Int J Clin Exp Med 2020;13(10):8068-8085
www.ijcem.com /ISSN:1940-5901/IJCEM0117492

Original Article

Prognostic value and interrelated expression profile of GINS complex subunits in lung adenocarcinoma

Fangyan Zhong[1*], Jianxiong Deng[1*], Xiaotong Duan[2], Junming Lai[3], Xin Long[4], Ruoyan Shen[5], Yi Zhong[6], Hong Lin[1], Hui Luo[1]

[1]Department of Oncology, The First Affiliated Hospital of Nanchang University, Nanchang University, Nanchang 330006, Jiangxi, P. R. China; [2]Department of Radiation Oncology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, P. R. China; [3]Department of Radiation Oncology, Yiwu Central Hospital, The Affiliated Yiwu Hospital of Wenzhou Medical University, Yiwu 322000, Zhejiang, P. R. China; [4]Department of Radiation and Chemotherapy Oncology, Zhongnan Hospital, Wuhan University, Wuhan 430000, Hubei, P. R. China; [5]Department of Second Institute of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou 310013, Zhejiang, P. R. China; [6]Department of Acupuncture and Immune Rheumatism, First People's Hospital of Tongxiang, Tongxiang 314500, Zhejiang, P. R. China. [*]Equal contributors.

Received July 3, 2020; Accepted August 30, 2020; Epub October 15, 2020; Published October 30, 2020

Abstract: Lung adenocarcinoma (LUAD) is one of the common pathological types of lung cancer, it has a poor 5-year overall survival rate. The GINS complex was found to be a core component of replication helicases in eukaryotes, and its abnormal function may lead to the cause of human cancers and other diseases. However, few attempts have been made to find an association between GINS complex subunits (GINSs) and LUAD. Therefore, in this study, we will attempt to explore the prognostic value of GINSs in patients with LUAD and the interrelated expression profile. In this study, ONCOMINE database, UALCAN and Human Protein Atlas were used to investigate the role of GINSs in LUAD transcription expression profiles. KM-plotter and DriverDBv3 database were used to analyzed the correlation between GINSs gene expression and prognosis. By using LinkedOmics database, we screened GINSs-related differentially expressed genes and performed GO, KEGG analysis on these genes. We found that all GINS genes are highly expressed in LUAD and are associated with poor survival of patients with LUAD, especially those who have a smoking history. Additionally, we found that co-overexpression of two or more GINSs was associated with poor prognosis, suggesting the GINSs complex may co-operate to promote LUAD progression. We had also obtained some GINSs-related kinase targets and transcription factor targets, which may serve as a future research direction. We propose that individual GINSs or combinations of multiple GINS genes may be potential biomarkers for the prognosis of LUAD.

Keywords: Lung adenocarcinoma, GINS complex subunits, prognostic value, online databases, biomarker

Introduction

Lung cancer has the highest morbidity in the world and is a leading cause of death [1, 2]. Traditionally, it is classified into two primary groups, small cell versus non-small cell types. The non-small cell lung carcinoma (NSCLC) accounts for 80-85% of all cases of lung cancer and is mainly classified into two distinct pathological subtypes: squamous cell carcinoma (LUSC) and the adenocarcinoma (LUAD) [3]. Currently, LUAD is the most common subtype of lung cancer found in men and women [4]. Before the application of targeted therapies for oncogenic drivers, Platinum-based chemotherapy was the most common treatment for advanced LUAD. Recently, targeted drugs of oncogenic drivers are more widely used, and the prevailing oncogenic drivers in the field of lung cancer are mutations in EGFR, BRAF, KRAS and other important key drivers [5]. Nowadays, as research into the role of the immune system in tumor immunosurveillance deepens, immunotherapy drugs are beginning to show their potential in lung cancer treatment [6]. However, despite significant advances in treatment, the prognosis for LUAD is not encouraging. Hence, it is urgent to discover valuable molecular tar-

geted therapy through investigation and trials so as to judge the prognosis or optimize the treatment efficacy of cancer treatment in patients.

The GINS complex subunits (GINSs) consists of 4 genes, namely GINS1, GINS2, GINS3, GINS4 in the human genome, corresponding to Psf1, Psf2, Psf3, Sld5 proteins, respectively. The GINS complex, which was constituted by Psf1, Psf2, Psf3, Sld5 proteins, is the basis for the development of DNA replication forks and replicators in eukaryotes, and it is one of the core components of replication helicases in eukaryotes (CMG complex, including Cdc45-MCM-GINS). The GINS complex initiates the circular structure, regulates multifunctional proteins, growth factor signals and receptor molecules at the beginning of DNA replication. The GINS also participates in the initiation of DNA replication and cell cycle progression [7-9]. DNA helicases are necessary for genomic stability, and their abnormal function leads to the failure of faithful DNA replication inducing a loss of genome integrity, which may be the cause of human cancers and other diseases [10]. Hence, GINSs may be considered as potential targets for human cancers.

In the past decade, a lot of studies have reported on the role of GINSs in different cancers. It is shown that high expression of GINSs indicates a poor prognosis in high-grade prostate cancer and hepatocellular carcinoma, and it can also promote the growth of breast cancer [11-13]. Over-expression of GINS2 promotes tumor progression in early-stage cervical cancer [14]. GINS3 may become a potential prognostic biomarker for colorectal cancer and uterine endometrial carcinomas [15, 16]. Also, GINS4 promotes the growth of gastric and colorectal cancer by distinct pathways [17, 18]. There are some studies reporting GINSs in LUAD. However, research has mainly focused on the association between GINS3 and LUAD [19-21]. In the current study, we extended the research field based on different databases in order to identify the prognostic values of GINSs in LUAD. Some correlated genes of GINSs were also found, and their functional enrichments were discussed in combination with GINSs. We also analyzed transcription factor targets and kinase targets for GINSs and their related genes.

Material and methods

Oncomine database analysis

Oncomine (https://www.oncomine.org/) is a publicly accessible online database with plentiful information on cancer microarrays to help facilitate discovery from genome-wide expression analyses [22]. In our study, we found transcriptional expression of GINSs are over-expressed in distinct cancers and then compared the transcriptional data between LUAD tissues and normal lung tissues by Oncomine. The difference of transcriptional expression was compared by student's t-test. The cut-off of *p*-values and fold changes are listed as follows: *p*-value: 1E-4, fold change: 1.5, gene rank: 10%, data type: mRNA.

UALCAN analysis

Ualcan (http://ualcan.path.uab.edu/) is a comprehensive, user-friendly, and interactive web resource for analyzing cancer OMICS (The Cancer Genome Atlas, TCGA and MET500) data. It can provide graphs and plots depicting gene expression [23]. We used Ualcan to analyze mRNA expression of GINSs in LUAD tissues and their association with clinicopathologic parameters. The difference of transcriptional expression was compared with student's t-test and a *p*-value <0.05 was considered statistically significant.

Human protein atlas

The Human Protein Atlas (HPA, https://www.proteinatlas.org/) is a database which aims to map all the human proteins in cells, tissues and organs using the integration of various omics technologies, including antibody-based imaging, mass spectrometry-based proteomics, transcriptomics and systems biology [24]. We compared protein expression patterns of each GINSs in patients with LUAD to normal lung tissues.

Kaplan-Meier plotter analysis

Kaplan-Meier plotter (KM-plotter, http://www.kmplot.com/) is an online tool mainly based on the Gene Expression Omnibus (GEO) and TCGA data sources for survival analysis between gene expression and a variety of cancer

data [25]. The desired probe ID was determined according to the file of probe sets provided by KM-plotter. Patients with LUAD were divided into a low expression group and a high expression group according to the median value as a cut-off point of mRNA expression which were analyzed by KM-plotter survival curves and Log-rank test. All the LUAD data are analyzed for overall survival (OS) in the KM-plotter. In addition, we analyzed OS by selecting for two subgroups: gender and smoking history. The number of at-risk cases, HRs, 95% CIs, and *P*-values were displayed accordingly. A *p*-value <0.05 was considered statistically significant.

DriverDBv3 database analysis

DriverDBv3 is a cancer omics database which incorporates somatic mutations, RNA expression, miRNA expression, methylation, copy number variation and clinical data in addition to annotation bases. There are three functions, 'Cancer', 'Gene', and 'Customized-Analysis', to help researchers visualize the relationships between cancers and driver genes [26]. After combining different numbers of GINSs genes, we used DriverDBv3 to analyze the correlation between their expression levels in LUAD and their prognostic value. Survival-relevant analysis with a log-rank *P*-value <0.05 was considered significant. The sample numbers were also divided into a high-expression and a low-expression groups according to median values. HRs, 95% CIs, and *P*-values were displayed accordingly.

LinkedOmics analysis

The linkedOmics (http://www.linkedomics.org/) is an online analysis platform that provides multi-omics data of 32 TCGA cancers types for multi-dimensional analysis [27]. In our study, we identified co-expressed genes associated with GINSs in LUAD, by employing 'LinkedFiinder' module in linkedOmics. Those genes were analyzed through Pearson's correlation coefficient. Then, by using the 'LinkedCompare' module to identify the overlapping genes associated with GINSs in LUAD, we displayed some of the significantly positive and negative correlated genes in the heat map and histogram. In addition, we conducted biological analysis of GINSs and related genes by Gene Set Enrichment Analysis (GSEA) in the 'LinkedInterpreter' module, mainly including GO, KEGG pathways, kinase-targets, miRNA-targets and transcription factor-targets. Through weighted set coverage to reduce redundancy in the enrichment results. The rank criterion was an FDR <0.25, and 500 simulations were performed.

