receptors in the treatment of lung adenocarcinoma.



Supplementary Figure 14. *KIAA0101* transcription in subgroups of patients with lung adenocarcinoma, stratified based on sex, age, and other criteria (UALCAN). (**A**) Individual cancer stages; (**B**) patient's ethnicity; (**C**) patient's sex; (**D**) patient's age; (**E**) patient's smoking habits; (**F**) patient's metastasis status. ※ P>0.05, ☆ P < 0.05; ☆☆, P < 0.01; ☆☆☆, P < 0.001.

Supplementary Tables

Supplementary Table 1. Details of the eight domain chains of KIAA0101.

Serial number	ID	POSITION	IDENTITY			MAX IDENTITY	RESOLUTION	IN COMP-LEX WITH	EXPERIMENT TYPE	INTERFACE CAVITY
			START	STOP	SCORE					
A	4D2G_D	52 - 69	52	69	100	100	2.65 Å	PCNA	X-RAY DIFFRACTION	√
B	4D2G_E	52 - 69	52	69	100	100	2.65 Å	PCNA	X-RAY DIFFRACTION	√
C	6EHT_D	52 - 71	52	71	100	100	3.2Å	PCNA	X-RAY DIFFRACTION	√
D	6EHT_E	52 - 71	52	71	100	100	3.2Å	PCNA	X-RAY DIFFRACTION	√
E	6GWS_D	41-72	41	72	100	100	3.2Å	PCNA	X-RAY DIFFRACTION	√
F	6GWS_E	41-72	41	72	100	100	2.9Å	PCNA	X-RAY DIFFRACTION	√
G	6GWS_F	41-72	41	72	100	100	2.9Å	PCNA	X-RAY DIFFRACTION	√
H	6IIW_B	2-11	2	11	100	100	1.699Å	UHRF1	X-RAY DIFFRACTION	√

Supplementary Table 2. Significantly enriched gene ontology (GO) annotations (cellular components) of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
condensed chromosome	66	0	RAD51, SPC25, CCNB1, BIRC5, NCAPG, ZWINT, MAD2L1, SKA3, NUF2, BUB1B, CENPA, SKA1, AURKB, NEK2, CENPW, HJURP, NDC80, CDCA5, NCAPH, BUB1, ZWILCH, CENPK, KIF2C, AURKA, CENPN, TOP2A, CENPM, PLK1, ERCC6L, CDT1, CHEK1, SPAG5, CENPH, SPC24, NUP37, BLM, CENPE, BUB3, CDK2, FANCD2, CENPO, CENPF, BRCA1, DSN1, MKI67, NCAPG2, H2AFX, HMGB2, SUV39H1, CBX3, TUBG1, KNTC1, PPP1CC, SMC2, BANF1, NCAPD2, SKA2, NUP107, BRCA2, NUP85, ITGB3BP, SYCE2, TOPBP1, DMC1, SMC4, INCENP.
chromosomal region	94	0	RAD51, OIP5, CDK1, SPC25, CCNB1, BIRC5, NCAPG, ZWINT, MAD2L1, SKA3, NUF2, BUB1B, CENPA, SKA1, AURKB, NEK2, ESCO2, CENPW, HJURP, TTK, NDC80, CDCA5, BUB1, ZWILCH, CENPK, KIF2C, AURKA, DSCC1, CENPN, CDCA8, CENPM, PLK1, MCM6, ERCC6L, CDT1, HELLS, CHEK1, SPAG5, CENPH, PCNA, SPC24, CENPI, NUP37, FEN1, CENPL, BLM, KIF18A, CENPE, MCM4, BUB3, SUV39H2, MCM2, CDK2, PIF1, DNA2, CENPO, CENPF, CHEK2, DSN1, H2AFX, MCM7, SUV39H1, MTBP, CBX3, RECQL4, KNTC1, PPP1CC, CENPP, CENPQ, PTGES3, NCAPD2, DYNLL1, SKA2, HAT1, NUP107, MCM5, MCM3, MSH2, BRCA2, NUP85, SSB, ITGB3BP, DMC1, INCENP, THOC3, XPO1, APEX1, XRCC5, KIF22, DCLRE1A, SEH1L, XRCC3, NSMCE2, RAD21.
