

DOCK2 is a Biomarker for Predicting the Prognosis of Lung Adenocarcinoma and Associated with Immune Infiltration

Jie He

Second affiliated hospital of Guangxi Medical university

Tongtong Zhang

Second affiliated hospital of Guangxi Medical University

Jian Sun

The First Affiliated Hospital of Chengdu Medical College: Chengdu Medical College The First Affiliated Hospital

Guangnan Liu (✉ gnliu63@hotmail.com)

Guangxi Medical University

Research Article

Keywords: Computational Biology; biomarkers, DOCK2

Posted Date: August 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-724569/v1>

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

[Read Full License](#)

Abstract

Background: dedicator of cytokinesis 2 is an atypical guanine exchange factor, which is particularly expressed in hematopoietic cells and modulates the activation along with the migration of immune cells by activating Ras-related C3 botulinum toxin substrate (Rac). Nevertheless, the role of DOCK2 in lung adenocarcinoma (LUAD) remains unclear.

Methods: Herein, we performed bioinformatics analysis of lung adenocarcinoma data abstracted from TCGA (The Cancer Genome Atlas) and GEO (Gene Expression Omnibus) data resources, and combined with web tools consisting of LinkedOmics, TIMER, and TISIDB. Finally, combined with clinical lung adenocarcinoma samples, we verified the expression of DOCK2 in tissue and its effect on the prognosis of lung adenocarcinoma.

Results: In the TCGA lung adenocarcinoma data set, the expression of DOCK2 was down-regulated in tumor tissues and verified in multiple independent cohorts. In addition, the low expression of DOCK2 indicates a poor overall survival(OS) in both TCGA and other GEO data sets and in our clinical samples. COX regression data illustrated that the low expression of DOCK2 was an independent predictor for OS. Functional network analysis shows that DOCK2 participates in immune response through interleukin production, neuroinflammatory response, acquired immune response, leukocyte migration and activation of lymph node cells, and is related to multiple immune-related pathways. Besides, the expression of DOCK2 was remarkably related with many kinds of tumor infiltrating immune cells.

Conclusion: combined with bioinformatics analysis and clinical sample verification, our study shows that DOCK2 can independently estimate the prognosis of lung adenocarcinoma and is related to immune infiltration. As a promising prognostic indicator and potential target of immunotherapy, the potential effect of DOCK2 on lung adenocarcinoma and its molecular mechanism are worthy of further discussion.

Introduction

Lung cancer is one of the tumors with the highest morbidity and mortality in the world. due to the lack of effective diagnostic markers, the early diagnosis rate of lung cancer is low. Once it is found that most of them are already in the late stage, the therapeutic effect and prognosis are poor, and the mortality is high(1). On the basis of histological characteristics, lung cancer is categorized into NSCLC (non-small cell lung cancer) and SCLC (small cell lung cancer), with an estimated 85% of lung cancer being NSCLC, in which lung adenocarcinoma is the primary tissue type of NSCLC (2). In general, the overall five-year OS (overall survival) of individuals diagnosed with advanced lung adenocarcinoma is less than 15%; and more than 60% of lung adenocarcinoma patients miss targeted genetic mutations that improve their OS (3). Although there are many ways to treat lung adenocarcinoma (lung adenocarcinoma, LUAD), such as surgery, radiotherapy, chemotherapy, targeted drug therapy and so on, the 5-year OS has not been remarkably improved(4–6). Therefore, the discovery of specific early screening indicators and treatment targets is critical in improving the OS of individuals with lung adenocarcinoma.

Dedicator of cytokinesis 2(DOCK2), a CDM family member, is a guanine exchange factor (Guanine nucleotide exchange factors), GEFs that mediates GTP-GDP exchange and specifically activates small G protein Ras-related C3 botulinum toxin substrate 1. It is primarily expressed in immune cells, modulates actin and cytoskeleton, and mediates cell adhesion and migration (7). As a regulatory molecule of immune cells, DOCK2 is involved in the modulation of lymph node migration, T cell activation and differentiation, cytotoxicity of NK cells, secretion of INF- γ , as well as bone marrow homing of many kinds of immune cells(8). In mechanism, the serine homology region (Src homology 3 (SH3) of DOCK2 contains 50 amino acids and has the ability to bind Engulfment and cell motility 1 (ELMOL1). ELMOL1 exists in intracellular vesicles and is a cytoplasmic junction protein. The binding of SH3 and ELMOL1 can inhibit the ubiquitination of DOCK2 and prevent DOCK2 degradation, thus activating Rac1, to play its immune mediating function(9). In addition, DOCK2 and DOCK5 can also synergistically stimulate PMA-triggered RAC activation, production of ROS (reactive oxygen species) and extracellular bactericidal network (Neutrophil extracellular traps, NETs (10). Studies have shown that DOCK2 expression is down-regulated in colorectal cancer, and overexpression of DOCK2 may participate in the recruitment of CD8 + T lymphocytes, which is an important indicator of good prognosis of colorectal cancer(11). Hu et al also showed that patients with acute concomitant leukocytes with high DOCK2 expression had a longer OS (12). However, some investigations have documented that DOCK2 is overexpressed in prostate cancer cell line PC3. After silencing DOCK2, the phosphorylation of Akt along with ERK1/2 decreased, and the cell proliferation decreased(13). These findings suggest that DOCK2 has many functions in many kinds of malignant tumors. Nonetheless, the possible function, as well as the mechanism of DOCK2 in lung adenocarcinoma and tumor immunology are not clear.

In this study, through TCGA and GEO database, more than 1200 patients with lung adenocarcinoma were selected as subjects to comprehensively analyze DOCK2 expression in lung adenocarcinoma and its relationship with the progression of lung adenocarcinoma, which was verified by clinical samples. Finally, the different expression levels of DOCK2 were associated with the changes of tumor immune microenvironment. The results show that the expression of DOCK2 plays an indispensable role in predicting the prognosis of lung adenocarcinoma and may become a new target for immune data of lung adenocarcinoma.

Materials And Methods

2.1 Data abstraction and processing

We abstracted the RNA expression patterns along with the clinical data of individuals with lung adenocarcinoma from the TCGA data resource. TCGA approved 513 lung adenocarcinoma samples that consist of prognostic data and 59 non-malignant lung tissue samples.

We abstracted the gene expression data sets (GSE10072, GSE116959, GSE31210, GSE7670, GSE32863, GSE75073, GSE72094) from GEO data resource. In these seven gene chip cohorts, GSE72094 contains

detailed clinical prognostic data, and the sample size is the largest, so it is used as a verification of prognostic analysis. The other 6 cohorts were employed to explore the differential expression of genes.

2.2 Analysis of differential expression of DOCK2 genes.

Evaluation of differential expression of DOCK2 by Wilcox test and Kruskal test.

2.3 Linked Omics data resource analysis.

The Linked Omics data resource (<http://www.linkedomics.org>) is used to analyze cancer-related multi-dimensional databases in 32 of the TCGA databases. The DEGs (differentially expressed genes) linked to DOCK2 were identified from the TCGA lung adenocarcinoma data set by the Link Finder module in the data resource, and Pearson correlation coefficient was employed to explore the correlation of the results, which were displayed on the volcano map and heat map respectively. Gene enrichment analysis in Link Interpreter module was used to assess the function module of gene ontology biological process (GO_BP) and (KEGG) pathway in Kyoto Encyclopedia of Genes and Genomes.