Results

Over-expression of GINSs in patients with LUAD

To begin our study, we first examined the expression levels of GINSs in different cancer types by using Oncomine database. We discovered that GINSs are over-expressed in most cancer types, except leukemia, melanoma, myeloma, and prostate cancer (**Figure 1A**). Especially when comparing LUAD tissues versus normal lung tissues, we found 14 datasets with high expression of GINSs in patients with LUAD (**Table 1**). GINS1 mRNA expression showed a 5.722-fold elevation from Hou lung (P=3.47e-18) [28], 3.861-fold elevation from Bhattacharjee Lung (P=5.33e-05) [29], 3.626-fold elevation from Su Lung (P=1.40e-10) [30], 7.817-fold elevation from Stearman Lung (P=1.01e-08) [31], 3.105-fold elevation from Okayama Lung (P=2.06e-18) [32], 2.295-fold elevation from Landi Lung (P=1.68e-16) [33], respectively. GINS2 mRNA expression showed a 9.927-fold elevation from Su lung (P=4.67e-12) [30], 3.601-fold elevation from Hou Lung (P=3.72e-17) [28], 2.508-fold elevation from Selamat Lung (P=3.80e-17) [34], 1.752-fold elevation from Landi Lung (P=8.94e-14) [33], 2.876-fold elevation from Okayama Lung (P=6.25e-13) [32]. Similarly, GINS3 mRNA expression showed 1.555-fold elevation from Hou lung (P=2.11e-09) [28] and GINS4 mRNA expression showed 2.56-fold elevation from Su Lung (P=2.19e-05) [30], 1.915-fold elevation from Hou Lung (P=6.45e-09) [28].

Next, we use Ualcan to explore the mRNA expression difference between the LUAD samples and normal samples (**Figure 1B**). mRNA expression of all four GINSs have statistically significant differences between tumor samples and normal samples (P<0.001).

Further, we tried to explore the protein expression patterns of GINSs in Human Protein Atlas. As is shown in **Figure 2**, GINS1 was lowly expressed in normal lung tissues, whereas a

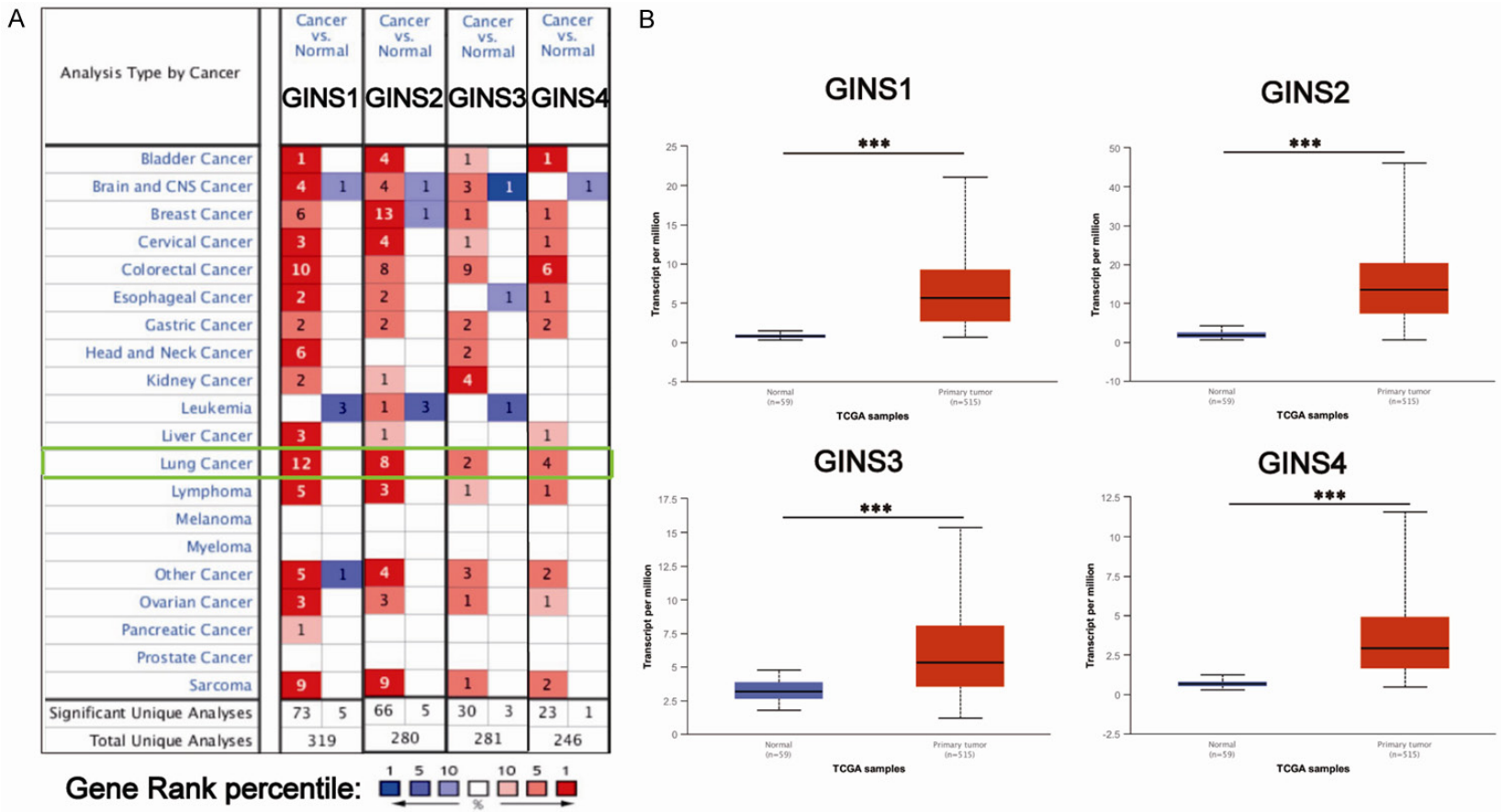


Figure 1. Transcriptional expression of GINSs in different types of cancer diseases (Oncomine database) and expression differences between the LUAD samples and normal lung samples (Ualcan). The graph demonstrated the numbers of datasets with statistically significant mRNA over-expression (red) or down-expression (blue) of the target genes. The number in each box signifies the number of analyses that meet the threshold within those analysis and cancer types. The gene rank was performed by means of percentile of target genes in the top of all genes measured in each research. The best gene rank percentile for the analyses within the cell was applied to determine the cell color. t-test was employed to compare the difference of transcriptional expression. Cut-off of *p*-value and fold change were listed as follows: *p*-value: 1e-4, fold change: 1.5, gene rank: 10%, data type: mRNA (A). mRNA expression of all four genes were of statistically significant difference between tumor samples and normal samples (B). GINSs: GINS complex subunits; LUAD: lung adenocarcinoma. ***P<0.001.

Table 1. Significant changes of GINSs expression in transcription level between LUAD and normal lung tissues (ONCOMINE)

GINSs	Types of LUAD vs Lung	Fold Change	t-test	*P*-value	Datasets
GINS1	Lung adenocarcinoma	5.722	11.85	3.47e-18	Hou Lung [30]
	Lung adenocarcinoma	3.861	4.781	5.33e-05	Bhattacharjee Lung [31]
	Lung adenocarcinoma	3.626	8.034	1.40e-10	Su Lung [32]
	Lung adenocarcinoma	7.817	7.126	1.01e-08	Stearman Lung [33]
	Lung adenocarcinoma	3.105	13.784	2.06e-18	Okayama Lung [34]
	Lung adenocarcinoma	2.295	10.462	1.68e-16	Landi Lung [35]
GINS2	Lung adenocarcinoma	9.927	8.795	4.67e-12	Su Lung [32]
	Lung adenocarcinoma	3.601	11.188	3.72e-17	Hou Lung [30]
	Lung adenocarcinoma	2.508	11.107	3.80e-17	Selamat Lung [36]
	Lung adenocarcinoma	1.752	8.9	8.94e-14	Landi Lung [35]
	Lung adenocarcinoma	2.876	10.83	6.25e-13	Okayama Lung [34]
GINS3	Lung adenocarcinoma	1.555	6.744	2.11e-09	Hou Lung [30]
GINS4	Lung adenocarcinoma	2.56	4.441	2.19e-05	Su Lung [32]
	Lung adenocarcinoma	1.915	6.76	6.45e-09	Hou Lung [30]

LUAD: lung adenocarcinoma; GINSs: GINS complex subunits.

medium expression was seen in LUAD tissues (**Figure 2A**). There was a low expression of GINS2 protein expression in normal lung tissues, yet a medium expression was observed in LUAD tissues (**Figure 2B**). The expression of GINS3 protein and GINS4 protein were both at an undetectable level in normal lung tissues, while the former was lowly expressed in LUAD tissues (**Figure 2C**) and the latter moderately expressed in LUAD tissues (**Figure 2D**).

Analysis of mRNA expression of GINSs in LUAD samples combine with clinicopathological parameters

After finding the over-expression of mRNA and protein, we further analyzed the mRNA expression of GINSs in LUAD tissue samples combine with clinicopathological parameters by Ualcan, including individual cancer stages, different genders and different smoking habits (**Figure 3**). In individual cancer stage groups (**Figure 3A-D**), mRNA expression of GINSs was expressed differently between normal stages and distinct tumor stages. Patients tended to express higher mRNA in more advanced stages and the highest mRNA expression of GINSs are all observed in stage 4. However, except for stage 2 and stage 3 of GINS3 gene, there was no statistically significant difference in expression between different cancer stages (**Figure 3C**). In the different gender groups (**Figure 3E-H**), GINSs mRNA expression levels were differentially expressed in both normal samples and tumor samples of different genders. In GINS1 and GINS2 genes, the mRNA expression levels between tumor samples from males and females showed statistical differences (**Figure 3E**, **3F**). In the different smoking habits groups (**Figure 3I-L**), similar to the first two groups, GINSs was highly expressed in tumor tissues compared with normal tissues. Among them, the expression of all the GINSs except GINS3 gene was higher in LUAD patients who smoked rather than non-smokers who never smoked (**Figure 3I**, **3J**, **3L**). In addition, the expression of GINSs was also related to smoking age. Compared with LUAD patients who have smoked for less than 15 years, all GINSs had higher expression in patients who have smoked for more than 15 years (**Figure 3I-L**). In addition to the GINS3 gene, there was no statistically significant difference in mRNA expression of the other three GINS genes in non-smokers and those who had smoked for less than 15 years (**Figure 3I**, **3J**, **3L**). Therefore, we compared the mRNA expression of these three genes in patients who had smoked for more than 15 years with that of non-smokers. Although the expression level of GINS4 was higher in patients who had smoked for more than 15 years, statistical differences were only found in the expression of GINS4 (**Figure 3L**).