mitochondrial protein complex	152	0	MRPL47, DNA2, MRPL11, MRPL42, TOMM5, MRPL3, MRPL21, NDUFA9, PPIF, MRPL13, NDUFA12, MRPL15, MRPS35, MRPL12, COX5A, CHCHD3, PNPT1, MRPS16, MRPL37, MRPS30, MRPL51, UQCRH, MRPS11, UQCRHL, MRPL52, MRPS15, MRPS22, TOMM40, MRPS12, MRPS17, TIMM8B, MRPS10, NDUFB5, TIMM9, MRPL35, COX7A2, NDUFB3, TIMM10, TIMM50, MRPL17, COX6A1, TOMM22, MRPL9, APOO, NDUFB4, MRPL48, MRPL27, MRPL30, MRPS7, CYC1, HSD17B10, MRPL44, MRPS33, VDAC1, MRPL36, MRPL22, PDK1, MRPS24, MRPL2, COX5B, UQCRFS1, CHCHD1, COX7A2L, C15orf48, ROMO1, NDUFS6, NDUFAB1, MRPL19, IMMT, MRPS18C, MRPL39, NDUFB9, MTX1, MRPL46, MRPL50, SUPV3L1, NDUFB6, MRPL32, SDHB, NDUFA8, TIMM17A, DAP3, MRPL16, NDUFB1, NDUFB8, UQCRC1, NDUFC2, GRPEL2, MFN1, MRPS5, MRPL18, NDUFA1, NDUFS3, GRPEL1, NDUFS1, MRPL10, NDUFS5, NDUFA6, NDUFV2, TIMM17B, MRPS28, DNAJC19, MTX2, UQCRQ, NDUFB11, MRPL33, MRPS9, IMMP1L, C12orf65, MRPL40, NDUFC1, SUCLG1, NDUFB2, MRPS14, NDUFS8, PARK7, DLAT, MRPS18A, KIAA0391, MRPL53, TOMM6, UQCRB, COX4I1, NDUFA11, NDUFA7, TIMM13, MRPL34, NDUFA4, NDUFA3, MRPS2, MRPS21, BCS1L, MTG1, MRPL24, CLPX, MRPL38, CHCHD10, MRPL28, TIMM22, FOXRED1, TOMM40L, SDHD, PMPCB, MRPL43, MRPS26, MRPL20, MRPL41, MPV17L2, NDUFS4, NDUFA5, MRPS34, NDUFA13.
ribosome	143	0	MRPL47, MRPL11, MRPL42, DENR, MRPL3, MRPL21, ZC3H15, MRPL13, MRPL15, MRPS35, MRPL12, HSPA14, PNPT1, MRPS16, MRPL37, MRPS30, MRPL51, MRPS11, MRPL52, MRPS15, MRPS22, MRPS12, MRPS17, MRPS10, MRPL35, MRPL17, NAA10, MRPL9, MRPL48, MRPL27, MRPL30, MRPS7, RPLP0, RPL39L, MRPL44, MRPS33, MRPL1, RPS7, MRPL36, RSL24D1, APEX1, MRPS23, MRPL22, PTCD3, MRPS24, MRPL2, EIF2AK2, CHCHD1, RPL22L1, NDUFAB1, MRPL19, MRPS18C, MRPL39, RPL27, MRPL46, MRPL50, MRPL32, DAP3, RPL26L1, MRPL16, RPL35A, RPL38, MRPS5, MRPL18, RPS19, RPS27A, LARP4, MRPL10, GADD45GIP1, RPL39, RPS16, MRPS28, RPS17, RPS10, RPL35, MCTS1, RPS21, NUFIP1, MRPL33, MRPS9, RPS3, C12orf65, MRPL40, RPS26, MRPS14, NCK1, RPSA, EIF2A, MRPS18A, MRPL53, RPL36A, RPS29, RPS18, RPL4, RPL7L1, RPL37, RPL24, RPL6, RPL8, RPL41, NDUFA7, MRPL34, RPL19, RPL23A, AURKAIP1, MRPS2, MRPS21, MTG1, MRPL24, ZNF622, MRPL38, RPS5, RPL30, MRPL28, RPLP1, RPS24, RPS15A, RPL31, EIF3H, MRPL43, NR0B1, RPS8, MRPS26, MRPL20, RPL18A, MRPL41, MPV17L2, MRPS34, RPL37A, RPL27A, RPL36AL, MRPL14, RPL18, RPL7A, RPS12, MRPL49, RPL5, RPS11, RPS15, RPS2, RPL23, RPL32, NSUN3.
spindle	70	0	CDK1, NUSAP1, CCNB1, BIRC5, MAD2L1, SKA3, KIF23, BUB1B, CDC6, PRC1, SKA1, AURKB, NEK2, TTK, CDC20, DLGAP5, KIF11, KIF20A, AURKA, RACGAP1, TPX2, KIFC1, KIF4A, CDCA8, PLK1, POC1A, CKAP2L, KIF15, SPAG5, SHCBP1, KIF18B, KIF14, ASPM, KIF18A, VRK1, ESPL1, ECT2, CENPE, TACC3, CKAP2, FAM83D, PSRC1, FBXO5, CDC7, KIF20B, CENPF, DSN1, CBX3, TUBG1, KNTC1, BCCIP, HAUS1, DYNLL1, SKA2, WDR62, RAE1, NUP85, TOPBP1, INCENP, HAUS2, HAUS6, MAPRE1, MAD2L2, HAUS8, KIF22, POC1B, CDC27, PRPF19, NEDD1, RAB11A.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 3. Significantly enriched gene ontology ()GO annotations (biological processes) of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
chromosome segregation	97	0	OIP5, NUSAP1, SPC25, CCNB1, BIRC5, NCAPG, ZWINT, MAD2L1, SKA3, NUF2, KIF23, BUB1B, CDC6, PRC1, SKA1, AURKB, NEK2, ESCO2, CENPW, CEP55, HJURP, TTK, CDC20, DLGAP5, NDC80, CDCA5, NCAPH, BUB1, KIF2C, RACGAP1, DSCC1, KIFC1, KIF4A, CENPN, CDCA8, TOP2A, PLK1, CDT1, SPAG5, RAN, KIF18B, KIF14, PTTG1, NUP37, FEN1, TRIP13, BLM, KIF18A, ESPL1, ECT2, CCNE2, EME1, CENPE, TACC3, FAM83D, BUB3, PSRC1, FBXO5, FANCD2, CENPF, BRCA1, DSN1, CCNE1, MKI67, BRIP1, TUBG1, KIF4B, CENPQ, ACTR3, SMC2, GEN1, NCAPD2, SKA2, NAA50, SYCE2, NAA10, DMC1, SMC4, INCENP, RAD51C, RAD18, RMI1, SRPK1, ANAPC5, MAD2L2, KIF22, KPNB1, FANCM, PHB2, SEH1L, XRCC3, CDC27, NSMCE2, RAD21, ANAPC11, RCC1, RAB11A.