2.4 TIMER and ESTIMATE data resource analysis

TIMER (Tumor Immune Estimation Resource) data resource (<https://cistrome.shinyapps.io/timer/>) is a comprehensive database for systematic assessment of immune infiltration in 32 different kinds of cancer. The TIMER data resource employs deconvolution to show the abundance of tumor invading immune cells (tumor-infiltrating immune cells, TIICs) in lung adenocarcinoma samples from TCGA and GEO data sets. Pearson correlation analysis was employed to explore the relationship of DOCK2 with the degree of immune cell invasion.

The "estimate" package of R software was used to generate immune score for gene expression data, and the immune score represented the degree of immune cell infiltration in tumor tissue.

2.5 TISIDB data analysis.

TISIDB data (<http://cis.hku.hk/TISIDB>) is a portal for the interaction between oncology and immune system that integrates a variety of data types. The website covers 988 reported anti-tumor genes related to immunity, high-throughput screening genes, molecular images and para-cancerous data, as well as a large number of immunological data collected from other public data resources. Herein, TISIDB provides us with the relationship between DOCK2 and lymphocytes, immunomodulators and chemokines.

2.6 Collection of clinical tissue specimens.

60 lung adenocarcinoma tissue samples and 60 para-cancerous tissue samples were obtained from the Department of Respiratory and critical Medicine of the first affiliated Hospital of Chengdu Medical College from March 2014 to December 2015. The para-cancerous tissue was at least 5cm away from the tumor. The inclusion criteria were as follows: (1) the histological type of lung adenocarcinoma was confirmed by pathology, and (2) the patients did not receive systematic radiotherapy and chemotherapy

before operation. (3) all the patients had complete clinical information. Including: age, sex and staging, and complete follow-up information in order to count the survival time of patients. This study has been approved by the Ethics Committee of the first affiliated Hospital of Chengdu Medical College, and the project number is 2020CYFYIRB-BA-121, in which all procedures involving human subjects are in line with the Helsinki Declaration (revised in 2013). All patients received informed consent.

2.7 Detection of mRNA transcription level in clinical tissue samples of DOCK2 by RT-qPCR.

The lung adenocarcinoma specimens frozen at -80 °C in our hospital were taken out and thawed into tissue blocks of about 50-100mg and put into 2ml enzyme-free tubes. The total RNA, extracted from the samples by TRIzol method was used to detect the purity and concentration of the prepared RNA by ultraviolet spectrophotometer. The ratio of OD260/OD280 was between 1.8 to 2.0. the quality of RNA was considered to be qualified. According to the instructions of cDNA synthesis kit (TaKaRa Company, Japan), the reverse transcription conditions of cDNA, were as follows: 37 °C, 30 min, 98 °C, 5 min, -20 °C. RT-qPCR preparation system: 10 μ L SYBR Green (Japan TaKaRa Co., Ltd.), upstream and downstream primers 0.8 μ L, 2 μ L cDNA, 6.4 μ L without enzyme water. The RT-qPCR reaction was carried out in ABI 8000 real-time quantitative PCR, with GAPDH as the internal reference, and the reaction conditions were set as follows: pre-denaturation at 95 °C, 30s at 95 °C, 5s at 60 °C, 34s at 60 °C, 40 cycles and annealing at 60 °C for 30s. DOCK2 upstream primer is 5'-TGAAGCTGGACCACGAGGTAGA, downstream primer is GCCTTTGACCAGGTTACGAAG-3'; GAPDH upstream primer is 5'- ATTTGCGTCATCCTTGCG, downstream primer is GACCTTCACCTTCCCATGG-3'. Gene expression levels were calculated relative to the housekeeping gene GAPDH.

3 Statistical analysis.

The R software version 3.6.1 was used to analyze the data statistically. The measurement data are expressed by $(\bar{x} \pm s)$, and the counting data are expressed by frequency and constituent ratio. The measurement data in accordance with normal distribution were compared by independent sample t-test and paired sample t-test, and the counting data were compared by chi-square test. The data that do not accord with normal distribution are tested by Wilcoxon test. Kaplan-Meier approach was employed to compare the OS between high and low DOCK2 expression groups. Univariate along with multivariate Cox regression analysis were employed to identify the independent predictors of OS. $P < 0.05$ defined statistical significance.

Results

4.1 Low expression of DOCK2 in lung adenocarcinoma.

In six lung adenocarcinoma studies based on GEO and TCGA databases, DOCK2 expression in lung adenocarcinoma was lower in contrast with that in non-malignant tissues (Fig. 1A-E). DOCK2 also showed low expression in 192 pairs of lung adenocarcinoma and the paired para-carcinoma tissue in GEO, as well as TCGA data resources (Fig. 1F-H). In TCGA data resource, ROC curve was employed to

explore the differentiation effect of DOCK2 in lung adenocarcinoma and para-cancerous tissues. The area under the curve of DOCK2 is (AUC) 0.792 (Fig. 1I), indicating that DOCK2 may be a potential molecular marker of lung adenocarcinoma.

4.2 Relationship of DOCK2 expression with clinic-pathological parameters in patients with lung adenocarcinoma.

As the mechanism of DOCK2 in lung adenocarcinoma is not clear, the study of the relationship of DOCK2 expression with clinic-pathological features is helpful to elucidate the role of DOCK2 in lung adenocarcinoma. The data demonstrated that the expression level of DOCK2 was remarkably related with the staging changes of T, M and TNM in the data of TCGA lung adenocarcinoma, as indicated in Fig. 2A, C, D. The patients with high expression of DOCK2 had lower T, M and TNM stages, indicating that DOCK2 may be involved in delaying the progression of lung adenocarcinoma.

Besides, to understand the prognostic impact of DOCK2 expression on lung adenocarcinoma, Cox proportional hazard regression model was employed to explore the prognostic predictors. On the basis of the median value of DOCK2 expression (the median value is 2.448), individuals with lung adenocarcinoma were stratified into two groups: DOCK2 high expression group and DOCK2 low expression group. Univariate analysis illustrated that low DOCK2 expression was related with short OS. Other clinical features, consisting of T, N, M and TNM, were also related to OS. In order to verify the prognostic value of DOCK2 in lung adenocarcinoma, multivariate analysis was performed. The data illustrated that only the expression of DOCK2, as well as TNM stage were independently correlated with OS (Fig. 2E, F), indicating that the expression of DOCK2 is not only helpful in the diagnosis of lung adenocarcinoma, but also plays a better role in assessing the clinical prognosis of patients than T stage, N stage and M stage.

4.3 DOCK2 co-expression network in lung adenocarcinoma.

To understand the biological function of DOCK2 in lung adenocarcinoma, the LindFinder module of LinkedOmics website was employed to detect the co-expression model of DOCK2 in TCGA-LUAD. As illustrated in Fig. 3A, 11013 genes (dark red dots) were positively linked to DOCK2 and 8975 genes (dark green dots) were negatively linked to DOCK2. Figures 3B and 3C represent the heat maps of the first 50 genes positively and negatively linked to DOCK2, respectively. GO annotations illustrated that the genes co-expressed by DOCK2 were primarily involved in interleukin production, neuroinflammatory response, acquired immune response, leukocyte migration, activation of lymph node cells and so on. As illustrated in Fig. 3E, the KEGG pathway analysis showed autoimmune thyroid disease, Staphylococcus aureus infection, intestinal immune network producing IgA, allograft rejection and so on (Fig. 3F).

These results suggest that DOCK2 expression network has a broad range of influences on immune activation in lung adenocarcinoma.