Prognostic value of expression of different single GINS genes in LUAD patients

We next evaluated the prognostic value of mRNA expression of GINSs in LUAD patients by using KM-plotter. We evaluated the OS as-

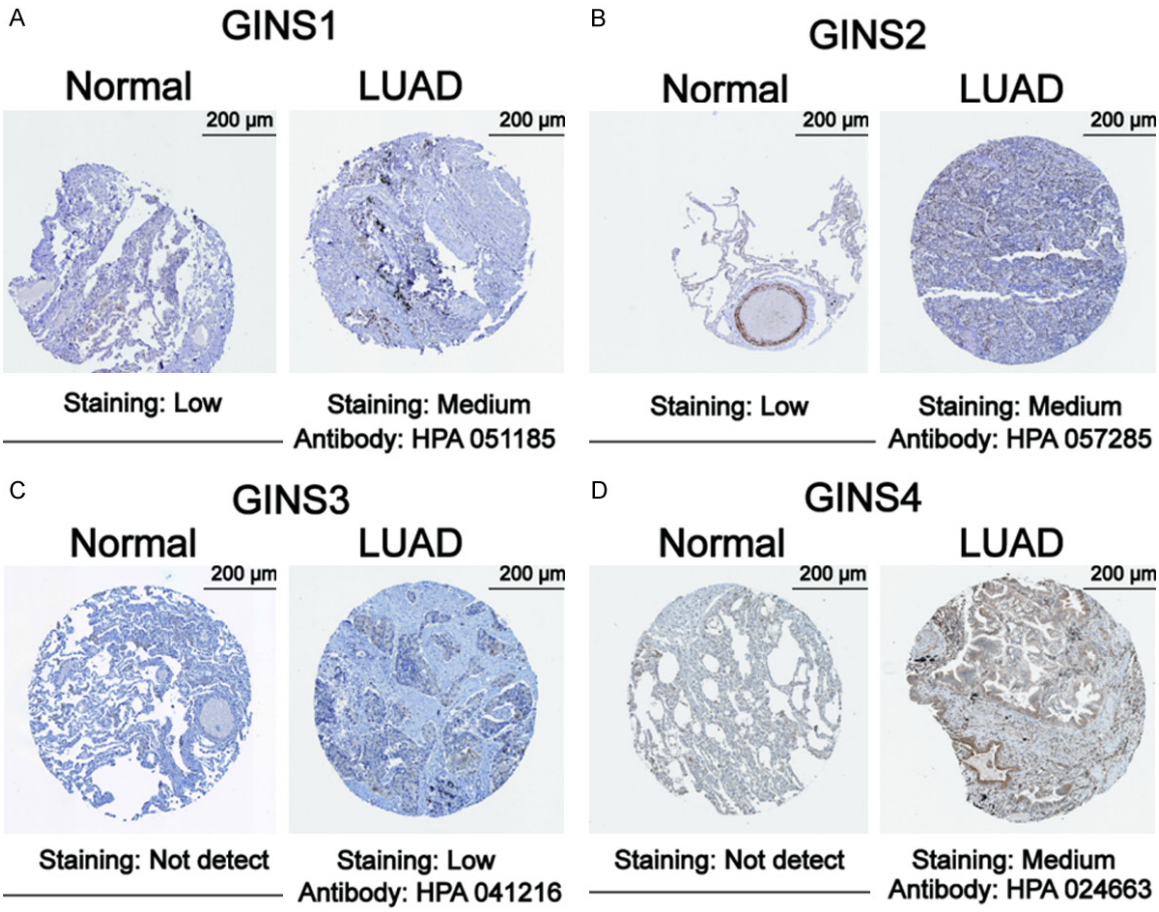


Figure 2. Representative immunohistochemistry images of distinct GINSs in normal lung tissues and LUAD tissues (Human Protein Atlas). Protein expression of GINS1 (A), GINS2 (B), GINS3 (C), and GINS4 (D) in normal lung tissues and LUAD tissues. Microscopic magnification of all samples is 200 μm. Different antibody types and staining intensities are indicated on the graph. GINSs: GINS complex subunits; LUAD: lung adenocarcinoma.

sociated with each mRNA expression of GINSs without restricting subtypes. As is presented in **Figure 4**, the higher mRNA expression of GINS1 (HR=1.37, P=0.0076), GINS2 (HR=1.93, P=5.9e-08), GINS3 (HR=1.31, P=0.023) and GINS4 (HR=1.86, P=3.2e-07) were significantly associated with poor OS in all patients with LUAD.

Then, we also correlated the mRNA expression of individual GINS genes in patient survival with subtypes sorted by gender or smoking history in LUAD. In the gender sub-cohort, higher mRNA expression of GINS1 (HR=1.58, P=0.0068), GINS2 (HR=1.66, P=0.0026) and GINS4 (HR=1.69, P=0.0019) were significantly associated with poor OS in male patients with LUAD (**Figure 5A**, **5B**, **5D**). High mRNA expression of GINS2 (HR=2.17, P=0.00012) and GINS4 (HR=2.58, P=3.1e-06) were significantly associated with poor OS in female patients with LUAD (**Figure 5G**, **5H**); however, GINS1 showed no statistically significant difference. No significant correlation was observed between the mRNA expression of GINS3 and OS of both genders. With respect to different smoking histories, high mRNA expression of GINS1 (HR=2.14, P=0.0017), GINS2 (HR=2.32, P=0.00056), GINS4 (HR=2.58, P=3.1e-06) was associated with poor OS in patients who smoked, except for GINS3 (**Figure 5I**, **5J**, **5L**). Whereas the mRNA expression of all four genes had no statistically significant difference with OS of patients who never smoked.

The co-expression of two or more GINS genes was associated with poor prognosis

Then, in order to explore the correlation between the co-expression of a different number

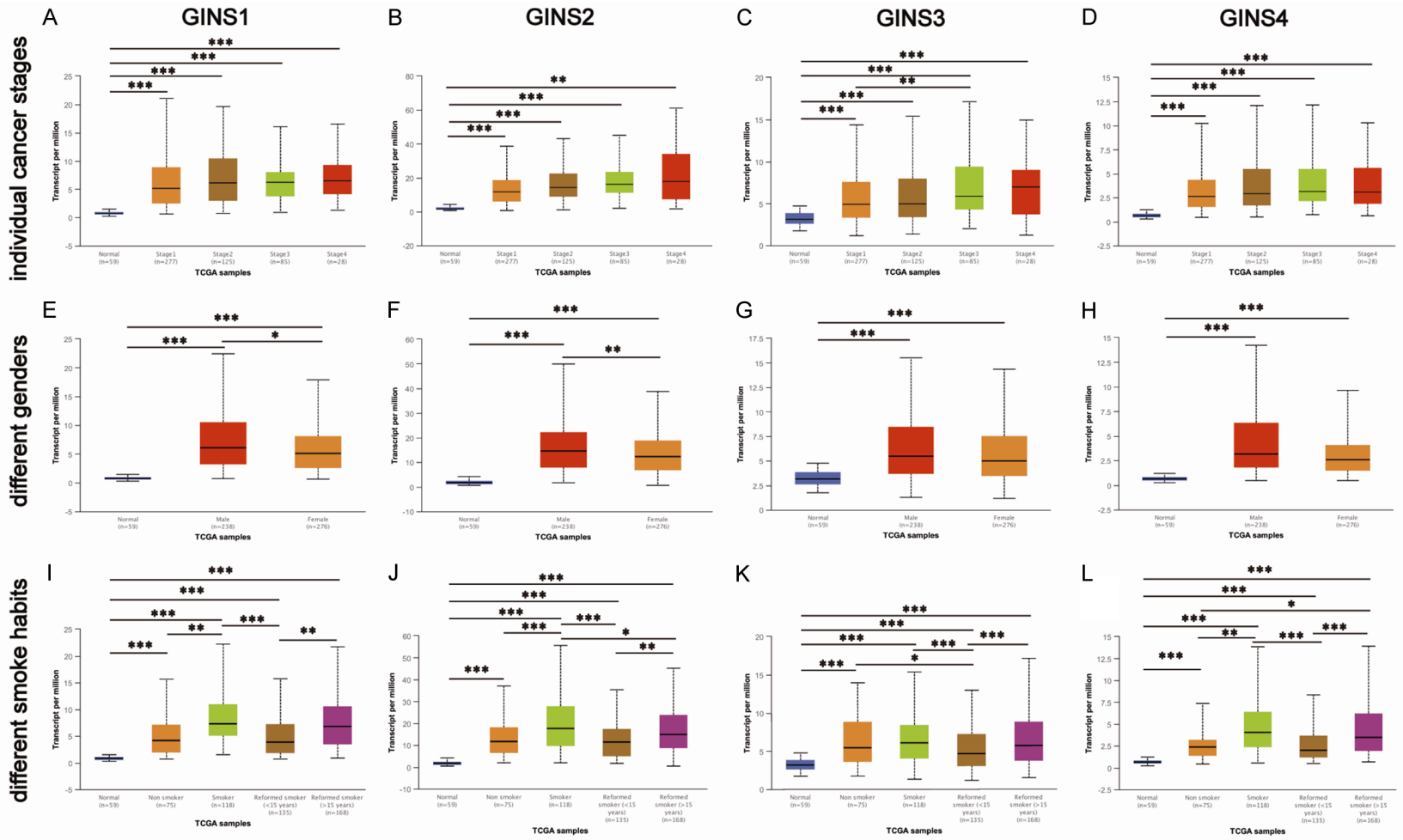


Figure 3. Analysis was performed on the mRNA expressions of different GINS family genes in LUAD samples in combination with clinicopathological parameters (Ualcan). Individual cancer stages groups (A-D), In different gender groups (E-H). Different smoking history groups (I-L). LUAD: lung adenocarcinoma. *P<0.05, **P<0.01, ***P<0.001.