DNA replication	96	0	RAD51, CDK1, CDC45, RRM2, CCNA2, CDC6, EXO1, ESCO2, MCM10, GINS1, GINS2, DSCC1, POLE2, MCM6, DTL, CDT1, CHEK1, RFC4, TIPIN, PCNA, DBF4, FEN1, BLM, GMNN, RNASEH2A, RFC5, RFC3, CCNE2, EME1, POLQ, MCM4, GINS4, RFC2, GINS3, MCM2, PRIM1, CDK2, FBXO5, CDC7, PIF1, DNA2, BRCA1, CHEK2, E2F8, CCNE1, POLA2, CHAF1B, BRIP1, WDHD1, DONSON, MCM7, TIMELESS, CLSPN, E2F7, RECQL4, SSBP1, SLBP, HMGA1, DUT, RPA3, CHAF1A, GEN1, STOML2, ATAD5, RRM1, RNASEH1, MCM5, MCM3, BRCA2, MCM8, GTPBP4, POLE3, DBF4B, RBBP7, RMI1, MSH6, SET, FANCM, RFWD3, POLD2, DNAJC2, PRIM2, NBN, FAF1, PPP2CA, KIN, CDK2AP1, RBBP8, POLE, SSRP1, LIG1, ATF1, POLD3, ZRANB3, DDX11, CDC34.
cell cycle checkpoint	60	0	CDK1, CCNB1, ZWINT, MAD2L1, CDC45, BUB1B, CDC6, AURKB, CDC25C, TTK, CDC20, NDC80, BUB1, ZWILCH, AURKA, GTSE1, TOP2A, PLK1, DTL, CDT1, CHEK1, TIPIN, PCNA, WDR76, TRIP13, BLM, EME1, BUB3, CDK2, E2F1, DNA2, CENPF, BRCA1, CHEK2, E2F8, BRIP1, H2AFX, DONSON, TIMELESS, CLSPN, E2F7, TRIAP1, KNTC1, GEN1, MSH2, PRMT1, TOPBP1, MSH6, MAD2L2, XRCC3, RFWD3, PRPF19, RINT1, TIPRL, NAE1, INTS7, NBN, ZNF207, PSMG2, TFDP1.
double-strand break repair	61	0	RAD51, CDC45, EXO1, ESCO2, RAD51AP1, CDCA5, GINS2, RAD54L, FOXM1, CHEK1, FEN1, TRIP13, BLM, EME1, RAD54B, PSMD14, FANCB, POLQ, GINS4, XRCC2, CDC7, DNA2, BRCA1, CHEK2, BRIP1, H2AFX, TIMELESS, UBE2N, RECQL4, UBE2V2, RPA3, GEN1, MSH2, BRCA2, MCM8, DMC1, SUMO1, RAD51C, PARP2, YY1, FIGNL1, RMI1, XRCC5, MAD2L2, PPP4C, DCLRE1A, DEK, XRCC3, RFWD3, NSMCE2, RAD21, PRPF19, DDX1, NBN, APTX, SFPQ, RBBP8, RECQL, TDP1, KDM1A, TDP2.