It is worth to note that the first 50 genes positively related to DOCK2 have a high likelihood of becoming low-risk markers of lung adenocarcinoma, and the HR values of most genes are less than 1. On the contrary, in the first 50 genes negatively associated with DOCK2, most of the genes had HR values greater than 1, suggesting that these genes are risk genes, as shown in Fig. 4A, 4B.

4.4. The relationship of DOCK2 with the level of immune invasion

We searched the TIMER database for whether the expression of DOCK2 affected the level of immune cell invasion in lung adenocarcinoma. Pearson correlation analysis illustrated that DOCK2 expression was positively linked to B cells, CD4 T cells, CD8 T cells, dendritic cells, macrophages and neutrophils (Fig. 5A). In the GSE72094 dataset, the positive relationship of DOCK2 expression with these immune cells in the TCGA-LUAD cohort has also been well verified (Fig. 5B).

We then used ETSIMATE algorithm to analyze if DOCK2 expression was related with the total immune infiltration level of lung adenocarcinoma. The results demonstrated that there was a remarkable relationship of DOCK2 with immune score (Immunescore) in both TCGA and GEO lung adenocarcinoma data sets (Fig. 5C-D). In addition, patients who had high immune scores had poorer OS in contrast with those with low immune scores, which was in agreement with the data of univariate prognostic analysis of DOCK2 expression (Fig. 5E-H).

4.5 the relationship between DOCK2 and immune molecules.

To deepen the understanding of the relationship of DOCK2 with immune infiltration, we studied the relationship of the expression of DOCK2 with diverse immune signals, consisting of immune-related signals, three immunomodulators, chemokines and receptors of 28 types of T lymphocyte studied by Charoentong et al(14).

The correlation between DOCK2 expression and various immune characteristics was abstracted from TISIDB data resource. Figure 6A indicates the correlation of DOCK2 with (TILs) in tumor-invading lymph node cells, including Treg_abundance, Tem_CD8_abundance, NK_abundance, and NKT_abundance. Immunomodulators can be further divided into three groups, which contain immunosuppressants, immunostimulators and (MHC) molecules. Figure 6B shows that the immunosuppressant associated with DOCK2 is CSF1R_exp, TIGIT_exp, BTLA_exp, and IL10_exp. Figure 6C shows that the immunostimulant associated with DOCK2 is TNFSF13B_exp, IL2RA_exp, CD27_exp and CD40_exp. Figure 6D shows that the MHC molecule associated with DOCK2 is HLA-DOA_exp, HLA-E_exp, HLA-DRB1_exp and HLA-B_exp. Figure 6E demonstrates that the chemokine associated with DOCK2 is CCL19_exp, CCL4_exp, CXCL12_exp as well as CCL18_exp. Figure 6F illustrates that the receptor associated with DOCK2 is CCR4_exp, CCR8_exp, CXCR5_exp and CXCR4_exp.

Hence, this study verified that DOCK2 is widely involved in the regulation of various immune molecules in lung adenocarcinoma and affects the immune infiltration of tumor microenvironment.

4.6 Expression of DOCK2 in lung adenocarcinoma in clinical tissue specimens

We collected tissue samples from 60 cases of lung adenocarcinoma treated by surgical resection at The First Affiliated Hospital of Chengdu Medical College, and DOCK2 expression in lung adenocarcinoma and corresponding non-malignant tissues was assayed by RT-qPCR. The data illustrated that the mRNA with DOCK2 was 0.531 (0.217,1.211), which was lower in contrast with that of 3.284 (2.271, 4.325), $Z = -4.999$, $P < 0.05$. (Fig. 7A). DOCK2 expression was remarkably linked to clinical stage and tumor size ($P < 0.05$) (Table1). Prognostic analysis demonstrated that the patients were stratified into high DOCK2 expression group and low DOCK2 expression group on the basis of the median expression of DOCK2 evaluated by RT-qPCR. The prognosis of low DOCK2 expression group was worse relative to that of high DOCK2 expression group (Fig. 7B). It is consistent with the results predicted by bioinformatics analysis.

Table 1
Correlation between DOCK2 expression and patient characteristics

Characteristics	n = 60	DOCK2(mRNA)		Chi-square	P-value
		High (n = 30)	Low (n = 30)		
Gender				0.271	0.602
Male	26	12	14		
Female	34	18	16		
Age (year)				0.069	0.793
≤ 60	35	18	17		
> 60	25	12	13		
Stage				5.079	0.024
I/II	42	25	17		
III/IV	18	5	13		
Tumor diameter (cm)				4.356	0.037
≤ 5	45	26	19		
> 5	15	4	11		
Differentiation				0.606	0.436
High/moderate	33	15	18		
Poor	17	15	12		

Discussion

As far as we know, DOCK2 expression in lung adenocarcinoma and its effect on the prognosis of lung adenocarcinoma have not been studied. In this study, bioinformatics analysis was conducted using gene

expression profile information from TCGA and GEO databases. The results showed that DOCK2 may be a potential moderate molecular marker of lung adenocarcinoma, and the low DOCK2 expression in lung adenocarcinoma was linked to higher clinico-pathological stage, shorter OS and poor prognosis. These results were confirmed in the clinical samples we collected. After that, we studied the co-expression genes and regulatory networks of DOCK2. Finally, we analyzed the relationship of DOCK2 with immune invasion or immune signal, and found that DOCK2 was linked to most immune marker genes. The purpose of our work is to guide the research on immunotherapy of lung adenocarcinoma in the future.

Dedicator of cytokinesis protein is a multi-domain guanine nucleotide exchange factor of RHO GTPases, which regulates the motility of intracellular actin(15). DOCK2 is an important member of the DOCK-A subfamily(16). In recent years, as a new mutation gene, DOCK2 gene has been recognized as a susceptible gene with high mutation rate in a variety of tumor populations(17). Nonetheless, there is a lack of research on the expression of DOCK2 in non-malignant lung tissue and lung adenocarcinoma. Herein, the results of bioinformatics analysis illustrated that DOCK2 expression in lung adenocarcinoma was down-regulated and verified by the RT-qPCR along with Western blot data. Moreover, DOCK2 expression in early lung adenocarcinoma was higher in contrast with that in advanced lung adenocarcinoma, demonstrating that DOCK2 might be a tumor repressor gene in lung adenocarcinoma. Ge(18) et al. showed that DOCK2 is also a tumor suppressor gene in colorectal cancer that exhibited high mutation, and the elevated DOCK2 expression indicates a good prognosis. Wang et al. (19) showed that DOCK2 is linked to the tumor immune microenvironment of Luminal breast cancer, and DOCK2 may be a prognostic marker. At present, most studies believe that DOCK2 gene plays a role by regulating Rac, but DOCK2 affects different functions in different immune cells. DOCK2 deficiency results in a reduction in the number and proportion of T lymphocytes in lymph nodes and spleen, which affects the positive, as well as negative selection of T cells in the process of T cell development. DOCK2 modulates the development and activation of B lymphocytes, affects the differentiation of plasma cells and the secretion of specific IgG antibodies. In addition, DOCK2 affects the cytotoxicity and degranulation of NK cells, and regulates the migration of neutrophils, the production of ROS and the formation of NETs. DOCK2 almost affects the whole process of the biological function of immune cells.