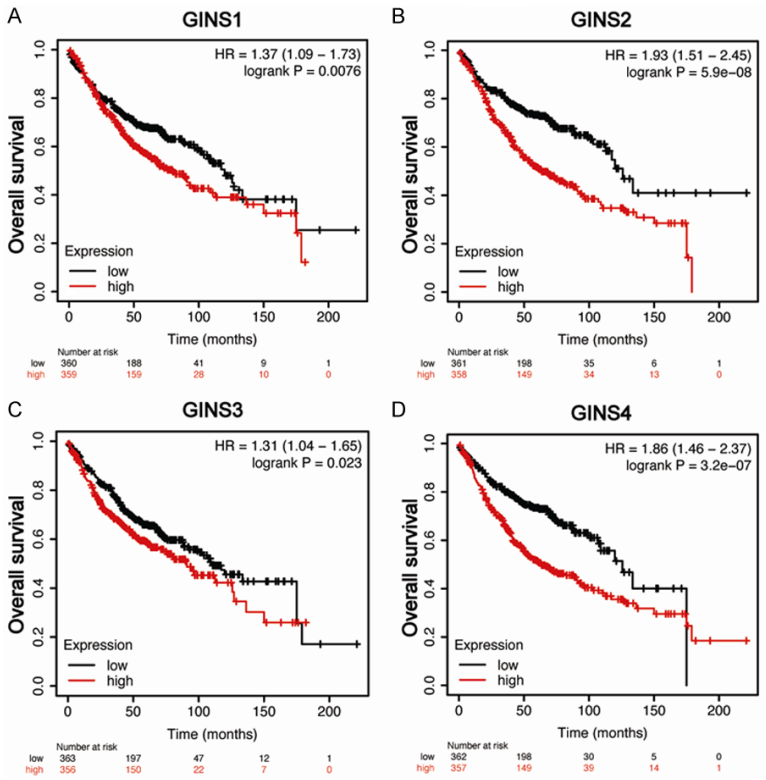


Figure 4. The associations between mRNA expression of each GINSs in tumor tissue and OS of LUAD patients (KM-plotter). mRNA expression of GINSs in tumor tissue was stratified into high or low expression using the median value as the cut-off point. Kaplan-Meier survival curves for (A) GINS1 (High expression, n=359; Low expression, n=360), (B) GINS2 (High expression, n=358; Low expression, n=361), (C) GINS3 (High expression, n=356; Low expression, n=363), and (D) GINS4 (High expression, n=357; Low expression, n=362), and the corresponding *P*-value for Log-rank test in all LUAD patients were showed. GINSs: GINS complex subunits; LUAD: lung adenocarcinoma.

of GINS genes and survival prognosis, we divided the different GINS gene combinations into 3 groups by number and analyzed them through DriverDBv3 database. In the two genes group (**Figure 6A-F**), the co-expression of genes in all subgroups was statistically significant, the GINS3-GINS4 subgroup had the most significant statistical significance (P= 0.00846), followed by the GINS1-GINS3 subgroup (P=0.0189), and then the GINS1-GINS2 subgroup (P=0.0256). The GINS1-GINS4 subgroup (P=0.0315), GINS2-GINS3 subgroup (P= 0.0367) and GINS2-GINS4 subgroup had similar statistical significance (P=0.0392). Among the three gene groups (**Figure 6G-I**), the subgroup of GINS1, GINS2 and GINS4 had the highest statistical difference (P=0.00944), followed by the subgroup of GINS1, GINS2 and GINS3 subgroups were not statistically significant. Finally, we found that the co-expression of four GINS genes in LUAD also suggested poor prognosis (P=0.0144) (**Figure 6J**).

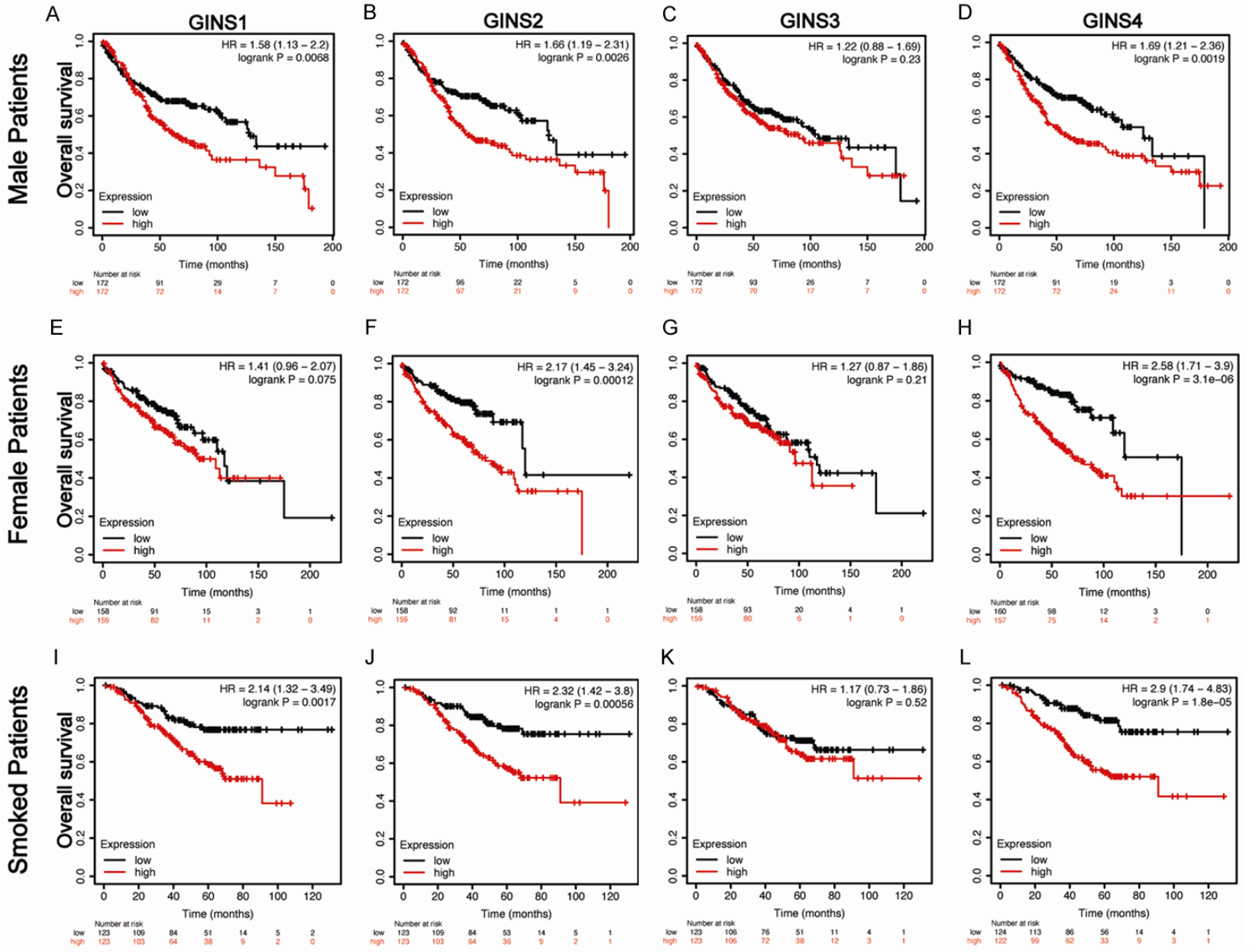
Screening of GINSs-related differentially expressed genes, and analysis of GO and KEGG pathway of GINSs and correlated genes in patients with LUAD

Afterward, by identifying GINSs-related overlapping genes in LUAD through LinkedOmics, we displayed some of these genes with a significant positive correlations and some with significant negative correlations in the histograms (**Figure 7A**) and heatmaps (**Figure 7B**), respectively. We also analyzed these GINSs-related overlapping genes for GO and KEGG pathways. The results show that the overlapped genes are mainly involved in the biological processes of chromosome separation, DNA replication, rRNA metabolic process, tRNA metabolic process, translational elongation and RNA splicing (**Figure 7C**). KEGG pathway analysis showed that these overlapped genes activate these pathways of cell cycles, spliceosomes, RNA transport, purine metabolisms, ribosomes and Parkinson disease (**Figure 7D**).

Enrichments of GINSs-related kinase targets, transcription factor targets and miRNA targets in LUAD

Due to the correlations between survival and the mRNA expression of GINSs, we tried to find some potential targets for further therapy. Therefore, we applied LinkedOmics to analyze GINSs-related kinase targets, transcription factor targets and miRNA targets in LUAD by GSEA.

As was shown in **Table 2**, the top 5 most significant kinase targets of correlated gene sets of


A
GINS1
HR = 1.58 (1.13 – 2.2)
logrank P = 0.0068
B
GINS2
HR = 1.66 (1.19 – 2.31)
logrank P = 0.0026
C
GINS3
HR = 1.22 (0.88 – 1.69)
logrank P = 0.23
D
GINS4
HR = 1.69 (1.21 – 2.36)
logrank P = 0.0019
Male Patients
E
HR = 1.41 (0.96 – 2.07)
logrank P = 0.075
F
HR = 2.17 (1.45 – 3.24)
logrank P = 0.00012
G
HR = 1.27 (0.87 – 1.86)
logrank P = 0.21
H
HR = 2.58 (1.71 – 3.9)
logrank P = 3.1e–06
Female Patients
I
HR = 2.14 (1.32 – 3.49)
logrank P = 0.0017
J
HR = 2.32 (1.42 – 3.8)
logrank P = 0.00056
K
HR = 1.17 (0.73 – 1.86)
logrank P = 0.52
L
HR = 2.9 (1.74 – 4.83)
logrank P = 1.8e–05
Smoked Patients
Overall survival
Expression
low
high
Time (months)
Number at risk