spindle organization	44	0	SPC25, CCNB1, NUF2, KIF23, PRC1, AURKB, NEK2, TTK, CDC20, NDC80, KIF11, AURKA, RACGAP1, TPX2, MYBL2, KIFC1, KIF4A, PLK1, POC1A, SPAG5, RAN, ASPM, ESPL1, CENPE, TACC3, STIL, PSRC1, FBXO5, CHEK2, STMN1, TUBB, TUBG1, KIF4B, BCCIP, HAUS1, WDR62, RAE1, HAUS2, HAUS6, MAPRE1, HAUS8, KPNB1, RCC1, RAB11A.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 4. Significantly enriched gene ontology (GO) annotations (molecular functions) of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading Edge Num	FDR	Leading Edge Gene
structural constituent of ribosome	109	0	MRPL47, MRPL11, MRPL42, MRPL3, MRPL21, MRPL13, MRPL15, MRPS35, MRPL12, MRPS16, MRPL37, MRPS30, MRPL51, MRPS11, MRPL52, MRPS15, MRPS22, MRPS12, MRPS17, MRPL35, MRPL17, MRPL9, MRPL27, MRPL30, MRPS7, RPLP0, RPL39L, MRPS33, MRPL1, RPS7, MRPL36, RSL24D1, MRPS23, MRPL22, MRPS24, MRPL2, RPL22L1, MRPL19, MRPS18C, RPL27, MRPL46, MRPL32, DAP3, RPL26L1, MRPL16, RPL35A, RPL38, MRPS5, MRPL18, RPS19, RPS27A, MRPL10, RPL39, RPS16, RPS17, RPS10, RPL35, RPS21, MRPL33, MRPS9, RPS3, RPS26, MRPS14, RPSA, MRPS18A, RPL36A, RPS29, RPS18, RPL4, RPL7L1, RPL37, RPL24, RPL6, RPL8, RPL41, NDUFA7, MRPL34, RPL19, RPL23A, MRPS2, MRPS21, MRPL24, RPS5, RPL30, MRPL28, RPLP1, RPS24, RPS15A, RPL31, MRPL43, RPS8, MRPL20, RPL18A, MRPL41, MRPS34, RPL37A, RPL27A, RPL36AL, MRPL14, RPL18, RPL7A, RPS12, MRPL49, RPL5, RPS11, RPS15, RPS2, RPL23, RPL32.
catalytic activity, acting on DNA	69	0	RAD51, CDC45, EXO1, GINS1, GINS2, TOP2A, POLE2, MCM6, RAD54L, ERCC6L, PCNA, NEIL3, FEN1, BLM, NME1, EME1, RAD54B, POLQ, MCM4, GINS4, UNG, PIF1, DNA2, POLA2, BRIP1, MCM7, TDG, RECQL4, DKC1, HMGA1, PTGES3, GEN1, ALKBH2, RUVBL1, DMC1, POLE3, DNMT3B, RAD51C, APEX1, XRCC5, DCLRE1A, SMUG1, FANCM, XRCC3, POLE4, RUVBL2, POLD2, SUPV3L1, DDX1, DHX36, NBN, APTX, TATDN1, RBBP8, RECQL, POLE, TERT, LIG1, TDP1, CHRAC1, POLD3, TDP2, ZRANB3, METTL4, DDX11, G3BP1, RPS3, POLB, APEX2.
single-stranded DNA binding	43	0	RAD51, CDC45, MCM10, RAD51AP1, MCM6, NEIL3, BLM, NME1, MCM4, PRIM1, HMGB2, HSPD1, MCM7, RECQL4, SSBP1, RPA3, SMC2, GEN1, POLR2H, MSH2, TSN, BRCA2, YBX1, POLR2D, DMC1, SMC4, SUB1, POLR2G, RAD18, NUP35, LRPPRC, DHX36, PRIM2, HMGB1, APTX, CNBP, RAD23B, RBBP8, WBP11, TDP1, TDP2, POT1, DDX11.
helicase activity	43	0	CDC45, GINS1, GINS2, MCM6, RAD54L, ERCC6L, HELLS, BLM, RAD54B, MCM4, GINS4, MCM2, PIF1, DNA2, BRIP1, MCM7, EIF4A3, RECQL4, MCM5, MCM3, RUVBL1, MCM8, DDX47, DDX52, TTF2, DDX55, EIF4A1, XRCC5, FANCM, RUVBL2, DDX18, SUPV3L1, DDX1, DHX36, NBN, DDX56, HLTF, RECQL, GTF2F2, DDX23, DDX10, ZRANB3, DDX11.
catalytic activity, acting on RNA	122	0	EXO1, FEN1, RNASEH2A, RAD54B, PRIM1, PIF1, FARSB, CPSF3, EIF4A3, RPP30, EXOSC2, POP7, PNPT1, TARS, PTRH2, METTL2A, EXOSC8, RNASEH1, DARS, POLR2H, MARS, POP1, TSN, YARS2, EXOSC3, EMG1, DUS4L, POLR2D, POLR3G, FBL, EXOSC9, POLR2K, MRPL44, POLR2G, RPP40, GARS, DARS2, EIF4A1, APEX1, PUS1, EXOSC1, NOP2, METTL8, RPP25, DDX18, TRMT61B, SUPV3L1, POLR2J, ZNRD1, GATC, TRMT112, POP5, RARS, DDX1, METTL1, POLR1C, DHX36, PRIM2, METTL2B, POLR2I, DDX56, PPP1R8, TSEN15, POLR2F, IARS, METTL6, RBMX2, RPP38, TERT, CARS, POP4, DDX23, EXOSC5, DDX10, TDP2, NSUN2, KARS, FTSJ1, MARS2, EXOSC4, RPP21, G3BP1, EDC3, POLR3F, DTD1, NARS, ERI1, KIAA0391, MED20, WDR4, DHX37, DHX15, RNASEH2B, QRSL1, POLR3K, TWISTNB, ERI3, TRIT1, NARS2, TRMT12, TRPT1, DDX21, FARSA, YARS, WARS, SARS2, TFB2M, CNOT7, TRMT5, TGS1, DCPS, PUS3, ISG20L2, POLR2B, THUMPD3, DBR1, CDKAL1, THUMPD2, TFB1M, RARS2, EXOSC10, POLR3D.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 5. Significantly enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotations of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
Cell cycle	48	0	CCNB2, CDK1, CCNB1, MAD2L1, CDC45, BUB1B, CCNA2, CDC6, CDC25C, TTK, CDC20, BUB1, CDC25A, PLK1, MCM6, CHEK1, PCNA, DBF4, PTTG1, ESPL1, PKMYT1, CCNE2, MCM4, BUB3, MCM2, CDK2, E2F1, CDC7, CHEK2, CCNE1, E2F2, MCM7, ANAPC7, YWHAQ, SKP2, MCM5, MCM3, HDAC2, CDK4, YWHAZ, ANAPC5, MAD2L2, YWHAG, RBL1, CDC27, E2F3, RAD21, ANAPC11.