To further confirm the relationship of DOCK2 with immune responses in lung adenocarcinoma, we sought to reveal the role of DOCK2 in the construction of the immune microenvironment in lung adenocarcinoma. To elucidate the prospective functions of DOCK2 in lung adenocarcinoma, we performed a bioinformatic analysis of public databases. Transcriptome analysis of over 1200 clinical samples from four geographical regions demonstrated that DOCK2 mRNA expression was remarkably lower in lung adenocarcinoma in contrast with that in para-cancerous lung tissue, both in the overall sample and in paired samples from individual patients. the expression level of DOCK2 had a remarkable impact on clinical parameters such as primary tumor status and tumor pathological stage. These results were validated in the clinical samples collected. Multifactorial analysis further illustrated that DOCK2 expression was an independent predictor affecting the prognosis of individuals with lung adenocarcinoma.

LinkedOmics data resource analysis further illustrated that not only DOCK2 gene has a remarkable influence on the prognosis of patients with lung adenocarcinoma, most of the genes co-expressed with DOCK2, whether positive or negative correlation, are remarkably related with the prognosis of patients with lung adenocarcinoma. In addition, these genes co-expressed with DOCK2 are mainly enriched in immune-linked cascades, which is consistent with the recognized role of DOCK2 involved in immune regulation. It is suggested that DOCK2 may be involved in modulating tumor immune microenvironment and improving the prognosis of lung adenocarcinoma.

The TIMER data resource analysis more directly confirmed the above conclusions. The DOCK2 expression level in TCGA, as well as GSE72094 data was highly correlated with the invasion level of B cells, dendritic cells, CD4 + T cells, neutrophils, CD8 + T cells, and macrophages. In addition, the expression of DOCK2 was positively correlated with the immune score of lung adenocarcinoma samples based on ESTIMATE algorithm. According to the immune score or DOCK2 expression, the patients were stratified into different groups, and the prognosis was also different. It is suggested that the high expression of DOCK2 may improve the immune microenvironment of lung adenocarcinoma patients through the invasion level of immune cells, so as to improve the prognosis of individuals with lung adenocarcinoma. Miao et al. [11] have shown that colorectal cancer patients with high DOCK2 protein expression have a better prognosis, and the mechanism may be that CD8 + T cells have a protective effect on the body, suggesting that DOCK2 may delay tumor progression through immune regulation. In this study, the relationship between DOCK2 and CD8 + T cells was shown to be $r = 0.451$ in TCGA lung adenocarcinoma and 0.5746 in GSE72094 lung adenocarcinoma. The result is similar to that of Miao. It has been reported that DOCK can participate in the innate immune response of microglia by regulating the secretion of microglial cytokines, consisting of tumor necrosis factor- α along with monocyte chemoattractant protein-1(20). Hence, we speculate that DOCK2 may enhance the invasion of CD8 + T cells by modulating the secretion of some cytokines, which is confirmed in the analysis of the relationship between DOCK2 and chemokines from TISIDB database. Then, according to the median value of DOCK2 mRNA, we divided the collected clinical samples into high DOCK2 expression group and low DOCK2 expression group. After 5 years of follow-up, we found that the OS of individuals with lung adenocarcinoma in high DOCK2 expression group was longer. Increasing the expression of DOCK2 is involved in delaying the progression of lung adenocarcinoma. The reason may be that the overexpression of DOCK2 is related to the promotion of CD8 + T lymphocyte infiltration. T cells are mainly responsible for the cellular immunity of the body, but also an important part of humoral immunity. There are T cell receptors (TCRs) on the surface of T cell membrane. TCR is the core structure of antigen recognition, as well as immune response. There are mainly two types of TCR $\alpha\beta$ and TCR $\gamma\delta$. Usually referred to as $\alpha\beta$ T cells, T cells are the primary lymphocyte subsets involved in adaptive immune response. Mature $\alpha\beta$ T cells are mostly CD4 or CD8 single positive cells. CD8 + T lymphocytes account for 30% of peripheral blood lymphocytes, and their main role is to directly kill target antigens (such as viruses and tumor cells), so they are called cytotoxic T lymphocyte (Cytotoxic T lymphocyte, CTL). Cytotoxic T lymphocytes first express Fas ligand FasL which docks to Fas indicated by antigen, which enhances apoptosis of tumor cells. On the other hand, CTL can also secrete perforin to dissolve the capsule of tumor cells, and can

also secrete interferon- γ and tumor necrosis factor- α , which inhibit the formation of target antigen DNA(21). Because CTL can kill tumor cells, this may explain why patients with high DOCK2 expression have a better prognosis. Lu et al. (22) also found that DOCK2 may be a potential prognostic index to improve the OS of brain metastasis in the genomic study of the immune microenvironment of brain metastasis of breast cancer. Many studies have shown that DOCK2 has extensive effects on the immune microenvironment, as well as immune cell invasion of patients with a variety of cancers. Therefore, DOCK2 is worthy of further research in the field of tumor immunotherapy.

Although the results of this study enhance our comprehension of the relationship of DOCK2 and lung adenocarcinoma, there are some limitations. First of all, in order to fully clarify the specific role of DOCK2 in the onset and progress of lung adenocarcinoma, several clinical factors and parameters need to be considered, such as the details of patients receiving treatment. However, this information is lacking or inconsistent in the public database, because these trials are conducted in different centers and the standards cannot be unified. Second, in the current study, the number of healthy subjects as a control group is relatively small, so the conclusion needs to be further confirmed by expanding the number of healthy subjects. Third, although the multicenter studies in the public database make up for the deficiency in the number of single-center studies, the retrospective study has its limitations, especially the disunity of intervention measures and the lack of specific treatment information. Finally, as the current studies are based on public databases and clinical samples, it is critical to further study the direct mechanism of DOCK2 in lung adenocarcinoma.

Conclusion

Herein, for the first time we established that the low expression of DOCK2 is closely linked to the progression, poor prognosis, as well as immune invasion of lung adenocarcinoma, which may promote tumorigenesis through abnormal inflammation and immune response. DOCK2 may become one of the targets of immunotherapy for lung adenocarcinoma. The direct mechanism between the low expression of DOCK2 and the progression and metastasis of lung adenocarcinoma needs to be further studied. This study provides a new idea for further elucidating the clinico-pathological significance and molecular pathogenesis of lung adenocarcinoma.

Abbreviations

DOCK2 Deducator of cytokinesis 2

Rac Ras-related C3 botulinum toxin substrate

LUAD lung adenocarcinoma

TCGA The Cancer Genome Atlas

GEO Gene Expression Omnibus

OS overall survival

NSCLC non-small cell lung cancer

SCLC small cell lung cancer

SH3 Src homology 3

ELMOL1 Engulfment and cell motility 1

GO Gene ontology

BP biological process

KEGG Kyoto Encyclopedia of Genes and Genomes

TIMER Tumor Immune Estimation Resource

TIICs tumor-infiltrating immune cells

Declarations

Data availability statement

The datasets GSE10072, GSE116959, GSE31210, GSE7670, GSE32863, GSE75073, GSE72094 used and analyzed in the current study are available from the public database (<http://www.ncbi.nlm.nih.gov/geo/>). The datasets TCGA are available from the public database(<https://portal.gdc.cancer.gov/>).

Notes

Ethics approval and consent to participate

This study was approved by the Ethical Committee of the First Affiliated Hospital of Chengdu Medical College. Project number:2020CYFYIRB-BA-121. Ethics, consent, and permissions: Written informed consents were signed by all the patients involved to ensure their approval of the data used in this research.

Consent for publication

All the patients signed the agreement consents for publishing their individual clinical data.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Competing interests

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

None

Authors' contributions

JH and TTZ collected data from public datasets and analyzed the data. JH performed the statistical analysis. JH carried out the clinical samples gathering and in-house qRT-PCR experiments. GNL participated in the design of the study and language modification. JH and TTZ drafted the manuscript. BS and GNL conceived of the study and helped to draft the manuscript. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Footnotes

Jie He and Tong-tong Zhang contributed equally to this work.