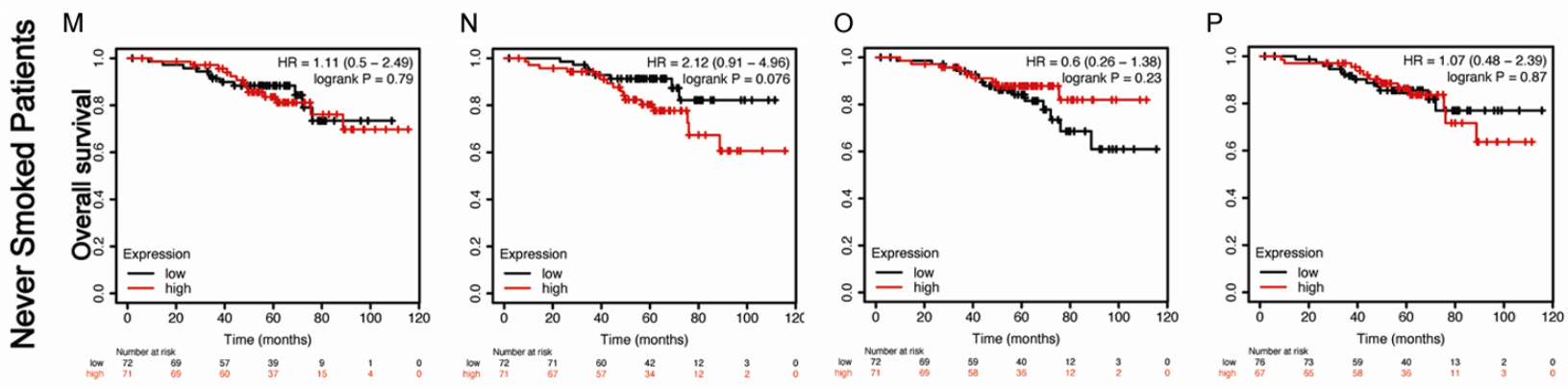


Figure 5. Two stratification factors are selected to analyze the associations between mRNA expression of GINSs and OS of LUAD patients (KM-plotter). According to gender, LUAD patients are divided into male patients (A-D) and female patients (E-H). Kaplan-Meier survival curves in male LUAD patients for (A) GINS1 (High expression, n=172; Low expression, n=172), (B) GINS2 (High expression, n=172; Low expression, n=172), (C) GINS3 (High expression, n=172; Low expression, n=172), and (D) GINS4 (High expression, n=172; Low expression, n=172), or in female LUAD patients for (E) GINS1 (High expression, n=159; Low expression, n=158), (F) GINS2 (High expression, n=159; Low expression, n=158), (G) GINS3 (High expression, n=159; Low expression, n=158), and (H) GINS4 (High expression, n=157; Low expression, n=160). Similarly, LUAD patients are divided into patients who smoked (I-L) and patients who never smoked (M-P). Kaplan-Meier survival curves in smoked LUAD patients for (I) GINS1 (High expression, n=123; Low expression, n=123), (J) GINS2 (High expression, n=123; Low expression, n=123), (K) GINS3 (High expression, n=123; Low expression, n=123), and (L) GINS4 (High expression, n=122; Low expression, n=124), or in LUAD patients who never smoked for (M) GINS1 (High expression, n=71; Low expression, n=72), (N) GINS2 (High expression, n=71; Low expression, n=72), (O) GINS3 (High expression, n=71; Low expression, n=72), and (P) GINS4 (High expression, n=67; Low expression, n=76). The corresponding *P*-value for Log-rank test were showed. GINSs: GINS complex subunits; LUAD: lung adenocarcinoma.

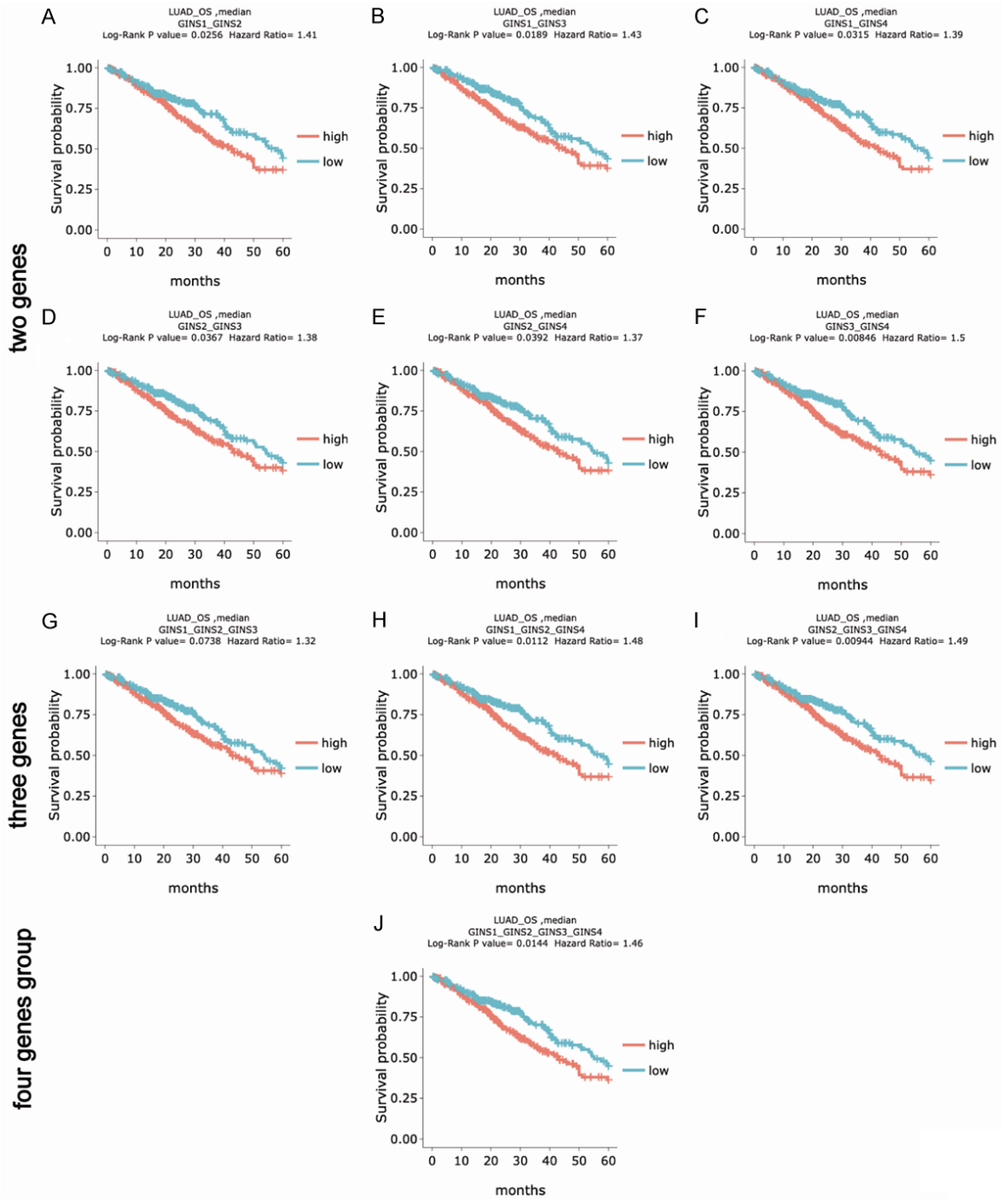
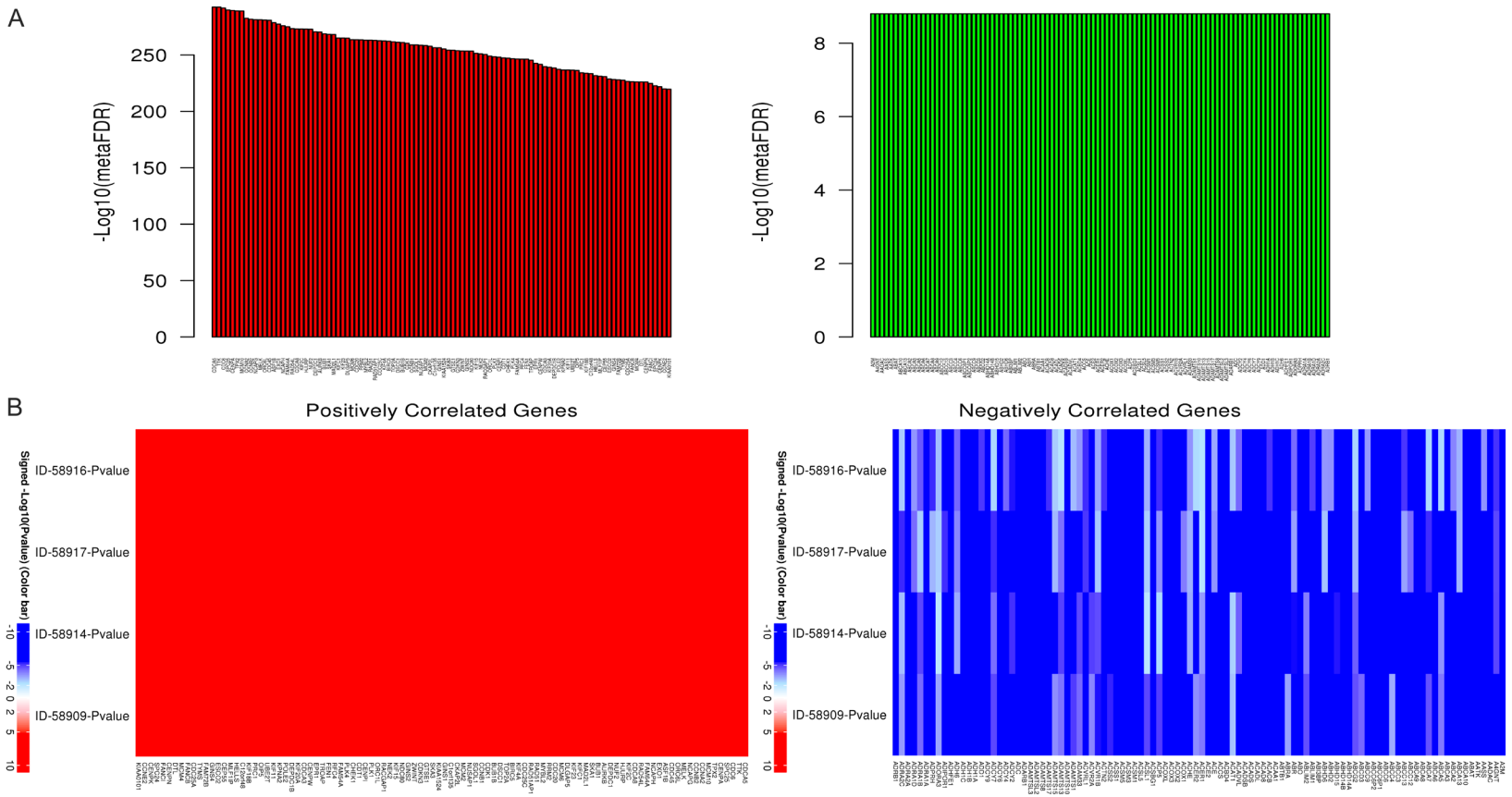


Figure 6. Association between co-expression of two or more genes and prognosis (DriverDBv3 database). There are two genes group (A-F), three genes group (G-I) and four genes group (J). The number of high-expression and low-expression patients in each subgroup was 250. The corresponding *P*-values for Log-rank test were showed in figure. LUAD: lung adenocarcinoma.