Ribosome	100	0	MRPL11, MRPL3, MRPL21, MRPL13, MRPL15, MRPL12, MRPS16, MRPS11, MRPS15, MRPS12, MRPS17, MRPS10, MRPL35, MRPL17, MRPL9, MRPL27, MRPL30, MRPS7, RPLP0, MRPL1, RPS7, MRPL36, RSL24D1, MRPL22, MRPL2, RPL22L1, MRPL19, MRPS18C, RPL27, MRPL32, RPL26L1, MRPL16, RPL35A, RPL38, MRPS5, MRPL18, RPS19, RPS27A, MRPL10, RPL39, RPS16, RPS17, RPS10, RPL35, RPS21, MRPL33, MRPS9, RPS3, RPS26, MRPS14, RPSA, MRPS18A, RPL36A, RPS29, RPS18, RPL4, RPL37, RPL24, RPL6, RPL8, RPL41, MRPL34, RPL19, RPL23A, MRPS2, MRPS21, MRPL24, FAU, RPS5, RPL30, MRPL28, RPLP1, RPS24, RPS15A, RPL31, RPS8, MRPL20, RPL18A, RPL37A, RPL27A, RPL36AL, MRPL14, RPL18, RPL7A, RPS12, RPL5, RPS11, RPS15, RPS2, RPL23, RPL32, UBA52, RPL36, MRPL4, RPL17, RPS20, RPS3A, RPL29, RPS13, RPL10L.
Proteasome	40	0	PSMD14, PSMA4, PSMD12, PSMA5, PSMB3, PSMA2, PSMB7, POMP, PSMB5, PSMA3, PSMD11, PSMC4, PSMA7, PSMC6, PSMA1, PSMA6, PSMC2, PSMB4, PSME2, PSMC1, PSMB2, PSMB1, PSME3, PSMD3, PSMB6, PSMC3, PSMD7, PSMD2, PSMD13, PSMD4, PSMD8, PSMD1, PSMD6, PSME4, ADRM1, IFNG, PSMB8, PSMB9, PSMC5, PSME1.
Spliceosome	63	0	SNRPA1, SNRPD1, SNRPF, SNRPG, LSM5, SNRPB, HNRNPC, MAGOHB, EIF4A3, LSM2, SMNDC1, SNRPE, PPIL1, SNRPC, PHF5A, PPIH, MAGOH, SNRPD2, PRPF4, SNRPB2, LSM4, LSM3, SNRNP27, SNRPA, TRA2B, BUD31, THOC3, RBM17, USP39, PRPF40A, EFTUD2, SNRPD3, PRPF19, LSM6, ISY1, SNRNP40, CWC15, NCBP1, WBP11, DDX23, LSM7, NCBP2, TXNL4A, SF3B5, BCAS2, U2AF2, PUF60, RBMX, HNRNPK, SNW1, U2AF1, PLRG1, DHX15, PCBP1, SF3B4, HNRNPA1, SF3A3, CDC5L, PRPF38A, TCERG1, PQBP1, HSPA1B, HNRNPA1L2.
DNA replication	27	0	POLE2, MCM6, RFC4, PCNA, FEN1, RNASEH2A, RFC5, RFC3, MCM4, RFC2, MCM2, PRIM1, DNA2, POLA2, MCM7, SSBP1, RPA3, RNASEH1, MCM5, MCM3, POLE3, POLE4, POLD2, PRIM2, POLE, LIG1, POLD3.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 6. Significantly enriched kinase-target networks of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
Kinase_CDK1	74	0	NUSAP1, CCNB1, BIRC5, PBK, NCAPG, RRM2, BUB1B, CENPA, PRC1, CDC25C, CEP55, CDC20, DLGAP5, CDCA5, BUB1, KIF11, KIF2C, TPX2, TK1, TOP2A, CDC25A, DTL, FOXM1, ERCC6L, CHEK1, RFC4, SPAG5, FEN1, BLM, ESPL1, UHRF1, RFC5, RFC3, ECT2, NME1, CKAP2, TMPO, RFC2, E2F1, UNG, CDC7, KIF20B, EZH2, LMNB1, BRCA1, PAICS, MKI67, STMN1, FANCG, EIF4EBP1, MCM7, GMPS, SLBP, MAPK6, HMGA1, DUT, NME2, BRCA2, LDHA, ZC3HC1, FBXO43, LMNB2, XPO1, KIF22, USP14, DNM1L, CDC27, ANAPC11, NEDD1, CSNK2B, RCC1, LBR, USP1, PPP1CA.