Acknowledgements

Not applicable

References

1. Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med*, 2020; 41(1): 1–24.
2. Nasim F, Sabath BF, Eapen GA. Lung Cancer. *Med Clin North Am*. 2019;103(3):463–73.
3. Kris MG, Johnson BE, Berry LD, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *JAMA*. 2014;311(19):1998–2006.
4. Hoy H, Lynch T, Beck M. Surgical Treatment of Lung Cancer. *Crit Care Nurs Clin North Am*. 2019;31(3):303–13.
5. Ruiz-Cordero R, Devine WP. Targeted Therapy and Checkpoint Immunotherapy in Lung Cancer. *Surg Pathol Clin*. 2020;13(1):17–33.
6. Hirsch FR, Scagliotti GV, Mulshine JL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet*. 2017;389(10066):299–311.

7. Chen Y, Meng F, Wang B, et al. Dock2 in the development of inflammation and cancer. *Eur J Immunol.* 2018;48(6):915–22.
8. Kunimura K, Uruno T, Fukui Y. DOCK family proteins: key players in immune surveillance mechanisms. *Int Immunol.* 2020;32(1):5–15.
9. Chang L, Yang J, Jo CH, et al. Structure of the DOCK2-ELMO1 complex provides insights into regulation of the auto-inhibited state. *Nat Commun.* 2020;11(1):3464.
10. Watanabe M, Terasawa M, Miyano K, et al. DOCK2 and DOCK5 act additively in neutrophils to regulate chemotaxis, superoxide production, and extracellular trap formation. *J Immunol.* 2014;193(11):5660–7.
11. Miao S, Zhang RY, Wang W, et al. Overexpression of dedicator of cytokinesis 2 correlates with good prognosis in colorectal cancer associated with more prominent CD8(+) lymphocytes infiltration: a colorectal cancer analysis. *J Cell Biochem.* 2018;119(11):8962–70.
12. Hu N, Pang Y, Zhao H, et al. High expression of DOCK2 indicates good prognosis in acute myeloid leukemia. *J Cancer.* 2019;10(24):6088–94.
13. El-Haibi CP, Singh R, Sharma PK, Singh S, Lillard JW. Jr. CXCL13 mediates prostate cancer cell proliferation through JNK signalling and invasion through ERK activation. *Cell Prolif.* 2011;44(4):311–9.
14. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 2017;18(1):248–62.
15. Gadea G, Blangy A. Dock-family exchange factors in cell migration and disease. *Eur J Cell Biol.* 2014;93(10–12):466–77.
16. Guo X, Chen SY. Dedicator of Cytokinesis 2 in Cell Signaling Regulation and Disease Development. *J Cell Physiol.* 2017;232(8):1931–40.
17. Yu J, Wu WK, Li X, et al. Novel recurrently mutated genes and a prognostic mutation signature in colorectal cancer. *Gut.* 2015;64(4):636–45.
18. Wu M, Hamaker M, Li L, Small D, Duffield AS. DOCK2 interacts with FLT3 and modulates the survival of FLT3-expressing leukemia cells. *Leukemia.* 2017;31(3):688–96.
19. Ge W, Cai W, Bai R, et al. A novel 4-gene prognostic signature for hypermutated colorectal cancer. *Cancer Manag Res.* 2019;11:1985–96.
20. Wang Y, Zhu M, Guo F, et al. Identification of Tumor Microenvironment-Related Prognostic Biomarkers in Luminal Breast Cancer. *Front Genet.* 2020;11:555865.
21. Cimino PJ, Sokal I, Leverenz J, Fukui Y, Montine TJ. DOCK2 is a microglial specific regulator of central nervous system innate immunity found in normal and Alzheimer's disease brain. *Am J Pathol.* 2009;175(4):1622–30.
22. Iwahori K. Cytotoxic CD8(+) Lymphocytes in the Tumor Microenvironment. *Adv Exp Med Biol.* 2020;1224:53–62.

Figures

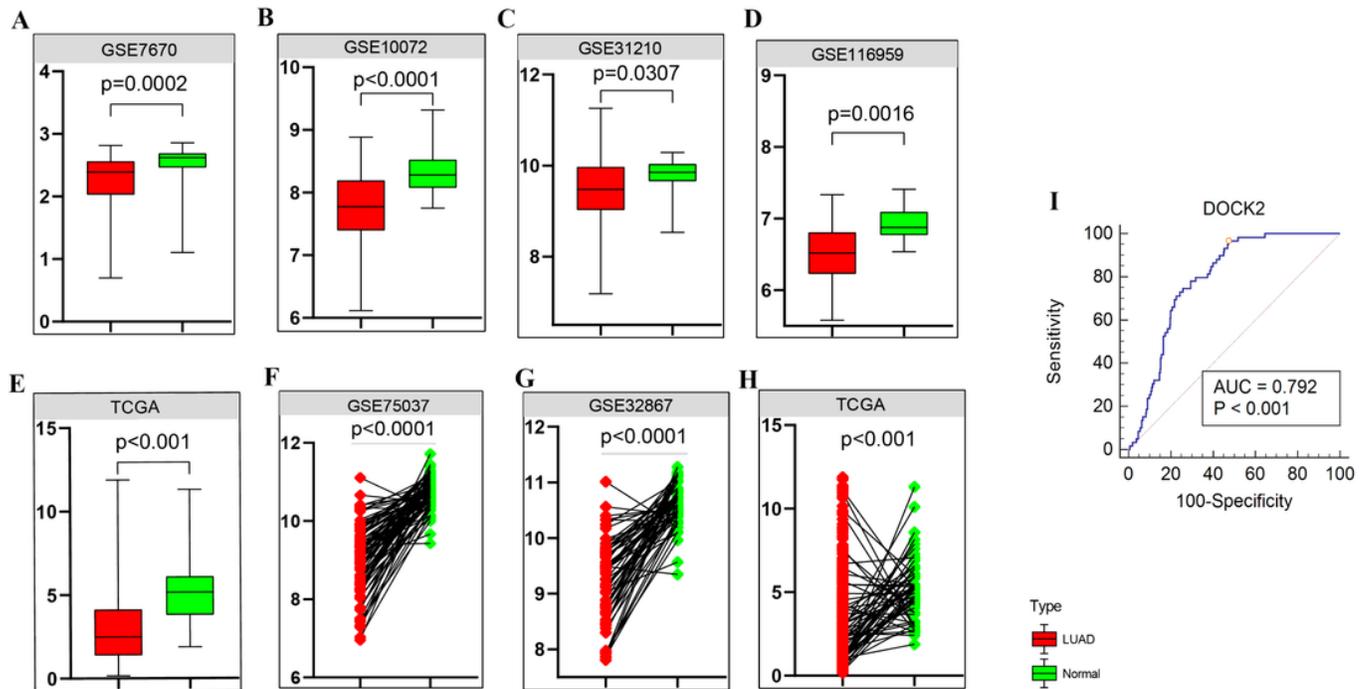


Figure 1

(A-E) Compared with normal lung tissue, DOCK2 was significantly down-expressed in LUAD tissue. (F-H) DOCK2 was down-regulated in LUAD tissue compared with 192 pairs of adjacent normal tissues from three different LUAD datasets. (I) The ROC curve for diagnosis of LUAD.