GINSs were related mainly to PLK1 (polo like kinase 1), CDK1 (cyclin dependent kinase 1), AURKB (aurora kinase B), CDK2 (cyclin dependent kinase 2) and CHEK2 (checkpoint kinase 2). We also found that transcriptional levels of these kinase genes, with the exception of CDK2, were significantly elevated in LUAD tissues. Also, the expression levels of all kinase genes were significantly associated with the OS in LUAD patient, except CDK1 (Supplementary


A
-Log10(metaFDR)
250
200
150
100
50
0
-Log10(metaFDR)
8
6
4
2
0
B
Positively Correlated Genes
Negatively Correlated Genes
Signed -Log10(Pvalue) (Color bar)
ID-58916-Pvalue
ID-58917-Pvalue
ID-58914-Pvalue
ID-58909-Pvalue
-10
-5
-2
0
2
5
10

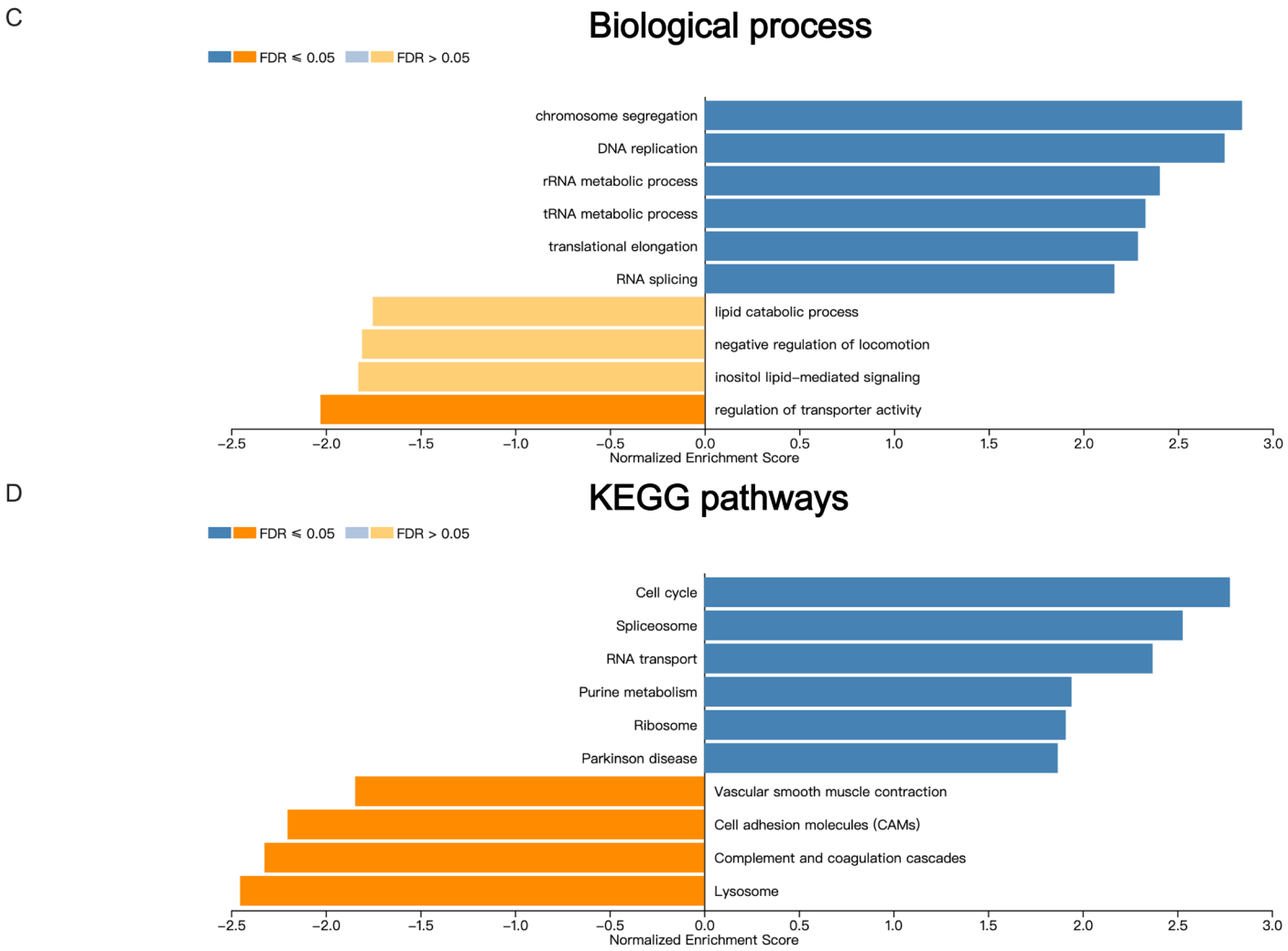


Figure 7. The differentially expressed genes were screened and analyzed by GO and KEGG (LinkedOmics). The most significantly different positive and negative correlated genes were plotted in red and green bars (A). Heatmaps of some correlated differentially expressed genes. Red represents positive correlation. Blue represents negative correlation (B). Biological process analysis and KEGG pathway analysis of GINSs and correlated genes in LUAD patients (C, D). GINSs: GINS complex subunits; LUAD: lung adenocarcinoma.

Table 2. The kinase targets, transcription factor targets, and miRNA targets of GINSs in LUAD (LinkedOmics)

Enriched category	Gene set	LEN	FDR
Kinase target	Kinase_PLK1 (polo like kinase 1)	90	0.00e+00
	Kinase_CDK1 (cyclin dependent kinase 1)	258	0.00e+00
	Kinase_AURKB (aurora kinase B)	87	0.00e+00
	Kinase_CDK2 (cyclin dependent kinase 2)	278	0.00e+00
	Kinase_CHEK2 (checkpoint kinase 2)	27	0.00e+00
Transcription factor target	V$E2F1_Q6	213	0.00e+00
	SGCGSSAAA_V$E2F1DP2_01	155	0.00e+00
	V$E2F_Q6	211	0.00e+00
	V$E2F4DP1_01	220	0.00e+00
	V$E2F_Q4	212	0.00e+00

LUAD: lung adenocarcinoma; GINSs: GINS complex subunits; LEN: Leading Edge Number.

Figure 1). The enrichment of transcription factor targets for GINSs and their related genes were E2F1_Q6, E2F1DP2_01, E2F_Q6, E2F4DP1_01 and E2F_Q4. It was reported that E2F family members with an abnormal transcriptional expression could serve as potential targets in LUAD [35]. We fond that overexpression of E2F1 and E2F4 indicated poor prognosis in patients with LUAD (Supplementary Figure 1). No miRNA targets were found to be significantly correlated.

Discussion

Although the current treatment methods for LUAD have been continuously improved, the actual survival rate is still not satisfactory. As core genes in DNA replication, abnormal expression of GINSs may lead to the dysfunction of the DNA helicase and further tumorigenesis [10]. In order to clarify the prognostic value of GINSs in LUAD, our study explored the relationship between the expression and survival of different GINSs in LUAD through online databases.

In the present study, we demonstrated that all GINS genes were highly expressed in LUAD and found that GINS1/2/4 was more highly expressed in LUAD patients with a history of smoking. Moreover, over-expression of GINSs was significantly correlated with poor prognosis, in which patients who smoked had high expression of GINS1/2/3 and poor prognosis. Actually, smoking has been shown to promote the development of lung cancer, and the incidence of lung cancer soared with the increase of smoking rate [36]. Therefore, smoking may promote the expression of GINSs and thus affect the prognosis of patients with LUAD. We also attempted to analyze the correlation between different GINS gene expressions grouped by gender and prognosis. The results showed that gender did not appear to be a factor affecting the correlation between GINS expression and prognosis. Next, we found that the co-expression of two or more genes were associated with poor prognosis, except for the expression of the GINS1, GINS2 and GINS3 gene subgroups that was not associated with prognosis. Enrichment of function and pathway in GINSs and their correlated genes in LUAD patients were also analyzed. Biological processes such as chromosome separation and pathways such as cell cycle were remarkably regulated by the GINSs mutations in LUAD. GINSs-related kinase targets (PLK1, CDK1, AURKB, CDK2 and CHEK2) and transcription factor targets (E2Fs) that we had explored can serve as the future research direction in LUAD.

In GINSs, previous studies respectively demonstrated that downregulation of GINS1 expression could inhibit cell proliferation and induced cell cycle arrest in lung cancer cells, high expression of GINS1 prompted poor prognosis in patients with NSCLC which were treated with surgery following by preoperative chemotherapy or chemoradiotherapy [37, 38]. In our study, we found that the mRNA expression of GINS1 increased with the development of cancer stage and high expression of GINS1 suggested poor prognosis. Interestingly, high mRNA expression of GINS1 showed poor prognosis in male patients and patients who smoked, respectively. As we all know, a large proportion of men regard smoking as their hobby, and smoking can promote the progression of lung can-

cer. Therefore, we have reason to believe that one of the reasons for the poor prognosis of male patients is that these male patients have a history of smoking.

GINS2 was discovered to enhance migration, invasion viability, and epithelial-mesenchymal transition (EMT) in NSCLC cells, which were dominantly mediated by PI3K/Akt and MEK/ERK signaling pathways [39]. Recently, a study showed that down-regulated expression of GINS2 can inhabit the cell proliferation and increase apoptosis through the p53/GADD45A pathway in NSCLC [40]. According to our results, the expression of GINS2 augmented the increase of cancer stage. We observed that high expression of GINS2 in patients with LUAD with smoking habits had apoor prognosis, so we hypothesed that smoking promoted the expression of the GINS2 gene, while the high expression of GINS2 reduced the survival of patients with LUAD. Among them, there was a statistical difference in GINS2 mRNA expression between male and female patients, and the high expression of GINS2 in both genders indicated a poor prognosis, which might suggest that gender had no decisive effect on GINS2 expression. In short, for smokers, smoking can promote the expression of GINS2 and have an impact on clinical outcomes.