Kinase_PLK1	32	0	RAD51, CCNB1, BIRC5, BUB1B, CDC6, PRC1, CDC25C, CEP55, KIF2C, RACGAP1, GTSE1, TOP2A, CDC25A, FOXM1, ERCC6L, RAN, ESPL1, PKMYT1, STIL, FBXO5, BRCA1, CHEK2, CLSPN, ANAPC7, CENPQ, RUVBL1, BRCA2, FBXO43, YY1, CDC27, NEDD1, SUZ12.
Kinase_AURKB	35	0	NUSAP1, BIRC5, KIF23, CENPA, AURKB, NDC80, CDCA5, KIF2C, RACGAP1, CDCA2, KIF4A, CDCA8, PLK1, SHCBP1, CKAP2, DSN1, MKI67, HIST1H3B, DDX52, INCENP, PPHLN1, YY1, CCDC86, DEK, HIST1H3C, HIST1H3I, HMGN2, MPHOSPH10, HIST1H3F, NSUN2, RPS10, HIST1H3G, HIST1H3J, RBMX, KRT8.
Kinase_CDK2	73	0	RRM2, CCNA2, CDC6, CDC25C, CDC20, DLGAP5, NCAPH, TPX2, MYBL2, TK1, DTL, FOXM1, CDT1, CHEK1, DIAPH3, BLM, UHRF1, MCM4, MCM2, CDK2, E2F1, UNG, CDC7, EZH2, CENPF, BRCA1, CCNE1, PAICS, MKI67, E2F2, STMN1, MCM7, GMPS, EIF4A3, TUBG1, C9orf40, ANAPC7, HMGA1, NUP107, SKP2, C2orf49, MCM3, TSN, BRCA2, TFAM, ITGB3BP, ZC3HC1, TOPBP1, LMNB2, RAD18, ANAPC5, KIF22, RBL1, CEP76, DNM1L, CDC27, E2F3, ANAPC11, CSNK2B, ERAL1, MTHFD1L, TBCE, PPP1CA, NBN, PYCR1, NPM1, SCML2, RBBP8, HIST1H1E, ANAPC10, MTA2, LIG1, TSR1.
Kinase_ATR	20	0	FANCI, GINS2, CHEK1, DBF4, BLM, MCM2, E2F1, FANCD2, BRCA1, CHEK2, H2AFX, CLSPN, FANCA, MCM3, DCK, XRCC3, NBN, NPM1, RBBP8, TDP1.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 7. Significantly enriched miRNA-target networks of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
GAGCCTG,MIR-484	40	0.010236	DPYSL2, EZH1, DLEC1, HSPG2, PTPRE, PLEKHH2, FOXO4, SCARA3, MYCBP2, HIPK1, PITPNA, NFIA, ZFYVE1, LBH, HTT, EDA, PRKCB, PTGER4, PRRT2, SLC6A1, TAF1L, SORBS2, FAM13A, FRMPD4, DACH1, TRIOBP, ZYG11B, KLF12, BCL11A, GAPVD1, KDM4A, MINK1, MAPKAPK2, HIVEP2, DENND5A, SNN, ACVR1B, HLA-DOB, PTPRF, WDR90.
CAGCACT,MIR-512-3P	53	0.025589	HLF, PLEKHM1, TAL1, RTN4RL1, ATXN7, C1orf21, SGSM2, TLN1, TRIM3, UBL3, ATXN1, PTPRT, FOXN3, FEM1C, MLLT6, IP6K1, RNF38, TNRC6B, ETV1, GRM7, 8-Mar, PPP3CA, USP47, TP53INP1, ESRP2, TRHDE, MBNL2, PDIK1L, PCDHAC2, TRIM2, PCDHA10, PPFIA2, NTNG1, GIGYF1, KLHL3, ARHGEF3, RSBN1, MRPS25, HCN4, SYT8, ARHGEF10, NFIB, ZDHHC9, NEO1, BHLHE41, PCDHA3, PCDH10, FRMD4A, BAHD1, KCNRG, SLC2A4, XIAP, CDK19.