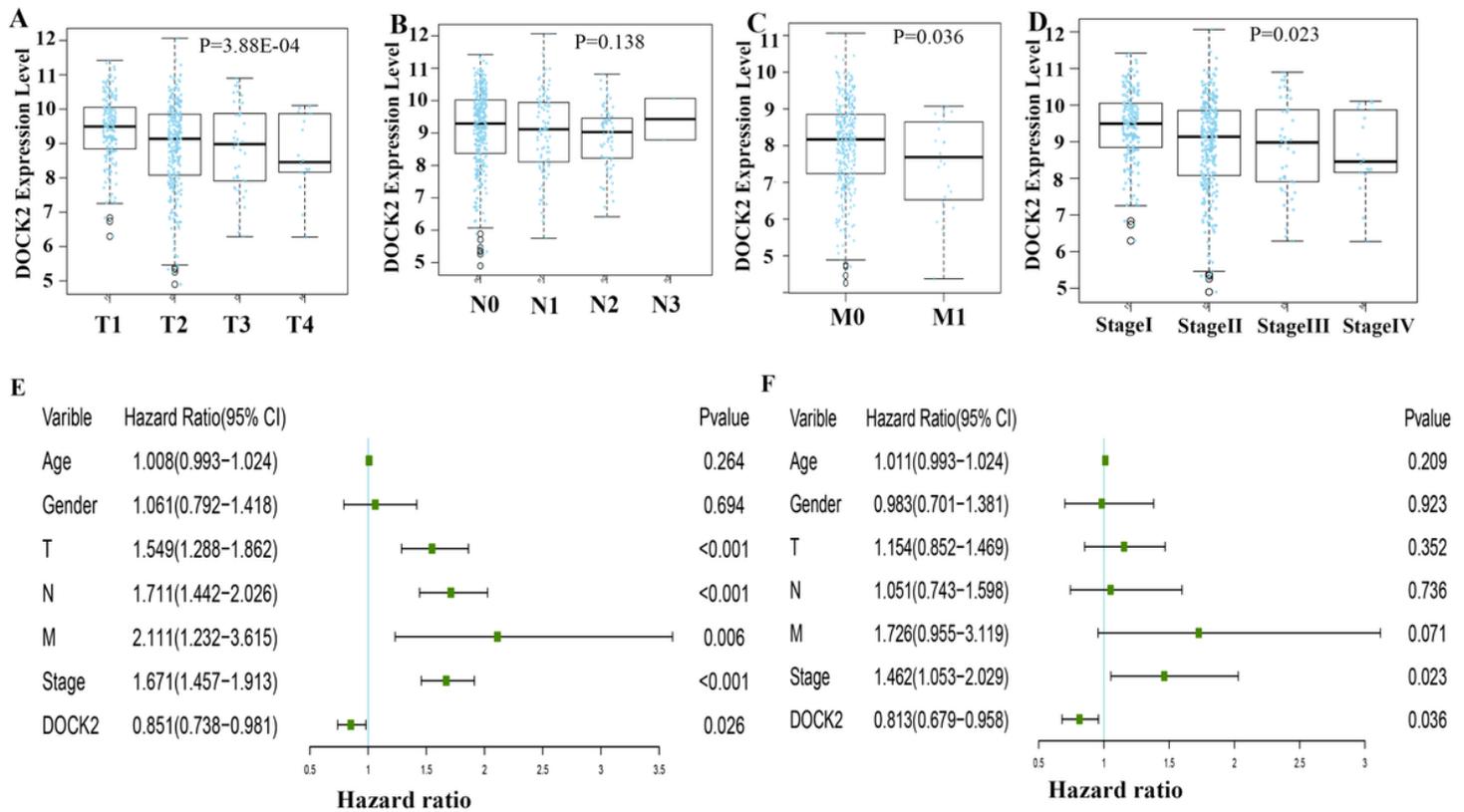


Figure 2

Association between DOCK2 expression and clinicopathological characteristics. (A) T stage; (B) N stage; (C) M stage; (D) TNM stage; (E-F) Univariate and multivariate Cox analysis of DOCK2 and pathological parameters

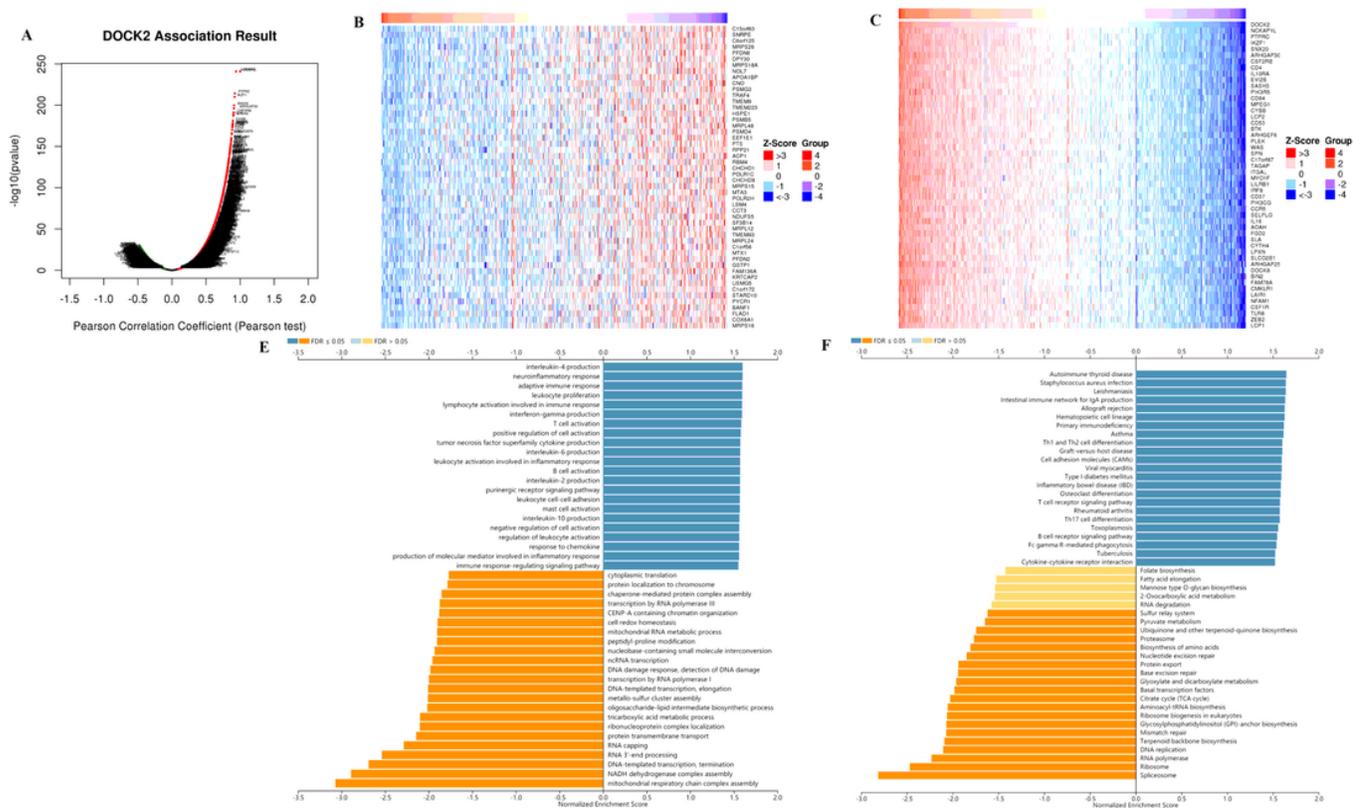


Figure 3

The co-expression genes with DOCK2 in LUAD from the LinkedOmics database. (A) The total significantly genes associated with DOCK2 from TCGA. (B-C) Top 50 genes positively and negatively linked to DOCK2 in LUAD exhibited by heat maps respectively. Red represents positively related genes and blue represents negatively related genes. (D-E) Significantly enriched GO terms and KEGG pathways of DOCK2.