In LUAD, the relationship between the expression of GINS3 and its prognosis had been studied most. High expression of GINS3 demonstrated a poor prognosis and may need additional treatment in LUAD, especially in the early stages of LUAD [19-21]. Nevertheless, its role as a prognostic biomarker is not as strong as previously suggested. We observed that mRNA expression of GINS3 was associated with OS in LUAD. However, in the sub cohort analysis concerning gender and smoking history, no statistically significant differences were observed. Therefore, we considered that smoking and gender are not influential factors in mRNA expression of GINS3 in LUAD. In combination with previous studies, we assumed that the statistical difference between GINS3 expression and survival was mainly caused by early stage, but according to this study, the mRNA expression of GINS3 in early stage LUAD was not high. Therefore, more studies are needed to verify the relationship between GINS3 expression and OS.

A study showed that GINS4 expression may be related to many procedures during tumorigenesis in NSCLC, including cell growth, epithelial-mesenchymal transition, clonal formation, migration and invasion [41], all of which lead to poor prognosis. Similar to the GINS2 gene analysis result, high mRNA expression of GINS4 in patients with LUAD who had asmoking history have an indicated poor OS. It is possible to hypothesize that GINS2 and GINS4 genes may be similar in structure and function, and that they are more susceptible to other adverse factors to promote tumor growth and development, but more research is needed to prove these hypotheses.

GINSs needed to be transcribed and assembled to form the GINS complex, which bind to other key components to form the CMG complex. Therefore, GINSs dod not necessarily play a role alone. We explored the correlation between the co-expression of two or more GINS genes and the prognosis [7]. We found that almost all subgroups showed statistical differences, which suggested that the combinatory use of GINS expression levels could be of clinical significance for prognosis prediction of patients with LUAD.

There are several limitations in our investigation. Firstly, our study of prognostic prediction was mainly based on the LUAD patient cohort from different online databases without preclinical or clinical research to verify our results. Secondly, our study only carried out single factor analyses due to an insufficient sample size. Hence, researchers in the future need to use a larger sample size to verify our findings through more basic and clinical trials, and to explore the clinical application of GINSs members in the treatment of LUAD.

In conclusion, the present research demonstrated that all GINS genes are highly expressed in LUAD and associated with poor survival of patients with LUAD, especially those who have a smoking history. The combinatory use of expression of GINSs could provide a good prognostic prediction for patients with LUAD.

Disclosure of conflict of interest

None.

Address correspondence to: Hui Luo, Department of Oncology, The First Affiliated Hospital of Nanchang University, Nanchang University, Nanchang 330006, Jiangxi, P. R. China. E-mail: 339749308@qq.com

References

[1] Barta JA, Powell CA and Wisnivesky JP. Global epidemiology of lung cancer. Ann Glob Health 2019; 85: 1-16.

[2] Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, Allen C, Hansen G, Woodbrook R, Wolfe C, Hamadeh RR, Moore A, Werdecker A, Gessner BD, Te Ao B, McMahon B, Karimkhani C, Yu C, Cooke GS, Schwebel DC, Carpenter DO, Pereira DM, Nash D, Kazi DS, De Leo D, Plass D, Ukwaja KN, Thurston GD, Yun Jin K, Simard EP, Mills E, Park EK, Catalá-López F, deVeber G, Gotay C, Khan G, Hosgood HD 3rd, Santos IS, Leasher JL, Singh J, Leigh J, Jonas JB, Sanabria J, Beardsley J, Jacobsen KH, Takahashi K, Franklin RC, Ronfani L, Montico M, Naldi L, Tonelli M, Geleijnse J, Petzold M, Shrime MG, Younis M, Yonemoto N, Breitborde N, Yip P, Pourmalek F, Lotufo PA, Esteghamati A, Hankey GJ, Ali R, Lunevicius R, Malekzadeh R, Dellavalle R, Weintraub R, Lucas R, Hay R, Rojas-Rueda D, Westerman R, Sepanlou SG, Nolte S, Patten S, Weichenthal S, Abera SF, Fereshtehnejad SM, Shiue I, Driscoll T, Vasankari T, Alsharif U, Rahimi-Movaghar V, Vlassov VV, Marcenes WS, Mekonnen W, Melaku YA, Yano Y, Artaman A, Campos I, MacLachlan J, Mueller U, Kim D, Trillini M, Eshrati B, Williams HC, Shibuya K, Dandona R, Murthy K, Cowie B, Amare AT, Antonio CA, Castañeda-Orjuela C, van Gool CH, Violante F, Oh IH, Deribe K, Soreide K, Knibbs L, Kereselidze M, Green M, Cardenas R, Roy N, Tillmann T, Li Y, Krueger H, Monasta L, Dey S, Sheikhbahaei S, Hafezi-Nejad N, Kumar GA, Sreeramareddy CT, Dandona L, Wang H, Vollset SE, Mokdad A, Salomon JA, Lozano R, Vos T, Forouzanfar M, Lopez A, Murray C and Naghavi M. The global burden of cancer 2013. JAMA Oncol 2015; 1: 505-527.

[3] Minna J, Roth J and Gazdar A. Focus on lung cancer. Cancer Cell 2002; 1: 49-52.

[4] Lortet-Tieulent J, Soerjomataram I, Ferlay J, Rutherford M, Weiderpass E and Bray F. International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women. Lung Cancer 2014; 84: 13-22.

[5] Calvayrac O, Pradines A, Pons E, Mazieres J and Guibert N. Molecular biomarkers for lung adenocarcinoma. Eur Respir J 2017; 49: 1601734.

[6] Sundar R, Soong R, Cho BC, Brahmer JR and Soo RA. Immunotherapy in the treatment of non-small cell lung cancer. Lung Cancer 2014; 85: 101-109.

[7] Kamada K. The GINS complex: structure and function. Subcell Biochem 2012; 62: 135-156.

[8] Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H and Takisawa H. A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. Genes Dev 2003; 17: 1141-1152.

[9] Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A and Araki H. GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. Genes Dev 2003; 17: 1153-1165.

[10] Bochman ML. Roles of DNA helicases in the maintenance of genome integrity. Mol Cell Oncol 2014; 1: e963429.

[11] Zhou L, Sun XJ, Liu C, Wu QF, Tai MH, Wei JC, Lei L, Meng FD, Qu K and Xu J. Overexpression of PSF1 is correlated with poor prognosis in hepatocellular carcinoma patients. Int J Biol Markers 2015; 30: e56-64.

[12] Nakahara I, Miyamoto M, Shibata T, Akashi-Tanaka S, Kinoshita T, Mogushi K, Oda K, Ueno M, Takakura N, Mizushima H, Tanaka H and Ohta T. Up-regulation of PSF1 promotes the growth of breast cancer cells. Genes Cells 2010; 15: 1015-1024.

[13] Tahara H, Naito H, Kise K, Wakabayashi T, Kamoi K, Okihara K, Yanagisawa A, Nakai Y, Nonomura N, Morii E, Miki T and Takakura N. Evaluation of PSF1 as a prognostic biomarker for prostate cancer. Prostate Cancer Prostatic Dis 2015; 18: 56-62.

[14] Ouyang F, Liu J, Xia M, Lin C, Wu X, Ye L, Song L, Li J, Wang J, Guo P and He M. GINS2 is a novel prognostic biomarker and promotes tumor progression in early-stage cervical cancer. Oncol Rep 2017; 37: 2652-2662.

[15] Sun X, Sui W, Huang M, Wang Y, Xuan Y and Wang Z. Partner of Sld five 3: a potential prognostic biomarker for colorectal cancer. Diagn Pathol 2014; 9: 217.

[16] Park JK, Otsuka N, Tomaru U, Suzuki H, Azuma M, Okamoto K, Yamashiro K and Kasahara M. Clinicopathological significance of PSF3 expression in uterine endometrial carcinomas. Hum Pathol 2018; 80: 104-112.

[17] Rong Z, Luo Z, Zhang J, Li T, Zhu Z, Yu Z, Fu Z, Qiu Z and Huang C. GINS complex subunit 4, a prognostic biomarker and reversely mediated by Kruppel-like factor 4, promotes the growth of colorectal cancer. Cancer Sci 2020; 111: 1203-1217.

[18] Zhu Z, Yu Z, Rong Z, Luo Z, Zhang J, Qiu Z and Huang C. The novel GINS4 axis promotes gas-

tric cancer growth and progression by activating Rac1 and CDC42. Theranostics 2019; 9: 8294-8311.
[19] Kimura K, Tanaka Y, Tauchi S, Kitamura Y, Nishio W, Sakai Y, Hayashi Y, Yoshimura M and Maniwa Y. Psf3 as a possible biomarker of postoperative chemotherapy for patients with early pulmonary adenocarcinoma. Thorac Cancer 2019; 10: 2300-2307.
[20] Hokka D, Maniwa Y, Tane S, Nishio W, Yoshimura M, Okita Y, Ohbayashi C, Sakai Y, Chen X and Hayashi Y. Psf3 is a prognostic biomarker in lung adenocarcinoma. Lung Cancer 2013; 79: 77-82.
[21] Tauchi S, Sakai Y, Fujimoto S, Ogawa H, Tane S, Hokka D, Tanaka Y, Nishio W, Yoshimura M, Yanagita E, Itoh T, Hayashi Y and Maniwa Y. Psf3 is a prognostic biomarker in lung adenocarcinoma: a larger trial using tissue microarrays of 864 consecutive resections. Eur J Cardiothorac Surg 2016; 50: 758-764.
[22] Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pander A and Chinnaiyan AM. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 2004; 6: 1-6.
[23] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[24] Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnström H, Glimelius B, Sjöblom T, Edqvist PH, Djureinovic D, Micke P, Lindskog C, Mardinoglu A and Ponten F. A pathology atlas of the human cancer transcriptome. Science 2017; 357: eaan2507.
[25] Chellappan SP, Győrffy B, Surowiak P, Budczies J and Lánczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. PLoS One 2013; 8: e82241.
[26] Liu SH, Shen PC, Chen CY, Hsu AN, Cho YC, Lai YL, Chen FH, Li CY, Wang SC, Chen M, Chung IF and Cheng WC. DriverDBv3: a multi-omics database for cancer driver gene research. Nucleic Acids Res 2020; 48: D863-D870.
[27] Vasaikar SV, Straub P, Wang J and Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res 2018; 46: D956-D963.
[28] Hou J, Aerts J, den Hamer B, van Ijcken W, den Bakker M, Riegman P, van der Leest C, van der Spek P, Foekens JA, Hoogsteden HC, Grosveld F and Philipsen S. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. PLoS One 2010; 5: e10312.
[29] Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M, Loda M, Weber G, Mark EJ, Lander ES, Wong W, Johnson BE, Golub TR, Sugarbaker DJ and Meyerson M. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 2001; 98: 13790-13795.
[30] Su LJ, Chang CW, Wu YC, Chen KC, Lin CJ, Liang SC, Lin CH, Whang-Peng J, Hsu SL, Chen CH and Huang CY. Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme. BMC Genomics 2007; 8: 140.
[31] Stearman R, Dwyer-Nield L, Zerbe L, Blaine S, Chan Z, Bunn PJ, Johnson G, Hirsch F, Merrick D, Franklin W, Baron A, Keith R, Nemenoff R, Malkinson A and Geraci M. Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. Am J Pathol 2005; 167: 1763-1775.
[32] Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, Furuta K, Tsuta K, Shibata T, Yamamoto S, Watanabe S, Sakamoto H, Kumamoto K, Takenoshita S, Gotoh N, Mizuno H, Sarai A, Kawano S, Yamaguchi R, Miyano S and Yokota J. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. Cancer Res 2012; 72: 100-111.
[33] Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, Dasgupta A, Mann FE, Fukuoka J, Hames M, Bergen AW, Murphy SE, Yang P, Pesatori AC, Consonni D, Bertazzi PA, Wacholder S, Shih JH, Caporaso NE and Jen J. Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. PLoS One 2008; 3: e1651.
[34] Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, Siegmund KD, Koss MN, Hagen JA, Lam WL, Lam S, Gazdar AF and Laird-Offringa IA. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 2012; 22: 1197-1211.
[35] Sun C, Zhou Q, Hu W, Li S, Zhang F, Chen Z, Li G, Bi Z, Bi Y, Gong F, Bo T, Yuan Z, Hu W, Zhan B, Zhang Q, Tang Q and Li D. Transcriptional E2F1/2/5/8 as potential targets and transcriptional E2F3/6/7 as new biomarkers for the prognosis of human lung carcinoma. Aging (Albany NY) 2018; 10: 973-987.
[36] Mao Y, Yang D, He J and Krasna MJ. Epidemiology of lung cancer. Surg Oncol Clin N Am 2016; 25: 439-445.

[37] Zhang J, Wu Q, Wang Z, Zhang Y, Zhang G, Fu J and Liu C. Knockdown of PSF1 expression inhibits cell proliferation in lung cancer cells in vitro. Tumour Biol 2015; 36: 2163-2168.

[38] Kanzaki R, Naito H, Kise K, Takara K, Eino D, Minami M, Shintani Y, Funaki S, Kawamura T, Kimura T, Okumura M and Takakura N. PSF1 (Partner of SLD Five 1) is a prognostic biomarker in patients with non-small cell lung cancer treated with surgery following preoperative chemotherapy or chemoradiotherapy. Ann Surg Oncol 2016; 23: 4093-4100.

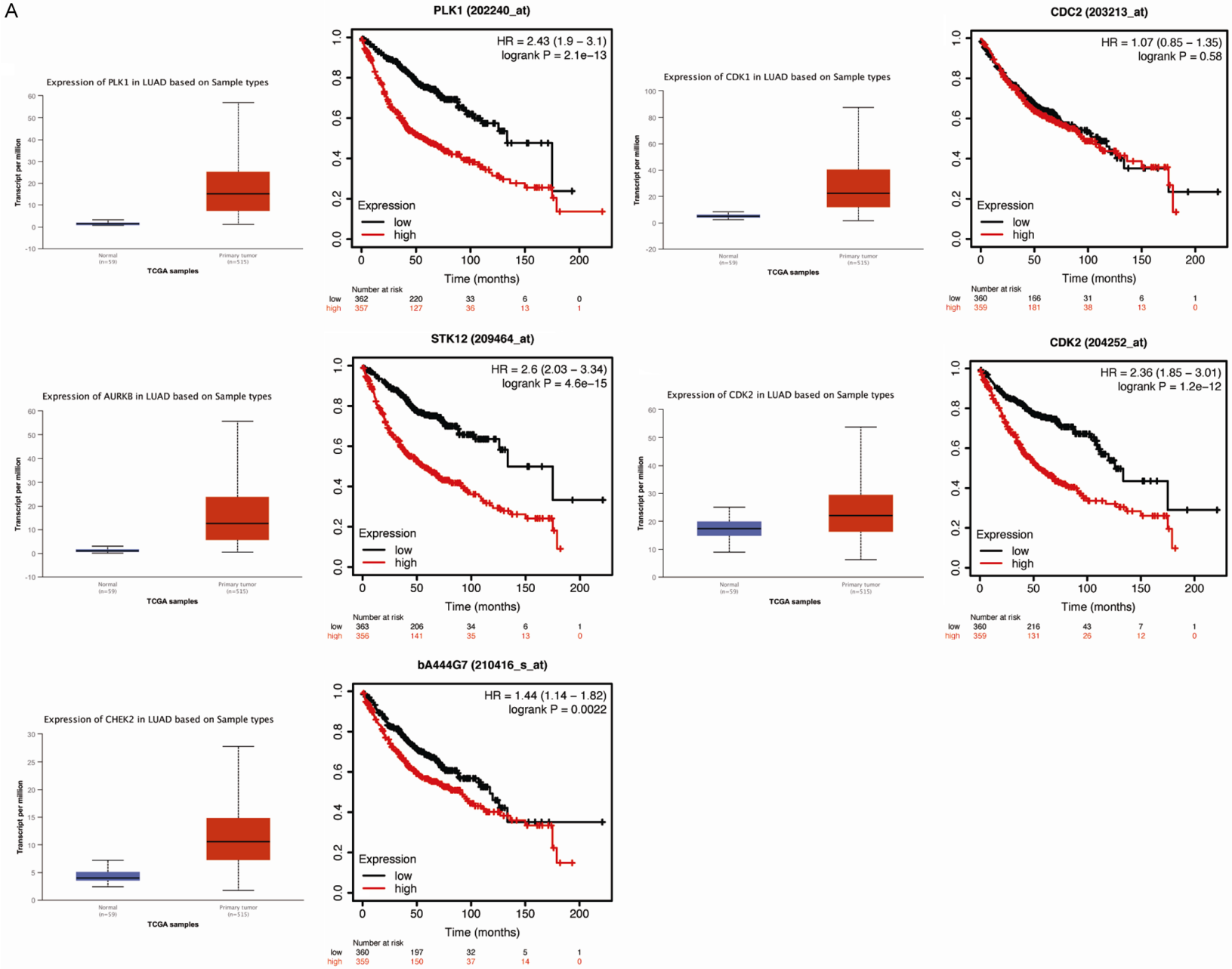
[39] Liu X, Sun L, Zhang S, Zhang S and Li W. GINS2 facilitates epithelial-to-mesenchymal transition in non-small-cell lung cancer through modulating PI3K/Akt and MEK/ERK signaling. J Cell Physiol 2019; 234: 1-10.

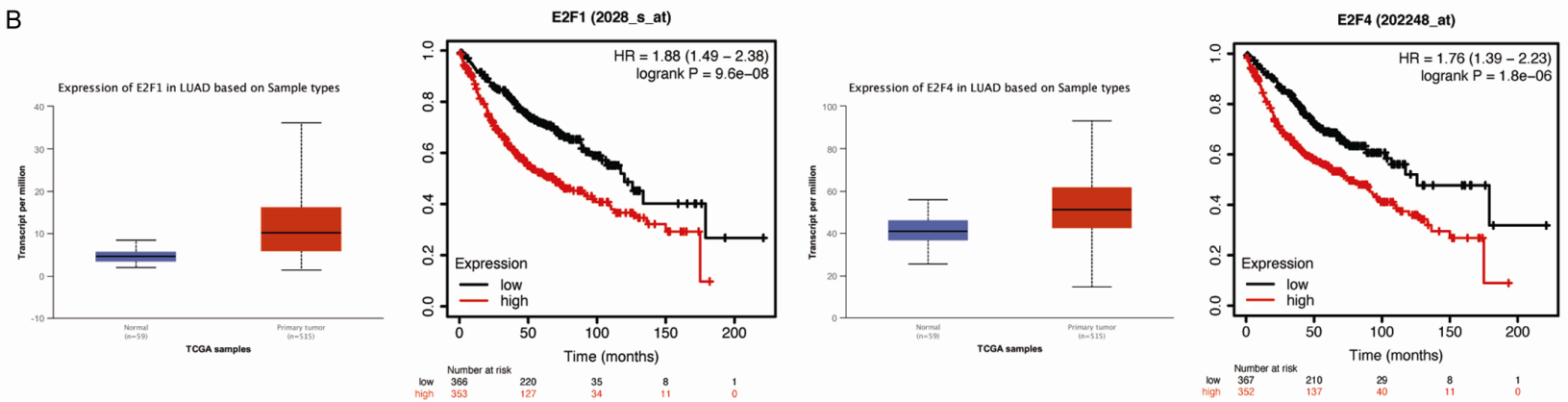
[40] Chi F, Wang Z, Li Y and Chang N. Knockdown of GINS2 inhibits proliferation and promotes apoptosis through the p53/GADD45A pathway in non-small-cell lung cancer. Biosci Rep 2020; 40: BSR20193949.

[41] Yang R, Liu N, Chen L, Jiang Y, Shi Y, Mao C, Liu Y, Wang M, Lai W, Tang H, Gao M, Xiao D, Wang X, Yu F, Cao Y, Yan Q, Liu S and Tao Y. LSH interacts with and stabilizes GINS4 transcript that promotes tumourigenesis in non-small cell lung cancer. J Exp Clin Cancer Res 2019; 38: 280.

GINSs could be potential biomarkers for LUAD

A





Supplementary Figure 1. Expression and survival outcome in GINSs-related kinase and transcription factor targets (UALCAN and KM-plotter). Expression and survival outcome of top 5 GINSs-related kinases targets in LUAD patients (A). E2F regulators of GINSs co-expressed genes (B). Aliases: CDC2 (CDK1); STK12 (AURKB); bA444G7 (CHEK2).