TTTGCAC,MIR-19A,MIR-19B	148	0.029684	ZBTB4, CBX7, ADCY9, RAI2, DLC1, ROBO2, PRICKLE2, SMARCA2, TGFBR2, HLF, SCARF1, MACF1, ST3GAL5, CACNA1C, RALGPS1, ERBB4, CNTFR, PARM1, WDFY3, RFX1, ZDHHC7, LRIG1, ARHGEF12, NCALD, PTK2B, RGL1, KLF13, RXRA, ATP11A, CYLD, MECP2, SRGAP3, KCNA4, RBMS3, FOXF2, OLFM1, ARHGAP1, MYLIP, ABR, SPEN, UBL3, ATXN1, ARC, SPRYD3, SH3D19, STAT5B, AFF1, TMEM63B, ATRX, HIPK1, FEM1C, CPEB4, TNRC6A, SLC24A4, MLLT6, RNF145, CREBL2, FOXP1, RTN1, RNF38, NBEA, TNRC6B, KCNS2, KLHL20, ETV1, MEF2D, S1PR1, SOX6, PTPRG, DLX3, LBH, TSC1, GRIN2A, INO80, CCND2, TGOLN2, SYBU, SLC9A1, GRM7, HECW2, ZEB2, PDE7B, MED26, ITPR1, ATP10A, MAGI2, BACE1, ADCY7, BMPR2, TRAK2, WDR47, LRRK1, PRRT3, TP53INP1, ARID4B, AKAP1, ESRP2, ANKRD12, PHLDA3, PDE5A, KIF3A, ZFPM2, ZFYVE26, COL19A1, EVI5L, ENPP5, SHANK2, EPC2, PLXNC1, BPTF, ARRDC4, MBNL2, MID1IP1, SMOC2, KIAA1217, ZMYND11, ARRDC3, UCP3, PCDHAC2, PHF12, MFSD6, EPN2, WBP2, CGN, PPARA, PCDHA10, RIN2, FZD8, BSN, VGLL4, SDC1, ID4, SLC24A3, OGT, IGF2R, DDX6, IGSF3, SOX5, TESK2, PCDHA3, MINK1, ARGLU1, PCDH10, SLAIN1, BTBD7, NAV3, MAP3K12, PRUNE2.
GACAATC,MIR-219	60	0.034546	CGNL1, AKAP13, TGFBR2, ZCCHC24, CELF2, EPHA4, ERG, CBFA2T3, FAM160A2, PTPRU, SH3D19, DDAH1, CPEB3, NR2C2, TACC1, HAS3, KLF9, SNRK, ZC3H12B, MEF2D, SOX6, ETV5, MKNK2, PIP5K1C, GTPBP1, SYNGAP1, AFF4, MFNG, INPP5J, TRHDE, THRB, TMEM98, PODXL, AGPAT3, RECK, PDZRN4, FBXL17, NCOA1, NTNG1, PDGFRA, CXXC5, PCDH17, ZNF827, EGR3, FZD4, BTBD7, FBXO3, ZNF609, CCDC28A, PHACTR2, MAPT, ERGIC1, CPEB2, SDK1, FAM120C, HOMER2, KIF1B, UBR1, RORB, FMNL2.
CCCAGAG,MIR-326	58	0.039237	PLXNA2, CORO2B, PTCH1, SCARF1, KIAA0513, CELF2, ANKFY1, NCALD, H6PD, ZMIZ1, TLN1, BSDC1, SMAD6, ATXN1, SPRYD3, RPS6KA1, SEMA6D, RPS6KA3, UBXN10, VPS39, CIC, GNAO1, AHCYL2, MMP24, PALM, DIDO1, PTK7, PPP1R9B, SYNGAP1, CEBPA, TCF4, ATP8B2, FAIM2, EPN2, GGT7, RPGR, OGT, C9orf24, LRRC32, SSH2, NAV3, ST3GAL3, PAPPA, ZNF609, KCNIP2, CRIM1, BRPF3, NRP1, LRRTM1, BCL11B, AGPAT4, DLGAP2, KLHL14, EGLN2, RALGAPA1, NHS, SEC63, ATP2B2.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).

Supplementary Table 8. Significantly enriched transcription factor-target networks of KIAA0101 in lung adenocarcinoma (LinkedOmics).

Description	Leading EdgeNum	FDR	Leading Edge Gene
E2F_Q6	87	0	RAD51, CDK1, CDC45, RRM2, CDC6, H2AFZ, POLE2, CDC25A, ARHGAP11A, MCM6, DNAJC9, CDT1, PCNA, GMNN, SNRPD1, PKMYT1, RANBP1, TMPO, MCM4, GINS3, MCM2, FBXO5, E2F1, FANCD2, UNG, ZNF367, EZH2, ATAD2, E2F8, STMN1, POLA2, GAPDH, MCM7, SASS6, CLSPN, E2F7, PPP1CC, EED, PHF5A, GEN1, ATAD5, WDR62, MCM3, MSH2, MCM8, TOPBP1, SUMO1, TRA2B, CDCA7, DCK, HMGXB4, HIST1H2AH, KPNB1, RBL1, POLE4, E2F3, MXD3, TRMT6, DCTPP1, INTS7, HMGN2, PPP1R8, CTDSPL2, KCND2, SYNGR4, POLD3, UXT, NOLC1, ZCCHC8, HNRNPR, SLC38A1, MRPL40, ACBD6, MTF2, NASP, PRPS1, MAZ, CAND1, SMC6, YBX2, HIST1H4A, EHBP1, HNRNPD, AP4M1, HNRNPA1, PRKDC, PCSK1.