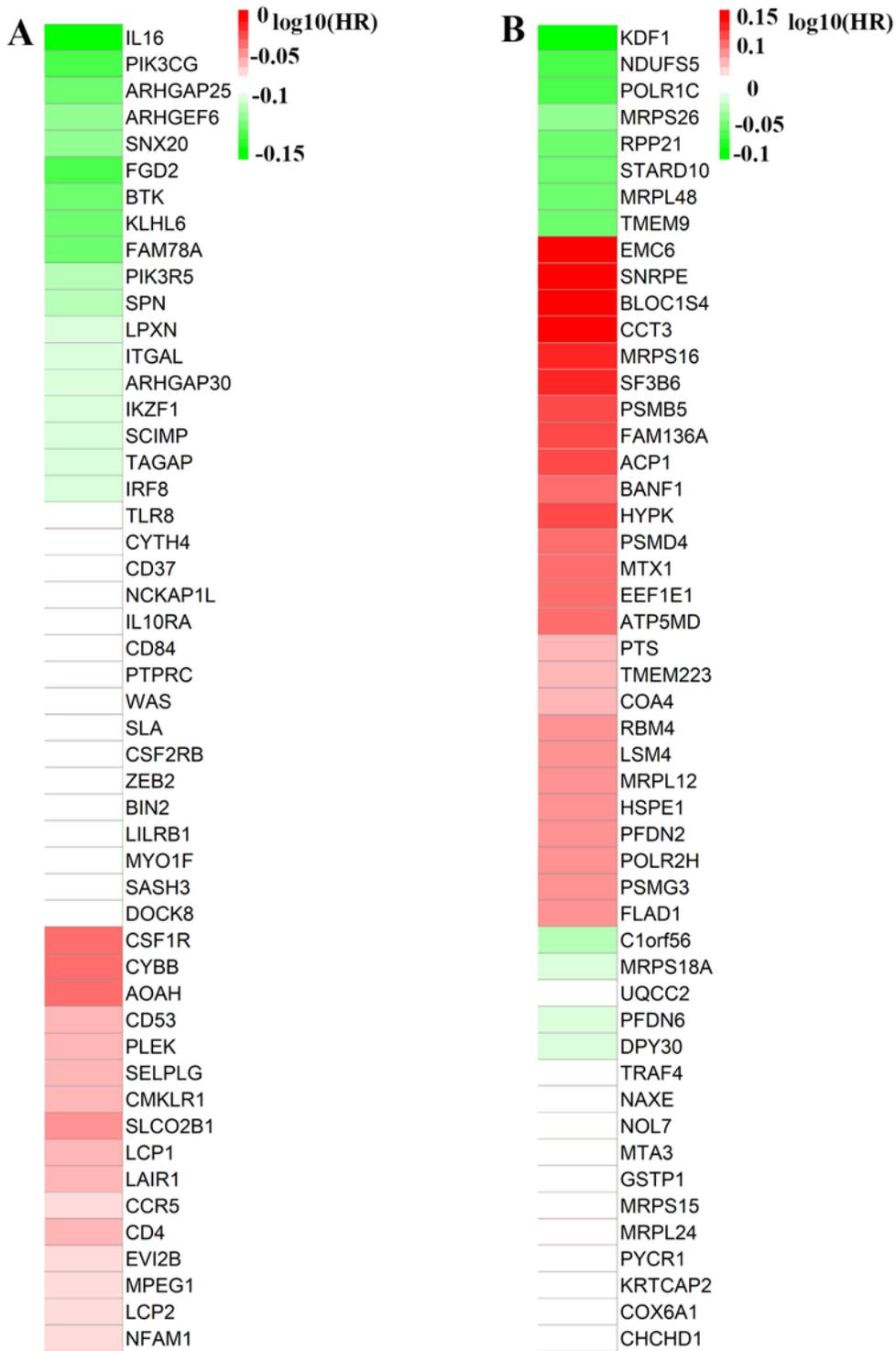


Figure 4

survival heatmap of the top 50 genes positively and negatively correlated with DOCK2 in LUAD.

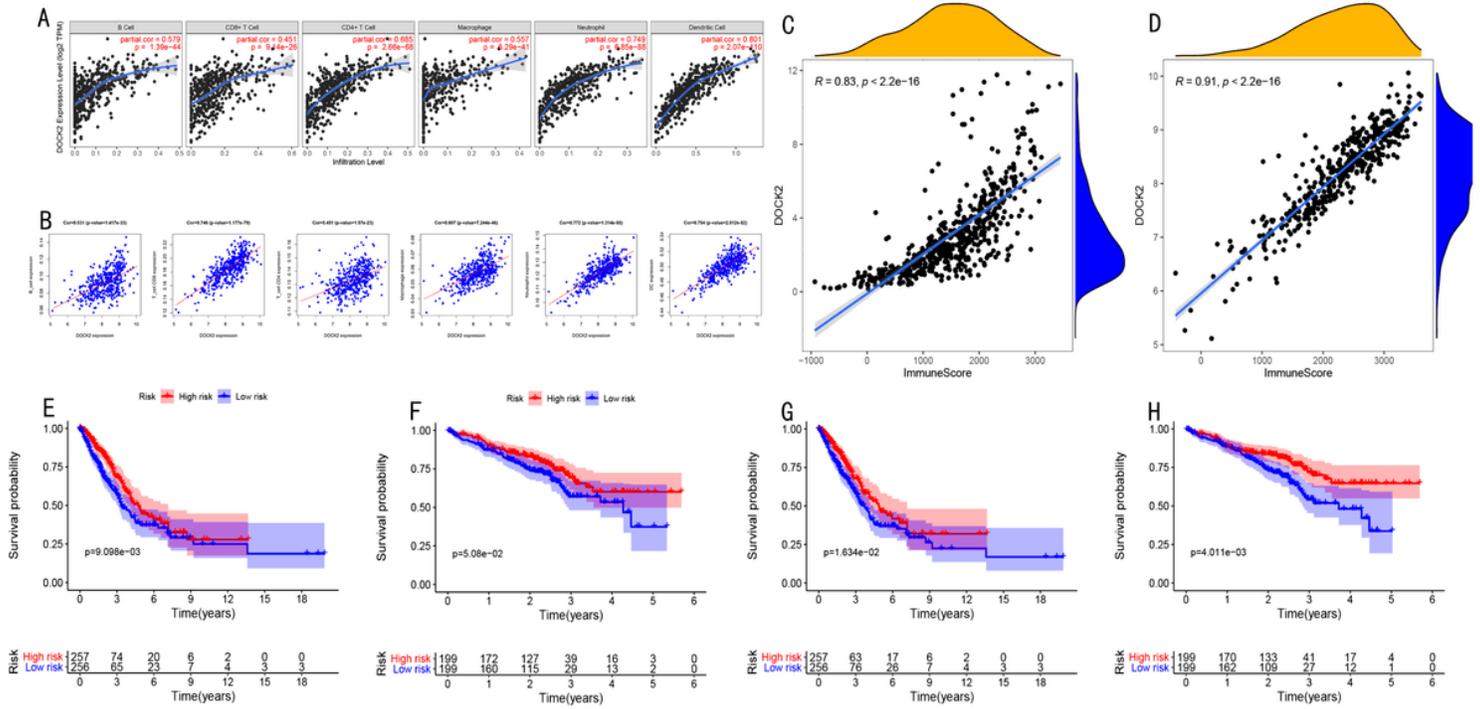


Figure 5

Correlations between the DOCK2 expression level and immune infiltration level based on TIMER database and R software. (A-B) The DOCK2 expression levels in the TCGA and GSE72094 dataset show a positive correlation with the immune infiltration level of B cells, CD4 T cells, CD8 T cells, dendritic cells, macrophages and neutrophils. (C-D) The DOCK2 expression has a significant positive correlation with the immune score of LUAD samples based on the ESTIMATE algorithm in the TCGA-LUAD And GSE72094 dataset. (E-F) Patients with higher immune scores have higher overall survival time in the TCGA-LUAD and GSE72094 dataset. (G-H) Patients with higher DOCK2 expression level have higher overall survival time.

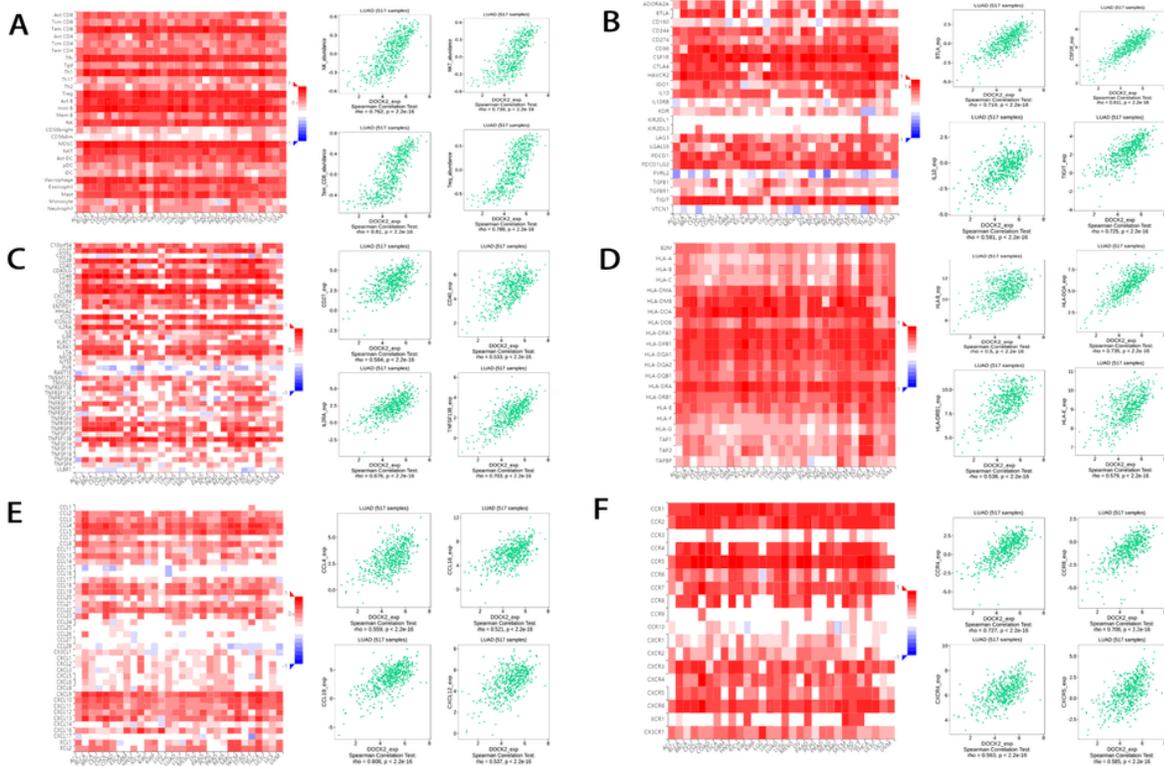


Figure 6

Associations of the DOCK2 expression level with lymphocytes, immunomodulators and chemokines in LUAD from TISIDB database. (A) Correlations between abundance of tumor-infiltrating lymphocytes (TILs) and DOCK2. (B-D) Correlations between immunomodulators and DOCK2. (E-F) Correlations between chemokines or receptors and DOCK2.

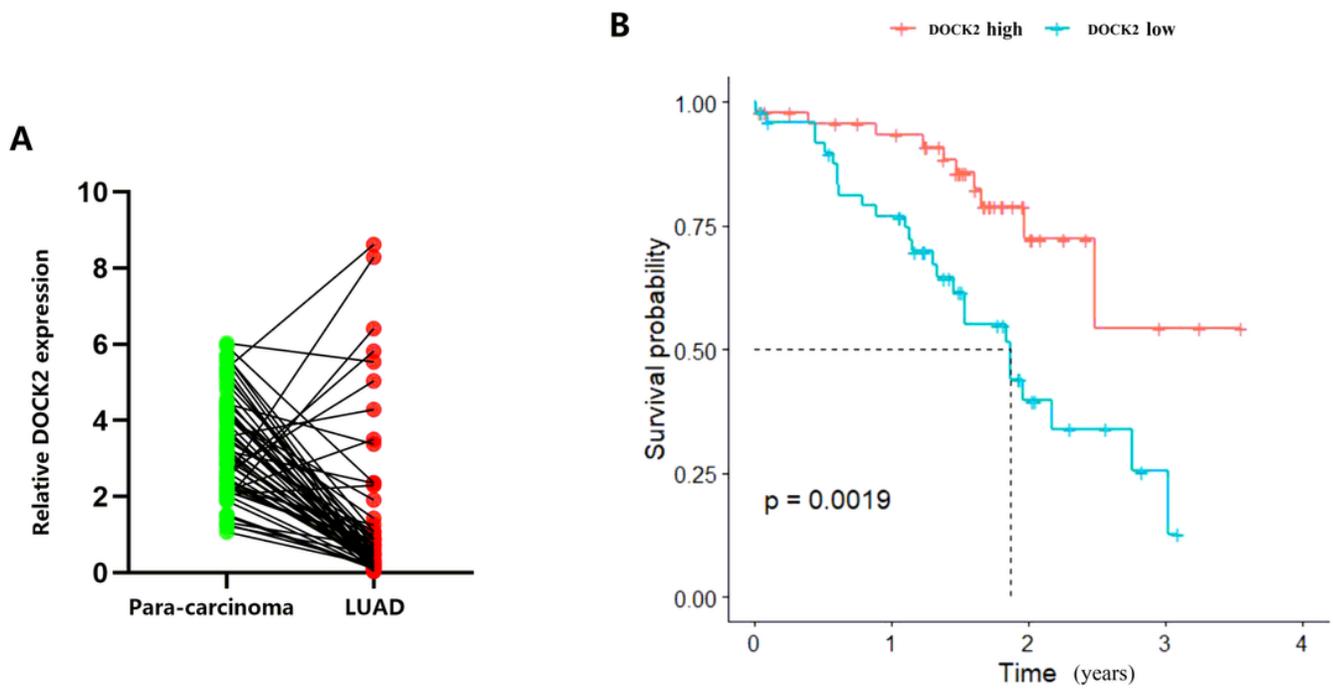


Figure 7

The correlation between the expression of DOCK2 from clinical tissue specimens and prognosis of LUAD
 (A) The DOCK2 expression levels in LUAD and adjacent non-cancerous tissues were detected by RT-qPCR.
 (B) Survival analysis of patients with low and high expression of DOCK2.