VE2F_Q4	86	0	RAD51, CDK1, CDC45, RRM2, CDC6, H2AFZ, POLE2, CDC25A, ARHGAP11A, MCM6, DNAJC9, CDT1, PCNA, GMNN, SNRPD1, PKMYT1, RANBP1, TMPO, MCM4, GINS3, MCM2, FBXO5, E2F1, FANCD2, UNG, ZNF367, EZH2, ATAD2, E2F8, STMN1, POLA2, GAPDH, MCM7, SASS6, CLSPN, E2F7, PPP1CC, EED, PHF5A, GEN1, ATAD5, WDR62, MCM3, MSH2, MCM8, TOPBP1, SUMO1, TRA2B, CDCA7, DCK, HMGXB4, HIST1H2AH, KPNB1, RBL1, POLE4, E2F3, MXD3, TRMT6, DCTPP1, INTS7, HMGN2, PPP1R8, CTDSPL2, KCND2, SYNGR4, POLD3, UXT, NOLC1, ZCCHC8, HNRNPR, SLC38A1, MRPL40, ACBD6, MTF2, NASP, PRPS1, MAZ, CAND1, SMC6, YBX2, EHBP1, HNRNPD, AP4M1, HNRNPA1, PRKDC, PCSK1.
E2F4DP1_01	94	0	CDK1, RRM2, CDC6, H2AFZ, POLE2, CDC25A, ARHGAP11A, MCM6, DNAJC9, PCNA, GMNN, SNRPD1, PKMYT1, RANBP1, TMPO, MCM4, GINS3, MCM2, FBXO5, E2F1, FANCD2, UNG, ZNF367, EZH2, ATAD2, E2F8, STMN1, FANCG, GAPDH, MCM7, SASS6, CLSPN, SUV39H1, E2F7, CBX3, MAPK6, HMGA1, EED, PHF5A, GEN1, ATAD5, WDR62, MCM3, MSH2, MCM8, TOPBP1, SUMO1, TRA2B, CDCA7, DCK, HMGXB4, RBBP7, EIF4A1, HIST1H2AH, RBL1, POLE4, E2F3, MXD3, TRMT6, PRPS2, DCTPP1, CTDSPL2, LIG1, SYNGR4, POLD3, NOLC1, H2AFV, ZCCHC8, HNRNPR, AP1S1, MRPL40, ACBD6, MTF2, NASP, PRPS1, MAZ, CAND1, SMC6, YBX2, HIST1H4A, EHBP1, HNRNPD, HNRNPA2B1, AP4M1, PRKDC, PCSK1, FANCC, NCL, USP37, POLD1, CDC5L, DNMT1, IER5L, NUP62.
E2F1_Q6	96	0	CDK1, RRM2, CDC6, H2AFZ, POLE2, CDC25A, ARHGAP11A, MCM6, DNAJC9, CDT1, PCNA, GMNN, SNRPD1, PKMYT1, RANBP1, TMPO, MCM4, GINS3, MCM2, FBXO5, E2F1, FANCD2, UNG, ZNF367, EZH2, ATAD2, E2F8, STMN1, FANCG, GAPDH, MCM7, SASS6, CLSPN, SUV39H1, E2F7, CBX3, HMGA1, EED, PHF5A, GEN1, ATAD5, WDR62, MCM3, MSH2, MCM8, GPN3, TOPBP1, SUMO1, TRA2B, CDCA7, DCK, HMGXB4, HIST1H2AH, KPNB1, RBL1, SLC25A3, POLE4, E2F3, MXD3, TRMT6, PRPS2, DCTPP1, CTDSPL2, SYNGR4, POLD3, NOLC1, H2AFV, SERBP1, ZCCHC8, HNRNPR, AP1S1, MRPL40, ACBD6, MTF2, NASP, ZBTB8OS, PRPS1, MAZ, CAND1, SMC6, YBX2, HIST1H4A, EHBP1, HNRNPD, HNRNPA2B1, AP4M1, PRKDC, PCSK1, FANCC, NCL, USP37, POLD1, CDC5L, DNMT1, IER5L, NUP62.
E2F_02	93	0	CDK1, RRM2, CDC6, H2AFZ, POLE2, CDC25A, ARHGAP11A, MCM6, DNAJC9, PCNA, GMNN, SNRPD1, PKMYT1, RANBP1, TMPO, MCM4, GINS3, MCM2, FBXO5, E2F1, FANCD2, UNG, ZNF367, EZH2, ATAD2, E2F8, STMN1, FANCG, GAPDH, MCM7, SASS6, CLSPN, SUV39H1, E2F7, CBX3, MAPK6, HMGA1, EED, PHF5A, GEN1, ATAD5, WDR62, MCM3, MSH2, MCM8, PTMA, TOPBP1, SUMO1, TRA2B, CDCA7, DCK, HMGXB4, EIF4A1, HIST1H2AH, RBL1, POLE4, E2F3, MXD3, TRMT6, PRPS2, DCTPP1, CTDSPL2, SYNGR4, POLD3, NOLC1, H2AFV, ZCCHC8, HNRNPR, AP1S1, MRPL40, ACBD6, MTF2, NASP, PRPS1, MAZ, CAND1, SMC6, YBX2, HIST1H4A, EHBP1, HNRNPD, HNRNPA2B1, AP4M1, PRKDC, PCSK1, FANCC, NCL, USP37, POLD1, CDC5L, DNMT1, IER5L, NUP62.

Abbreviations: LeadingEdgeNum, the number of leading edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA).