Am J Cancer Res 2022;12(3):986-1008
www.ajcr.us /ISSN:2156-6976/ajcr0141269

Original Article

A pan-cancer analysis of GINS complex subunit 4 to identify its potential role as a biomarker in multiple human cancers

Muhammad Usman[1], Mohammad K Okla[2], Hafiz Muhammad Asif[3], Gehad AbdElgayed[4], Fatima Muccee[5], Shakira Ghazanfar[6], Mukhtiar Ahmad[1], Muhammad Junaid Iqbal[7], Aamina Murad Sahar[8], Ghania Khaliq[9], Rabbia Shoaib[10], Hira Zaheer[1], Yasir Hameed[1]

[1]Department of Biochemistry and Biotechnology, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Pakistan; [2]Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia; [3]University College of Conventional Medicine, Faculty of Pharmacy and Alternative Medicine, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan; [4]Integrated Molecular Plant Physiology Research, Department of Biology, University of Antwerp, 2020 Antwerp, Belgium; [5]Department of Biotechnology, Virtual University of Pakistan, Lahore 54000, Pakistan; [6]Functional Genomics and Bioinformatics, National Agricultural Research Centre, Islamabad 45500, Pakistan; [7]Department of Biosciences, COMSATS University, Islamabad 44000, Pakistan; [8]Department of Biosciences, COMSATS University Islamabad, Islamabad 4400, Pakistan; [9]Department of Zoology, Cholistan University of Veterinary and Animal Sciences Bahawalpur, Bahawalpur 63100, Pakistan; [10]Department of Chemistry, Government College University Faisalabad, Faisalabad 3800, Pakistan

Received December 22, 2021; Accepted February 16, 2022; Epub March 15, 2022; Published March 30, 2022

Abstract: This study was initiated to explore the expression variation, clinical significance, and biological importance of the GINS complex subunit 4 (GINS4) in different human cancers as a shared biomarker via pan-cancer analysis through different platforms including UALCAN, Kaplan Meier (KM) plotter, TNMplot, GENT2, GEPIA, DriverDBv3, Human Protein Atlas (HPA), MEXPRESS, cBioportal, STRING, DAVID, MuTarge, Enrichr, TIMER, and CTD. Our findings have verified the up-regulation of GINS4 in 24 major subtypes of human cancers, and its overexpression was found to be substantially associated with poor overall survival (OS), relapse-free survival (RFs), and metastasis in ESCA, KIRC, LIHC, LUAD, and UCEC. This suggested that GINS4 plays a significant role in the development and progression of these five cancers. Furthermore, we noticed that GINS4 is also overexpressed in ESCA, KIRC, LIHC, LUAD, and UCEC patients with different clinicopathological characteristics. Enrichment analysis revealed the involvement of GINS4 associated genes in a variety of diverse GO and KEGG terms. We also explored few significant correlations between GINS4 expression and promoter methylation, genetic alterations, CNVs, other mutant genes, tumor purity, and immune cells infiltration. In conclusion, our results elucidated that GINS4 can serve as a shared diagnostic, prognostic biomarker, and a potential therapeutic target in ESCA, KIRC, LIHC, LUAD, and UCEC patients with different clinicopathological characteristics.

Keywords: GINS4, cancer, expression variations, biomarker, tumor purity

Introduction

Cancer is one of the major health threats worldwide and is triggered by several factors, including viral infections, previous history of cancer development, excessive alcohol intake, lack of physical activity, autoimmune, and metabolic disorders [1, 2]. According to recent reports, the overall global burden of cancer has risen to 19.3 million new cases, and 10 million deaths in 2020 [3], relative to 18.1 million and 9.6 million, respectively, in 2018 [4]. Despite the rapid and precise interventions in cancer detection approaches developed during the last decade, the prognosis of cancer patients is poor due to distant metastasis occurrence and recurrence [5, 6]. In addition, maximum cancer cases are initially detected at advanced stages owing to the lack of reliable and sensitive diagnostic biomarkers, with a 5-year survival rate of less than

20% in many cancer subtypes [7, 8]. Therefore, a detailed understanding of the molecular processes governing cancer progression is needed to explore the novel diagnostic and prognostic biomarkers for cancer detection and the development of more effective therapeutic strategies.

The GINS complex is consist of four different subunits, including Sld5, Psf1, Psf2, and Psf3, which are also known as GINS4, GINS1, GINS2, and GINS3. In eukaryotes, the GINS complex binds to Cdc45 and Mcm2-7 to form the replicative helicase CMG complex, which unties double-stranded DNA before moving the replication fork in the replication process [9]. According to previous studies, during the replication process, the GINS complex mainly enhances the enzymatic activity of the minichromosome maintenance (MCM) complex by binding to it, which further helps to recruit the other essential factors involved in the formation of replisome progression complex that leads to the initiation and elongation of replication [10, 11]. Furthermore, newly emerging evidence has also reported that GINS may act as a key factor for regulating eukaryotic DNA polymerases such as DNA polymerase (Pol) ε [12] and the DNA Pol α-primase complex [13]. In the GINS complex, GINS4 or sld5 is the most important component that is required for the GINS complex assembly and to initiate and elongate the replication process in eukaryotes [14]. In addition, GINS4 also plays a key role in regulating embryogenesis in mice and cell cycle regulation and maintenance of genomic integrity in Drosophila [15, 16]. Previous reports have revealed the GINS4 up-regulation in different human cancers, including breast cancer (BRCA) [17], adrenal cortex adenocarcinoma (ACC) [18], colorectal cancer (CRC) [19], non-small cell lung cancer (NSCLC) [20], bladder cancer [21], pancreatic cancer [22], and gastric cancer [23]. Additionally, it was also observed that elevated GINS4 expression is significantly associated with the lower overall survival (OS) duration of gastric cancer, CRC, NSCLC, and pancreatic cancer patients [19, 20, 23]. Altogether, GINS4 has a vital contribution to the progression of cancers, and we speculate that it can probably be utilized as an important target for cancer detection and treatment potentially. Moreover, no previous studies about the GINS4 based on pan-cancer analysis.

Therefore, in this study, we attempted to systematically analyze and validate the GINS4 expression across multiple human cancer subtypes using various online available databases and bioinformatics tools. In addition, we analyzed the correlation among GINS4 expression and various other parameters in distinct cancer subtypes, including OS duration, RFS duration, genetic mutations, copy number variations (CNVs), promoter methylation level, tumor purity, and immune cells infiltration. Then, we also identified the GINS4-associated miRNAs, TFs, genes, and performed their Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, and finally developed a gene-drug interaction network.

Materials and methods

UALCAN

UALCAN (http://ualcan.path.uab.edu/) is an online database that is created to analyze TCGA multi-omics cancer-related data [24]. With the help of UALCAN, we analyzed the transcription expression level of GINS4 in distinct human cancer subtypes through pan-cancer analysis using default settings. The transcription expression level of GINS4 was measured in terms of transcript per million (TPM) reads, and a student t-test was used for statistics purpose. A *P*-value <0.05 represents the significant scores.

Kaplan-Meier plotter

Kaplan-Meier Plotter tool (https://kmplot.com/analysis/) is developed to check the impact of the gene(s) of interest on the survival duration of patients suffering from distinct types of cancer [25]. In our study, we utilized Kaplan-Meier Plotter tool with default settings to find the association between the GINS4 expression and distinct cancer types related OS and RFS survival rates. For this purpose, the cancer specimens were categorized into two categories based on their median expression level (high expression level v/s low expression level), and a *P*-value <0.05 was used to represent the significant scores.

TNMplot database

TNMplot (https://www.tnmplot.com/) [26] was used in this study to analyze GINS4 expression in normal and metastatic tissues of different cancers. For statistics purpose, a student's

t-test was employed in this database, and a *P*-value <0.05 was used to represent the significant scores.

GENT2, GEPIA, DriverDBv3, and HPA databases

GENT2 (http://gent2.appex.kr/), GEPIA (http://gepia.cancer-pku.cn/) DriverDBv3 (http://driverdb.tms.cmu.edu.tw/), and Human Protein Atlas (HPA) (https://www.proteinatlas.org/) database offer a reliable multi-omics analysis of the cancer-related TCGA data [27-29]. In this study, to validate the transcription and translation expression levels of GINS4 in distinct cancer subtypes, we employed these databases to analyze the GINS4 differential expression patterns in new independent cancer cohorts with default settings. In GENT2, GEPIA, and DriverDBv3 databases, the transcription expression level was measured in terms of transcript per million (TPM) reads, and a student t-test was used for statistics purpose. While in HPA, the protein expression level was graded as not detected, low, medium, and high, based on the intensity of staining and fraction of the stained cells. A *P*-value <0.05 represents the significant scores.

MEXPRESS

MEXPRESS (https://mexpress.be/) is developed to visualize the TCGA expression data and identify correlations between promoter methylation and expression level [30]. In this study, the correlation between GINS4 transcription expression and promoter methylation levels in distinct cancer subtypes were computed via this tool using Pearson correlation analysis. A *P*-value <0.05 represents the significant scores.

The cBioportal database

cBioPortal (http://www.cbioportal.org/) is a user-friendly application that offers data on genetic mutations, copy number variations (CNVs), and transcription expression from samples of various cancer subtypes [31]. In this study, we chose TCGA PanCancer Atlas datasets to investigate GINS4-associated genetic mutations and mutational hotspots in different human cancers using default settings.

PPI network construction, visualization, functional, and pathway analysis

In the current study, STRING (Search Tool for the Retrieval of Interacting Genes) biological tool [32] was used to obtain the protein-protein interaction (PPI) network of GINS4-associated genes with a confidence score of ≥ 0.7. Later, functional, and pathway analysis of GINS4 enriched genes was performed via DAVID (v6.8, http://david.ncifcrf.gov/summary.jsp) [33] and a *P*-value was used <0.05 to represents the significant scores.

Correlation between GINS4 and its associated genes across different cancers

The GEPIA (http://gepia.cancer-pku.cn/) was conducted in this study to evaluate pairwise gene correlations between GINS4 and its other associated genes using the "Correlation Analysis" module. A *P*-value <0.05 represents the significant scores.

Enrichr database analysis

Enrichr (https://maayanlab.cloud/Enrichr/) [34] was used in this study with default settings to identify GINS4 targeted miRNAs and TFs from TRRUST 2019 and miRTarBase 2017 sources. The top 10 significantly (P<0.05) enriched items were displayed using Enrichr.

MuTarget analysis

The MuTarget (https://www.mutarget.com/result) is an online platform that associates gene expression alterations with mutational status in human cancers. Via this platform, mutant genes altering the expression of a gene of interest could be identified [35]. In our study, we used this platform to identify the mutant genes responsible for the expression alteration in the GINS4 gene in different cancers with default thresh-holds of P<0.05 and FC >1.4.

Tumor purity, immune cells infiltration, and GINS4 expression in cancer patients of distinct subtypes

The TIMER database (https://cistrome.shinyapps.io/timer/) offers helpful services to analyze the association between gene expression, tumor purity, and the infiltration level of different immune cells [36]. In this study, GINS4 was queued in the 'Gene module' tool of TIMER to find the Spearman correlation between tumor purity, immune cells infiltration such as B cells, macrophages, neutrophils, CD4+ T cells, and CD8+ T cells, and GINS4 expression in distinct cancer subtypes using default settings. A *P*-value <0.05 represents the significant scores.

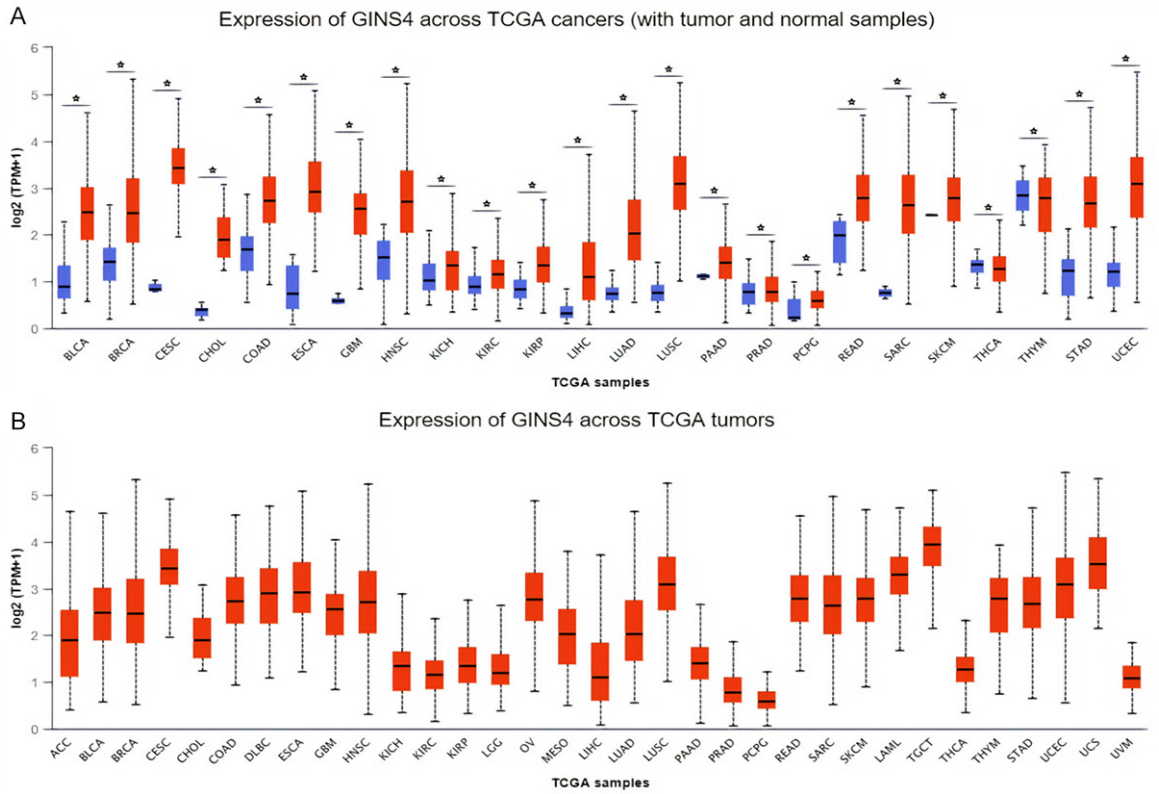


Figure 1. Differential transcription expression analysis of GINS4 gene in cancerous and normal tissues via pan-cancer cancer analysis using UALCAN. (A) Pan-cancer expression analysis results of GINS4 across cancerous samples paired with normal controls, and (B) Pan-cancer expression analysis results of GINS4 in only cancer samples. Blue color represents the normal samples while red color indicates the cancer samples. *P<0.05.

GINS4 gene-drug interaction network analysis

The GINS4 gene-drug interaction network was built via Cytoscape 3.8.0 based on the data obtained from the Comparative Toxicogenomics Database (CTD) with default settings [37]. By queuing the CTD database, different potential compounds that are capable to regulate GINS4 expression were identified through this network.

Results

GINS4 expression in pan-cancer

In this study, we used UALCAN to analyze the GINS4 transcription expression across 24 major human cancers relative to controls. Our results showed that GINS4 expression was elevated significantly (P<0.05) in all 24 cancer subtypes, especially in Liver hepatocellular carcinoma (LIHC), Cholangiocarcinoma (CHOL), Kidney renal papillary cell carcinoma (KIRP), Esophageal carcinoma (ESCA), Colon adenocarcinoma (COAD), Cervical squamous cell carcinoma (CESC), Breast invasive carcinoma (BRCA), and Stomach adenocarcinoma (STAD) (**Figure 1**).

Correlation analysis of GINS4 expression with OS, RFS, and metastasis

We used the KM plotter tool to analyze the association between GINS4 expression and OS or RFS in 24 human cancer subtypes. The obtained KM curves highlighted that elevated expression of GINS4 was significantly (P<0.05) linked to the reduced OS and RFS duration in five subtypes of cancer including ESCA (HR =2.8, 95% CI: 1.35-4.93, P=0.0029, HR =3.46, 95% CI: 0.49-24.72, P=0.019), KIRC (HR =1.51, 95% CI: 1.11-2.05, P=0.008, HR =2.15, 95% CI: 0.78-6.78, P=0.018), LIHC (HR =1.79, 95% CI: 1.24-2.59, P=0.0017, HR =1.55, 95% CI: 1.12-2.16, P=0.0084), LUAD (HR =1.59, 95% CI: 1.18-2.12, P=0.0018, HR =2.01, 95% CI: 1.29-3.14, P=0.0017), and UCEC (HR =2.13, 95% CI: 1.39-3.27, P=0.00038, HR =1.7, 95% CI: 1.01-2.87, P=0.044) UCEC (**Figure 2A**, **2B**). Furthermore, GINS4 notable overexpression

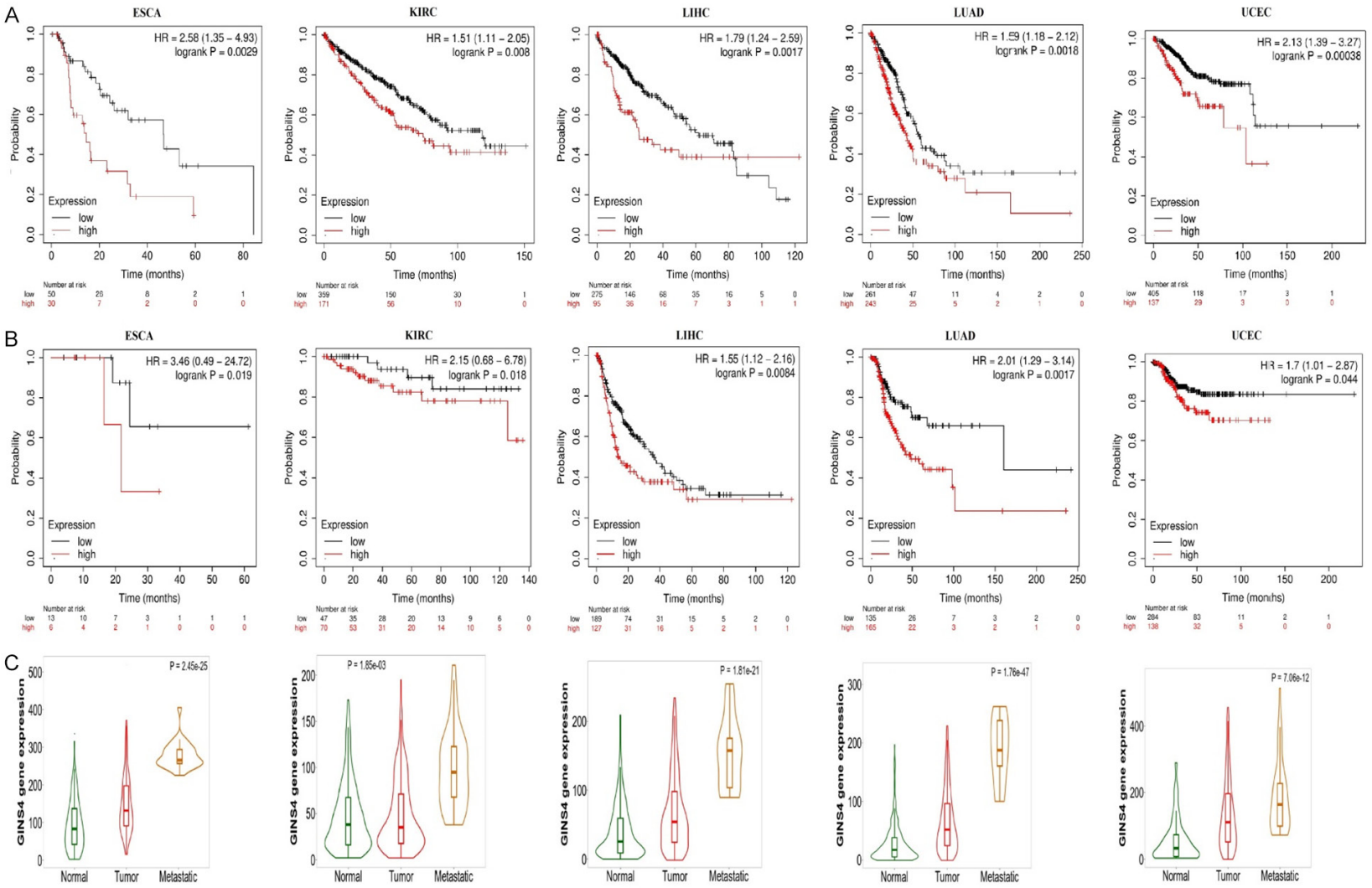


Figure 2. High expression level of GINS4 expression is an adverse prognostic factor in ESCA, KIRC, LIHC, LUAD, and UCEC. (A) Survival analysis revealed that higher GINS4 expressions reduced OS duration in ESCA, KIRC, LIHC, LUAD, and UCEC, (B) Survival analysis revealed that higher GINS4 expressions reduced RFS duration in ESCA, KIRC, LIHC, LUAD, and UCEC, and (C) A correlation analysis of GINS4 with metastasis in ESCA, KIRC, LIHC, LUAD, and UCEC tissues. A *P*-value <0.05 was considered as significant.

Table 1. Clinicopathalogical features-specific expression pattern of GINS4 in ESCA, KIRC, LIHC, LUAD, and patients

GINS4 expression across ESCA patients with distinct clinicopathological features			
Different cancer stages-based GINS4 expression pattern relative to normal (n=11) control samples	Stage 1 (n=13)	↑ (up-regulation)	*P*-value <0.05
	Stage 2 (n=78)	↑ (up-regulation)	
	Stage 3 (n=55)	↑ (up-regulation)	
	Stage 4 (n=9)	↑ (up-regulation)	
Different patient's races-based GINS4 expression pattern relative to normal (n=11) control samples	Caucasian (n=113)	↑ (up-regulation)	*P*-value <0.05
	African-American (n=5)	↑ (up-regulation)	
	Asian (n=46)	↑ (up-regulation)	
Different patient's gender-based GINS4 expression pattern relative to normal (n=11) control samples	Male (n=157)	↑ (up-regulation)	*P*-value <0.05
	Female (n=26)	↑ (up-regulation)	
Different patient's ages-based GINS4 expression pattern relative to normal (n=11) control samples	21-40 Yrs (n=3)	↑ (up-regulation)	*P*-value <0.05
	41-60 Yrs (n=89)	↑ (up-regulation)	
	61-80 Yrs (n=76)	↑ (up-regulation)	
	81-100 Yrs (n=15)	↑ (up-regulation)	
GINS4 expression across KIRC patients with distinct clinicopathological features			
Different cancer stages-based GINS4 expression pattern relative to normal (n=72) control samples	Stage 1 (n=267)	↑ (up-regulation)	*P*-value <0.05
	Stage 2 (n=57)	↑ (up-regulation)	
	Stage 3 (n=123)	↑ (up-regulation)	
	Stage 4 (n=84)	↑ (up-regulation)	
Different patient's races-based GINS4 expression pattern relative to normal (n=72) control samples	Caucasian (n=462)	↑ (up-regulation)	*P*-value <0.05
	African-American (n=56)	↑ (up-regulation)	
	Asian (n=8)	↑ (up-regulation)	
Different patient's gender-based GINS4 expression pattern relative to normal (n=72) control samples	Male (n=345)	↑ (up-regulation)	*P*-value <0.05
	Female (n=185)	↑ (up-regulation)	
Different patient's ages-based GINS4 expression pattern relative to normal (n=72) control samples	21-40 Yrs (n=26)	↑ (up-regulation)	*P*-value <0.05
	41-60 Yrs (n=238)	↑ (up-regulation)	
	61-80 Yrs (n=246)	↑ (up-regulation)	
	81-100 Yrs (n=23)	↑ (up-regulation)	
GINS4 expression across LIHC patients with distinct clinicopathological features			
Different cancer stages-based GINS4 expression pattern relative to normal (n=50) control samples	Stage 1 (n=168)	↑ (up-regulation)	*P*-value <0.05
	Stage 2 (n=84)	↑ (up-regulation)	
	Stage 3 (n=82)	↑ (up-regulation)	
	Stage 4 (n=6)	↑ (up-regulation)	
Different patient's races-based GINS4 expression pattern relative to normal (n=50) control samples	Caucasian (n=177)	↑ (up-regulation)	*P*-value <0.05
	African-American (n=17)	↑ (up-regulation)	
	Asian (n=157)	↑ (up-regulation)	
Different patient's gender-based GINS4 expression pattern relative to normal (n=50) control samples	Male (n=245)	↑ (up-regulation)	*P*-value <0.05
	Female (n=117)	↑ (up-regulation)	
Different patient's ages-based GINS4 expression pattern relative to normal (n=50) control samples	21-40 Yrs (n=27)	↑ (up-regulation)	*P*-value <0.05
	41-60 Yrs (n=140)	↑ (up-regulation)	
	61-80 Yrs (n=181)	↑ (up-regulation)	
	81-100 Yrs (n=10)	↑ (up-regulation)	
GINS4 expression across LUAD patients with distinct clinicopathological features			
Different cancer stages-based GINS4 expression pattern relative to normal (n=59) control samples	Stage 1 (n=277)	↑ (up-regulation)	*P*-value <0.05
	Stage 2 (n=125)	↑ (up-regulation)	
	Stage 3 (n=85)	↑ (up-regulation)	
	Stage 4 (n=28)	↑ (up-regulation)	
Different patient's races-based GINS4 expression pattern relative to normal (n=59) control samples	Caucasian (n=387)	↑ (up-regulation)	*P*-value <0.05
	African-American (n=51)	↑ (up-regulation)	
	Asian (n=08)	↑ (up-regulation)	
Different patient's gender-based GINS4 expression pattern relative to normal (n=59) control samples	Male (n=238)	↑ (up-regulation)	*P*-value <0.05
	Female (n=276)	↑ (up-regulation)	

Different patient's ages-based GINS4 expression pattern relative to normal (n=59) control samples	21-40 Yrs (n=12)	↑ (up-regulation)	*P*-value <0.05
	41-60 Yrs (n=90)	↑ (up-regulation)	
	61-80 Yrs (n=149)	↑ (up-regulation)	
	81-100 Yrs (n=32)	↑ (up-regulation)	
GINS4 expression across UCEC patients with distinct clinicopathological features			
Different cancer stages-based GINS4 expression pattern relative to normal (n=35) control samples	Stage 1 (n=341)	↑ (up-regulation)	*P*-value <0.05
	Stage 2 (n=52)	↑ (up-regulation)	
	Stage 3 (n=124)	↑ (up-regulation)	
	Stage 4 (n=29)	↑ (up-regulation)	
Different patient's races-based GINS4 expression pattern relative to normal (n=35) control samples	Caucasian (n=374)	↑ (up-regulation)	*P*-value <0.05
	African-American (n=107)	↑ (up-regulation)	
	Asian (n=20)	↑ (up-regulation)	
Different patient's gender-based GINS4 expression pattern relative to normal (n=35) control samples	Male (n=268)	↑ (up-regulation)	*P*-value <0.05
	Female (n=147)	↑ (up-regulation)	
Different patient's ages-based GINS4 expression pattern relative to normal (n=35) control samples	21-40 Yrs (n=18)	↑ (up-regulation)	*P*-value <0.05
	41-60 Yrs (n=189)	↑ (up-regulation)	
	61-80 Yrs (n=292)	↑ (up-regulation)	
	81-100 Yrs (n=45)	↑ (up-regulation)	

was also found in the metastatic samples of ESCA, KIRC, LIHC, LUAD, and UCEC relative to primary tumor samples and normal controls (**Figure 2C**). Altogether, our data suggested that GINS4 might have a significant contribution to the development and progression of ESCA, KIRC, LIHC, LUAD, and UCEC, thus the next part of this study will primarily focus on the unique role of GINS4 in those five types of human cancers.

GINS4 expression in ESCA, KIRC, LIHC, LUAD, and UCEC patients with different clinicopathological features

Generally, gene expression is often varied clinicopathological features-wise. We analyzed the relationship between GINS4 expression and different clinicopathological features of ESCA, KIRC, LIHC, LUAD, and UCEC using the UALCAN database. Our results demonstrated that GINS4 expression level was closely correlated with the clinicopathological features of ESCA, KIRC, LIHC, LUAD, and UCEC including cancer stages, races, genders, and ages (**Table 1**). The clinicopathological features of the ESCA, KIRC, LIHC, LUAD, and UCEC cohorts are provided in Supplementary Tables 1, 2 and 3.

GINS4 expression validation in new cohorts

Based on GENT2, GEPIA, DriverDBv3, and HPA databases, we further validated GINS4 expression at both transcriptional and translational levels using independent cohorts of ESCA, KIRC, LIHC, LUAD, and UCEC. As per expectations, our results were in agreement with the results of UALCAN, indicating the robustness of the evidence. The expression analysis via GENT2, GEPIA, and DriverDBv3 revealed the significant ($P<0.05$) higher expression of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC patients relative to normal controls at the transcriptional level (**Figure 3A-C**), moreover, the expression analysis of GINS4 via HPA also revealed that normal esophageal, kidney, liver, lung, and endometrial tissues had low GINS4 IHC staining, while cancer tissues had medium or high staining (**Figure 3D**). Taken together, our results have validated that GINS4 is overexpressed at both transcriptional and translational levels in ESCA, KIRC, LIHC, LUAD, and UCEC as compared to the normal controls.

GINS4 promoter methylation negatively correlated its expression

Hypermethylation of the gene promoter region regulates transcriptional silencing. On the other hand, hypomethylation can result in the enhanced gene expression. A variety of cancers has been linked to the promoter-specific methylation levels and accompanied gene dysregulation [38]. In this study, we have chosen GINS4 methylation sites from the MEXPRESS database. This is one of the most reliable databases built to analyze the association between gene expression and methylation levels at CpG islands. As shown in **Figures 4** and **5**, we observed that the promoter methylation values

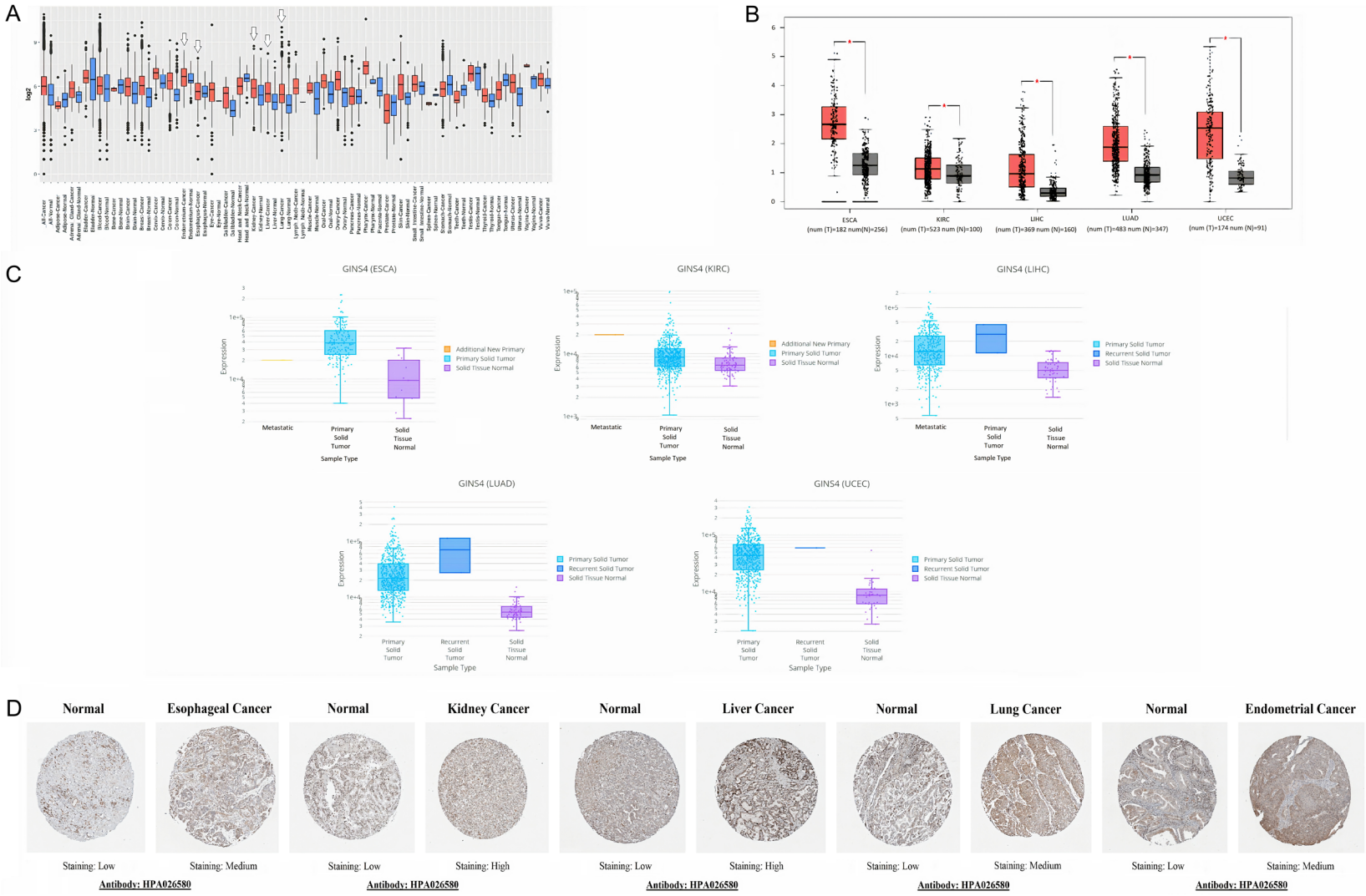


Figure 3. Transcription and translational level expression validation of GINS4 in new independent cohorts of ESCA, KIRC, LIHC, LUAD, and UCEC via GENT2, GEPIA, DriverDBv3 and HPA databases. (A) Transcription level expression validation of GINS4 via GENT2, (B) Transcription level expression validation of GINS4 via GEPIA, (C) Transcription level expression validation of GINS4 via DriverDBv3, and (D) Translation level expression validation of GINS4 via HPA. A *P*-value of <0.05 was selected as cutoff criterion.

Figure 4. A MRXPRESS based correlation analysis between GINS4 expression and its promoter methylation in ESCA, KIRC, and LIHC. (A) In ESCA, (B) In KIRC, and (C) In LIHC. A negative sign indicates the negative correlation between GINS4 expression and its promoter methylation using a specific probe at a specific CpG island. A *P*-value of <0.05 was selected as cutoff criterion.

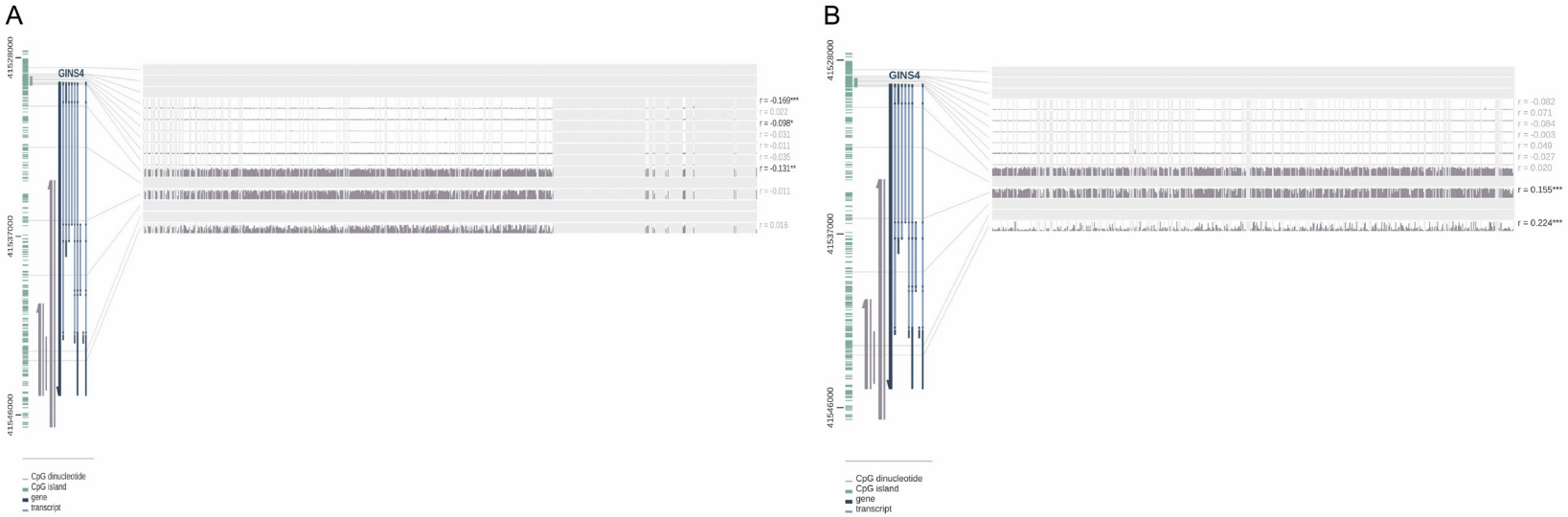


Figure 5. A MRXPRESS based correlation analysis between GINS4 expression and its promoter methylation in LUAD and UCEC. (A) In LUAD, and (B) In UCEC. A negative sign indicates the negative correlation between GINS4 expression and its promoter methylation using a specific probe at a specific CpG island. A *P*-value of <0.05 was selected as cutoff criterion.

obtained from the different CpG dinucleotides in BLCA, HNSC, KIRP, LUAD, and UCEC were significant (P<0.05) negatively correlated with GINS4 expression levels.

Genetic alterations of GINS4

For inquiring about GINS4-associated genetic alterations and CNVs we used the cBioportal database. In this analysis, PanCancer Atlas ESCA, KIRC, LIHC, LUAD, and UCEC datasets were queued and genetic alterations and CNVs were observed in only 7%, 1.1%, 6%, 6%, and 6% cases of ESCA, KIRC, LIHC, LUAD, and UCEC, respectively (**Figure 6A**). Deep amplifications abnormality was most common in these cancers followed by deep deletions (**Figure 6A**). Taken together, it is speculated that GINS4 harbors genetic alteration in small numbers of ESCA, KIRC, LIHC, LUAD, and UCEC samples.

Mutational hotspot analysis of GINS4

To further identify the mutational hotspots of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC cancer subtypes, we analyzed PanCancer Atlas ESCA, KIRC, LIHC, LUAD, and UCEC datasets using cBioportal. In ESCA, LIHC, and LUAD, the mutational hotspots of the most frequently observed mutations, including one nonsense mutation (Q567*) in ESCA, one missense mutation (D6G) in LIHC, and other one missense mutation (A145S) in LUAD lie outside the Sld5 domain of the GINS4, which plays an important role in the initiation of the replication process. On the other hand, in UCEC, the GINS4 mutational hotspots of the most frequently observed missense mutation (P119H) lies within the Sld5 domain (**Figure 6B**). Moreover, no GINS4 mutation was identified in case of KIRC (**Figure 5B**). Taken together, we observed different GINS4 mutational hotspots in ESCA, LIHC, LUAD, and UCEC which overall suggested a high level of complexity regarding GINS4 mutations.

A PPI network and enrichment analysis of GINS4

We further conducted STRING and Cytoscape analysis to identify the GINS4 enriched genes. Functional interaction network analysis showed that GINS4 physically interacts with 23 different other genes (**Figure 7**). We next performed the GO and KEGG analysis of GINS4 associated genes via DAVID tool to determine the GINS4 associated genes functions and pathways. Results revealed the enrichment of GINS4-associated genes in biological processes (BP), molecular function (MF), cell composition (CC), and KEGG pathways. GINS4-associated genes were significantly (P<0.05) enriched in DNA replication BP, DNA helicase activity MF, MCM complex CC, and different KEGG terms, including DNA replication, Cell cycle, Glucagon signaling pathway, Biosynthesis of antibiotics, and Cysteine and methionine metabolism (**Figure 7**; Supplementary Table 4).

Correlation analysis between GINS4 and the expression of its other associated genes

Via GEPIA, we further analyzed the correlations among GINS4 and its other physically associated 23 genes expression across ESCA, KIRC, LIHC, LUAD, and UCEC samples. In view of our results, GINS4 expression was found to be positively correlated with the expressions of all of its associated genes including RECQL4, SSRP1, TIPIN, NCDN, LDHB, GINS3, DONSON, KIF16B, GINS2, MCM2, MCM3, PPIL3, PAICS, MCM5, AHCYL1, ADHA1, MCM7, GINS1, WDHD1, SIK1, DUSP13, CD2BP2, and POLA1 (**Figure 8**).

Identification of miRNAs and TFs that potentially regulate GINS4 expression

Through enrichr, we predicted ten highly significant miRNAs (hsa-miR-193b-3p, hsa-miR-215-5p, hsa-miR-192-5p, hsa-miR-3613-3p, hsa-miR-372-5p, hsa-miR-373-5p, hsa-miR-371b-5p hsa-miR-616-5p, hsa-miR-371a-5p, and hsa-miR-6849-5p), and ten TFs (E2F4, RBL2, MEN1, E2F1, E2F3, BRCA1, E2F4, OTX2, FOXM1, and FOXO3) which could potentially regulate GINS4 expression (**Figure 9**). Ultimately, all these clues indicate that GINS4 expression can be regulated through different miRNAs and TFs.

Correlations among GINS4 expression and crucial mutant genes

To correlate GINS4 expression with different other mutant genes, we used MuTarget to select top the 3 mutant genes associated with GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC, respectively, with default settings. The selected top 3 mutant genes which are positively correlated with the expression GINS4 are

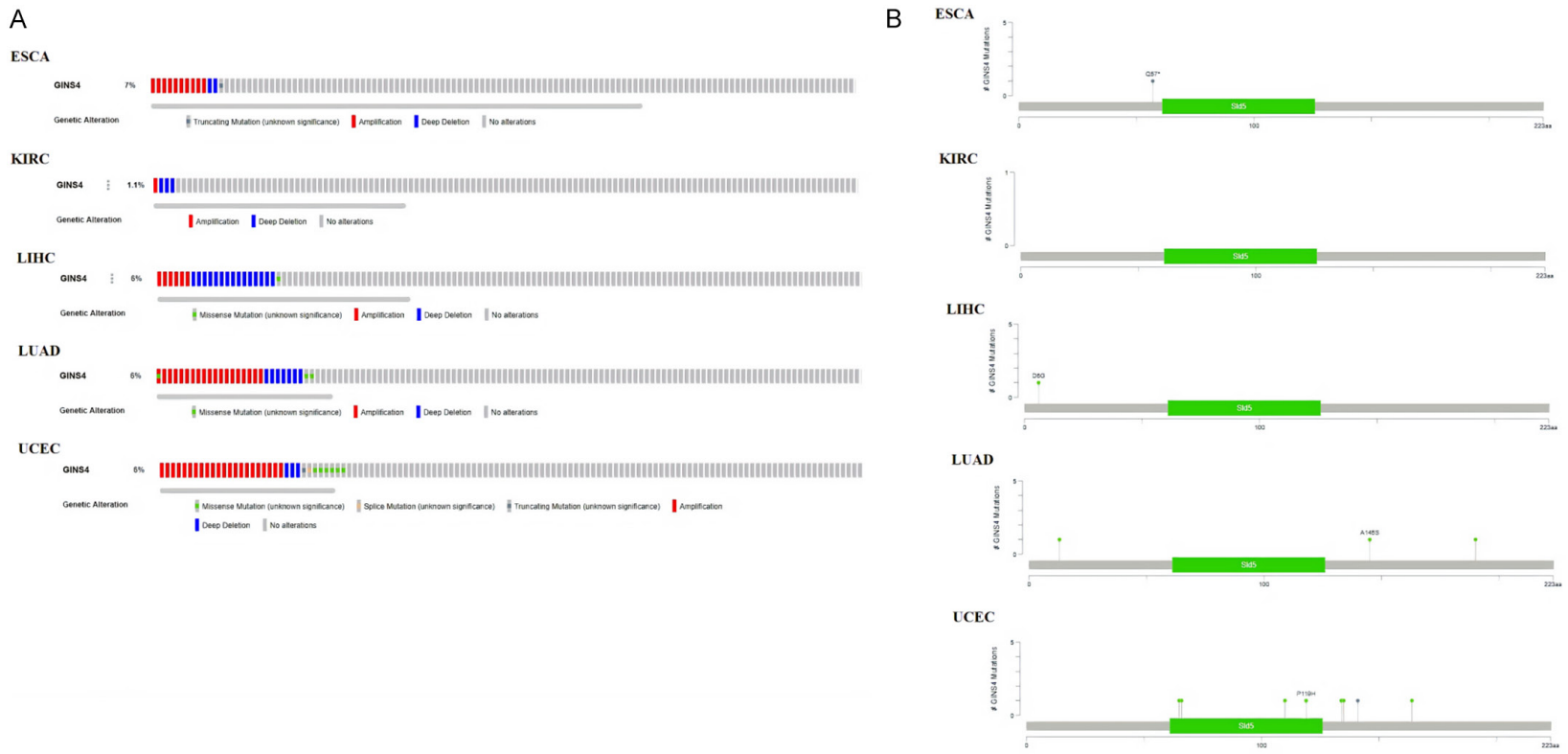


Figure 6. GINS4 genetic alterations, CNVs, and mutational hotspot status in ESCA, KIRC, LIHC, LUAD, and UCEC. (A) Genetic alterations and CNVs status of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC and (B) Mutational hotspot analysis of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC.

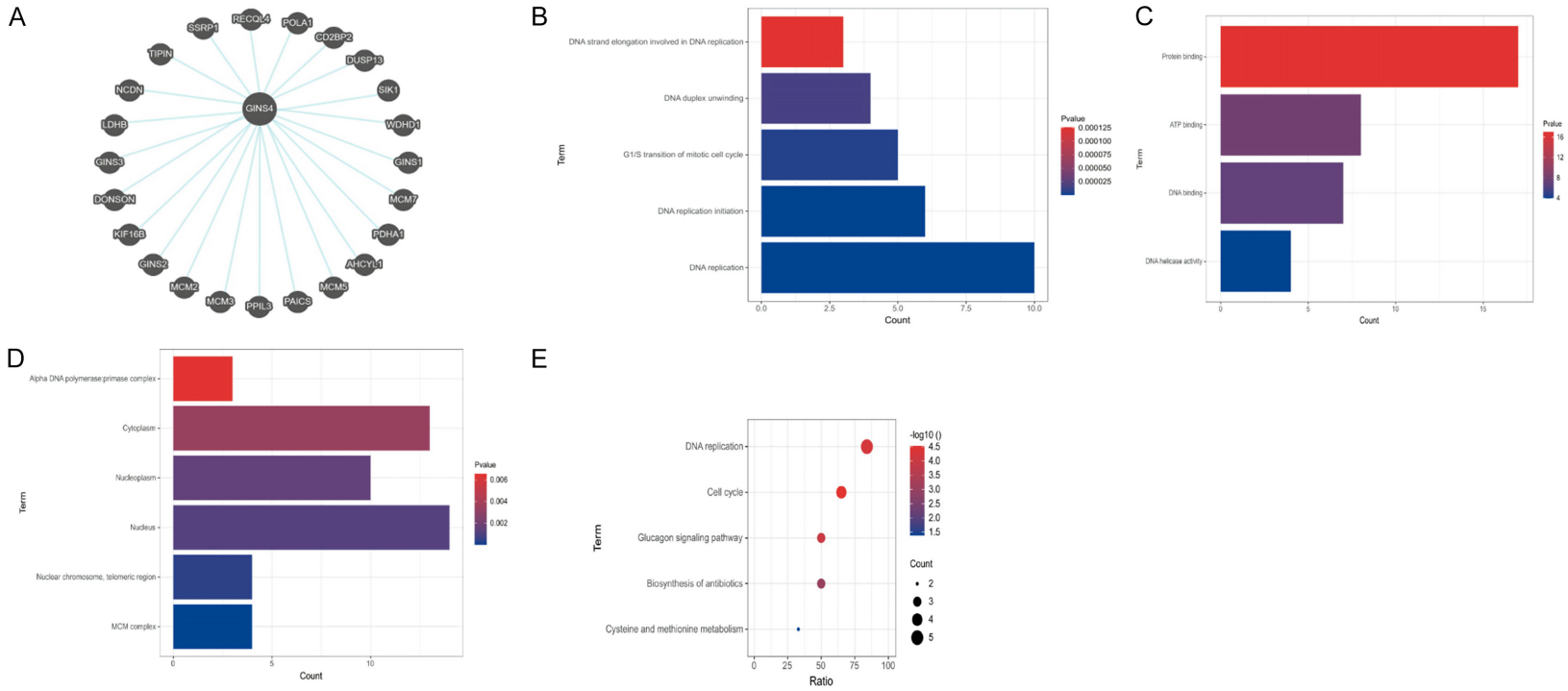


Figure 7. A PPI network, GO and KEGG analysis of the GINS4 enriched genes. (A) A PPI network of GINS4 enriched genes, (B) BP functional classification terms of the GINS4 enriched genes, (C) MF functional classification terms of the GINS4 enriched genes, and (D) CC functional classification terms of the GINS4 enriched genes, and (E) KEGG classification terms of the GINS4 enriched genes. A *P*-value <0.05 was considered as significant.

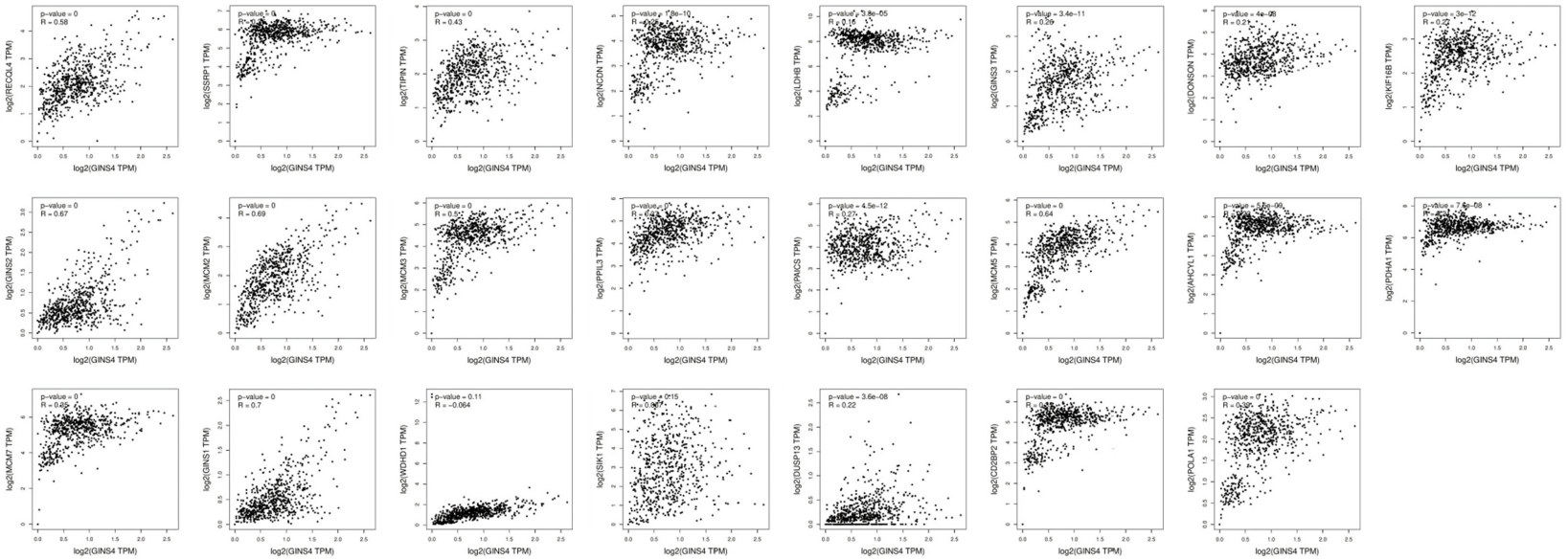
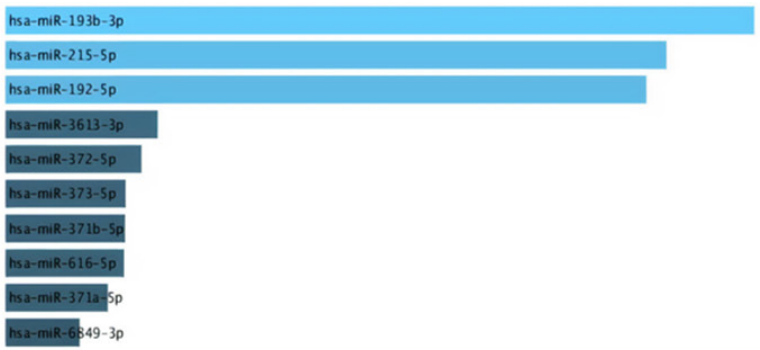


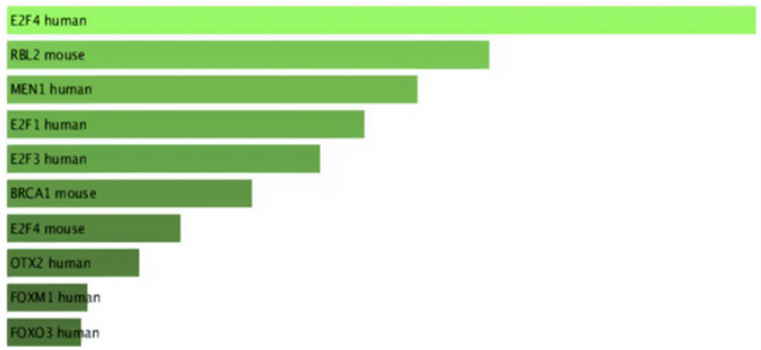
Figure 8. A GEPIA-based correlation analysis among GINS4 and its other associated genes expression across ESCA, KIRC, LIHC, LUAD, and UCEC samples. A *P*-value <0.05 was considered to indicate a statistically significant result.

A



Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	hsa-miR-193b-3p	1.697e-26	2.459e-23	14.37	852.74
2	hsa-miR-215-5p	1.094e-23	7.926e-21	13.70	724.49
3	hsa-miR-192-5p	4.955e-23	2.393e-20	11.64	597.71
4	hsa-miR-8088	0.001581	0.09049	40.59	261.80
5	mmu-miR-3078-5p	0.003594	0.1561	25.36	142.75
6	mmu-miR-6951-3p	0.02963	0.3639	40.19	141.43
7	mmu-miR-7116-3p	0.02963	0.3639	40.19	141.43
8	hsa-miR-372-5p	7.320e-7	0.0002121	7.66	108.21
9	hsa-miR-6813-3p	0.0003749	0.03395	12.91	101.88
10	hsa-miR-3713	0.005848	0.1893	19.32	99.33

B



Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	E2F4 human	1.265e-9	5.312e-8	74.65	1529.56
2	RBL2 mouse	0.00001419	0.0003266	87.89	981.15
3	MEN1 human	0.00001944	0.0003266	76.90	834.26
4	E2F1 human	1.313e-13	1.103e-11	24.43	724.50
5	E2F3 human	0.00003345	0.0004683	61.52	633.95
6	BRCA1 mouse	0.0006796	0.008155	67.67	493.57
7	E2F4 mouse	0.001085	0.01139	50.74	346.39
8	OTX2 human	0.001581	0.01207	40.59	261.80
9	FOXM1 human	0.003205	0.02071	27.05	155.37
10	FOXO3 human	0.003594	0.02157	25.36	142.75

Figure 9. Identification of GINS4 targeted miRNAs and TFS via Enrichr database. (A) GINS4 targeted miRNAs, and (B) GINS4 targeted TFS. A *P*-value <0.05 was considered as significant.

TP53, NELL2, and RUNX1 in ESCA, SLC22A4, PTPRZ1, and VARS2 in KIRC, TP53, CSMD3, and CDH10 in LIHC, TP53, KIF19, and RB1 in LUAD, and TP53, TCOF1, and ZNF780A in UCEC (**Figure 10**). Collectively, this information revealed that GINS4 strongly correlates with different other mutant genes in ESCA, KIRC, LIHC, LUAD, and UCEC. This new information may also enhance the knowledge of cancer development in those cancer subtypes.

Tumor purity and immune cells infiltration analysis of GINS4

Considering the involvement of GINS4 in the regulation of different pathways, including cell cycle and DNA replication, it was hypothesized that GINS4 expression level variations may contribute to alterations in the immune cells infiltration and may also associate with tumor purity. Therefore, we used the TIMER algorithm to evaluate the correlation among tumor purity, immune cells infiltrations including B cells, macrophages, neutrophils, CD4+ T cells, and CD8+ T cells level and GINS4 expression in ESCA, KIRC, LIHC, LUAD, and UCEC. As per the tumor purity analysis, we observed a negative correlation between GINS4 expression and tumor purity in KIRC (Rho =-0.098, *P*-value =3.56e-02) and LUAD (Rho =-0.002, *P*-value =961e-01) while positive correlation in ESCA (Rho =0.254, *P*-value =5.60e-04), LIHC (Rho =0.122, *P*-value =2.38e-02), and UCEC (Rho =0.054, *P*-value =3.53e-01) (**Figure 10**). Moreover, we also observed a different correlations between GINS4 expression and immune cells infiltration in those cancers, like in case of B cells, our results revealed a negative correlation between B cells infiltration and GINS4 expression in KIRC (Rho =-0.073, *P*-value =1.16e-01), LUAD (Rho =-0.188, *P*-value =2.78e-05), and UCEC (Rho =-0.175, *P*-value = 1.03e-01) while a positive correlation in LIHC (Rho =0.359, *P*-value =6.34e-12), and ESCA (Rho =0.047, *P*-value =5.33e-01) (**Figure 11**). In case of macrophages, a positive correlation was revealed between macrophages infiltration and GINS4 expression in ESCA (Rho =0.009, *P*-value =9.06e-01), KIRC (Rho =0.252, *P*-value =4.17e-08), LIHC (Rho =0.35, *P*-value =2.14e-11), and LUAD (Rho =0.162, *P*-value =2.97e-04) while a negative correlation in UCEC (Rho =-0.28, *P*-value =8.26e-03). In case of neutrophils, a positive correlation was observed between neutrophils infiltration and GINS4 expression in ESCA (Rho =0.08, *P*-value =2.87e-01), KIRC (Rho =0.364, *P*-value =7.40e-16), LIHC (Rho =0.193, *P*-value =3.01e-04), LUAD (Rho =0.217, *P*-value =1.13e-06), and UCEC (Rho =0.14, *P*-value =1.93e-01). In case of CD4+ T cells, a positive correlation was seen between CD4+ T cells infiltration and GINS4 expression in ESCA (Rho =0.044, *P*-value =5.56e-01), KIRC (Rho =0.212, *P*-value =4.26e-06), LIHC (Rho =0.161, *P*-value =2.64e-03), and UCEC (Rho =0.01, *P*-value =9.26e-01) while a negative correlation in LUAD (Rho =-0.124, *P*-value =5.77e-03). Finally, in case of CD8+ T cell, a positive correlation was seen between CD8+ T cells infiltration and GINS4 expression in ESCA (Rho =0.006, *P*-value =9.36e-01), LIHC (Rho =0.163, *P*-value =2.41e-03), and LUAD (Rho =0.124, *P*-value =5.76e-03) while a negative correlation in KIRC (Rho =-0.115, *P*-value =7.54e-01) and UCEC (Rho =0.31, *P*-value =3.26e-03) (**Figure 11**).

Gene-drug interaction network analysis of the GINS4

To identify different available potential compounds targeting GINS4, a gene-drug interaction network was carried out using the Comparative Toxicogenomics Database (CTD) and Cytoscape. As highlighted in **Figure 12**, a total of 18 compounds were identified that could impact GINS4 expression. For example, aflatoxin B1 and dorsomorphin could elevate the expression level of GINS4 while cyclosporine and bisphenol A could reduce GINS4 expression level (**Figure 12**).

Discussion

Cancer is characterized by poor clinical outcomes and a higher rate of mortality [39]. Therefore, cancer patients often have the worst prognosis and thus, it is urgent to disclose the potentially shared ideal molecular biomarker for different cancers together that could help to enhance the diagnosis and treatment efficacy of these cancers as a shared target.

GINS complex was initially discovered by *Boskovic et al.* [40]. Recent data suggested that one of the main GINS complex subunits, the GINS4, is overexpressed in a few cancer subtypes including breast cancer (BRCA) [17], adrenal cortex adenocarcinoma (ACC) [18],

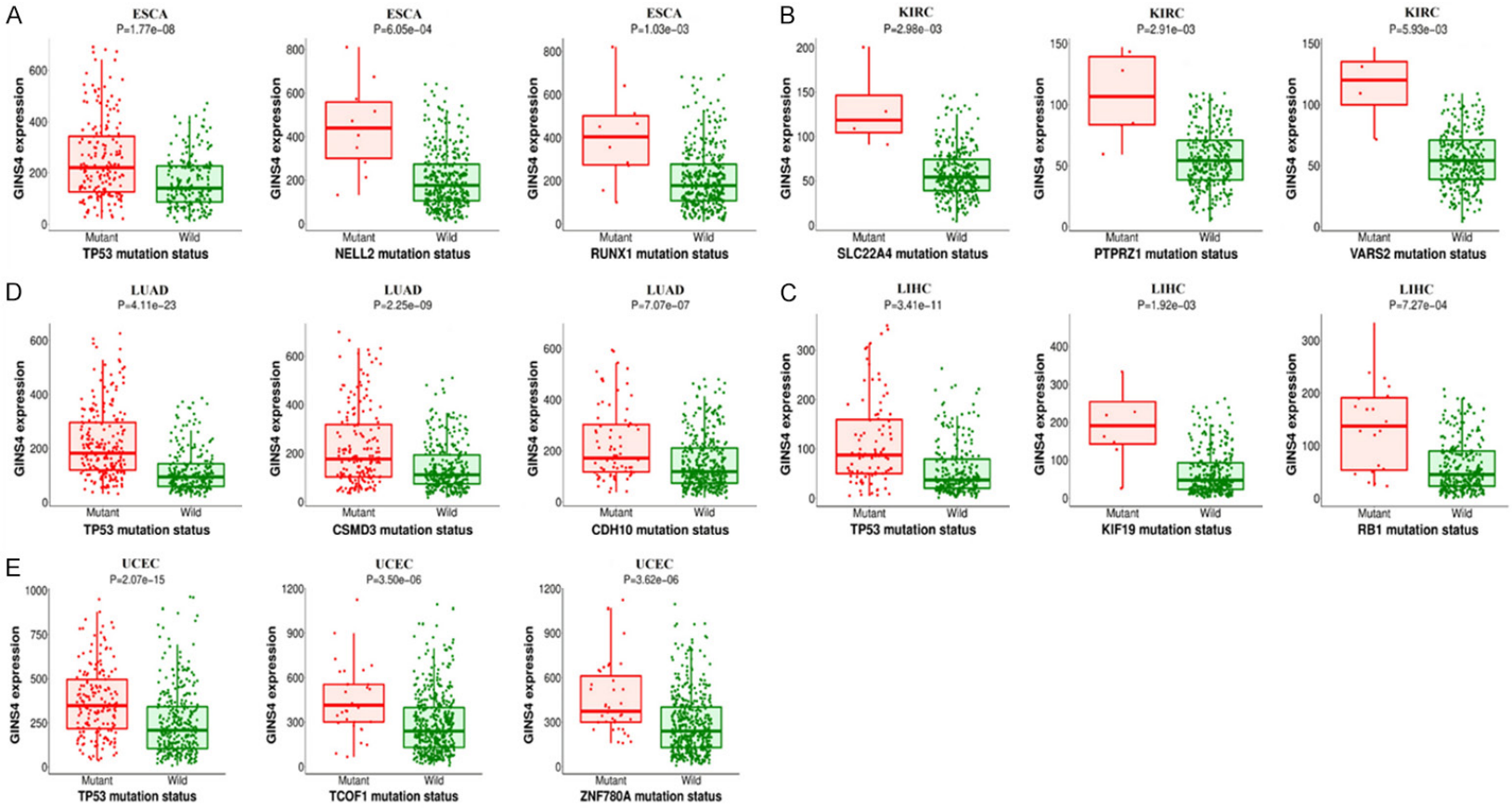


Figure 10. Positively correlated mutant genes with GINS4 in CESC, ESCA, HNSC, and KIRC from MuTarget. (A) Top 3 correlated genes with GINS4 in ESCA, (B) Top 3 correlated genes with GINS4 in KIRC, (C) Top 3 correlated genes with GINS4 in LIHC, (D) Top 3 correlated genes with GINS4 in LUAD, and (E) Top 3 correlated genes with GINS4 in UCEC. A *P*-value <0.05 was consider as significant.

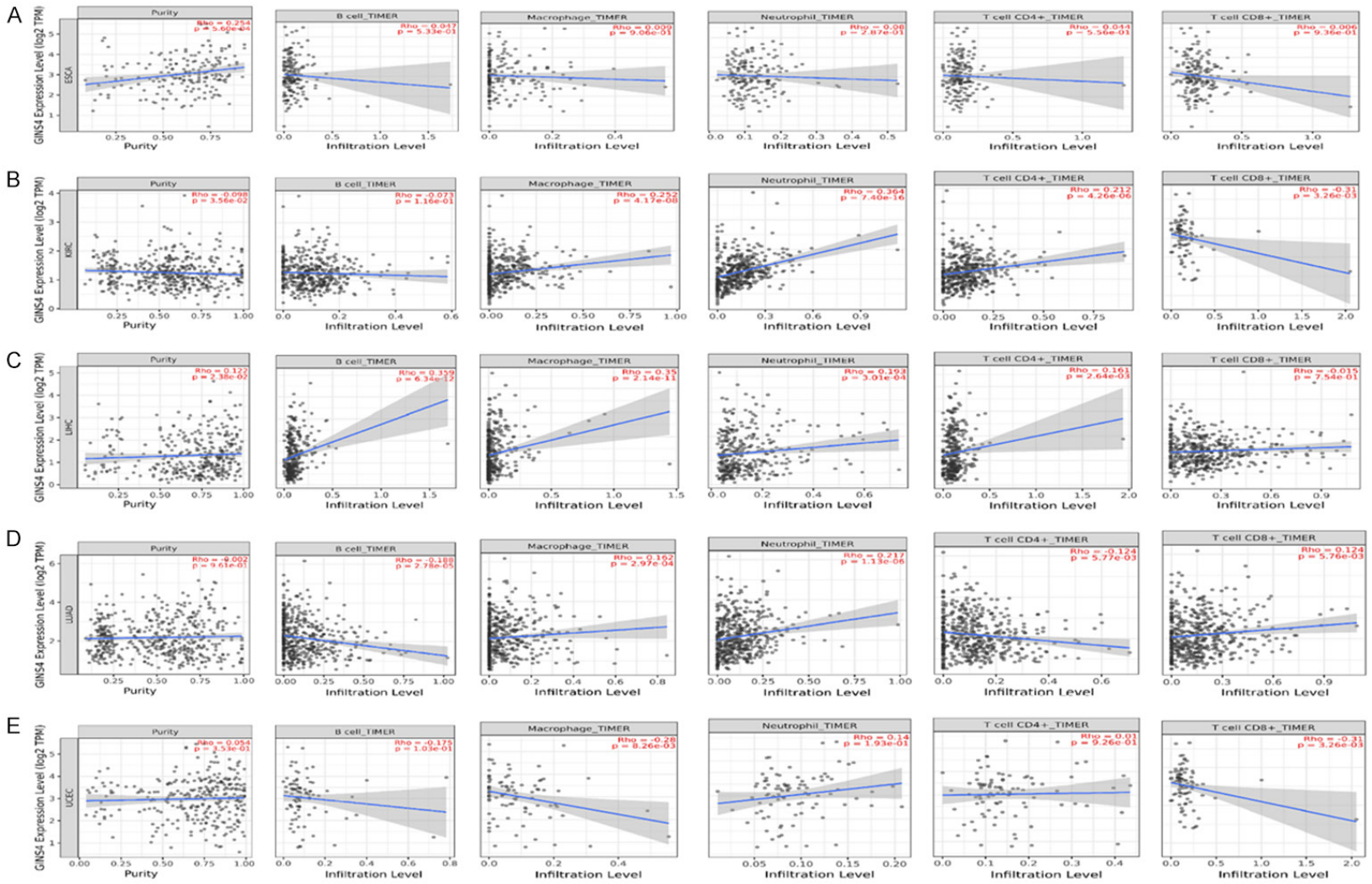


Figure 11. GINS4 correlation with tumor purity and immune cells infiltration in ESCA, KIRC, LIHC, LUAD, and UCEC. (A) GINS4 correlation with tumor purity and immune cells infiltration in ESCA, (B) GINS4 correlation with tumor purity and immune cells infiltration in KIRC, (C) GINS4 correlation with tumor purity and immune cells infiltration in LIHC, (D) GINS4 correlation with tumor purity and immune cells infiltration in LUAD, and (E) GINS4 correlation with tumor purity and immune cells infiltration in UCEC. A *P*-value (<0.05) was considered as statistical significant.

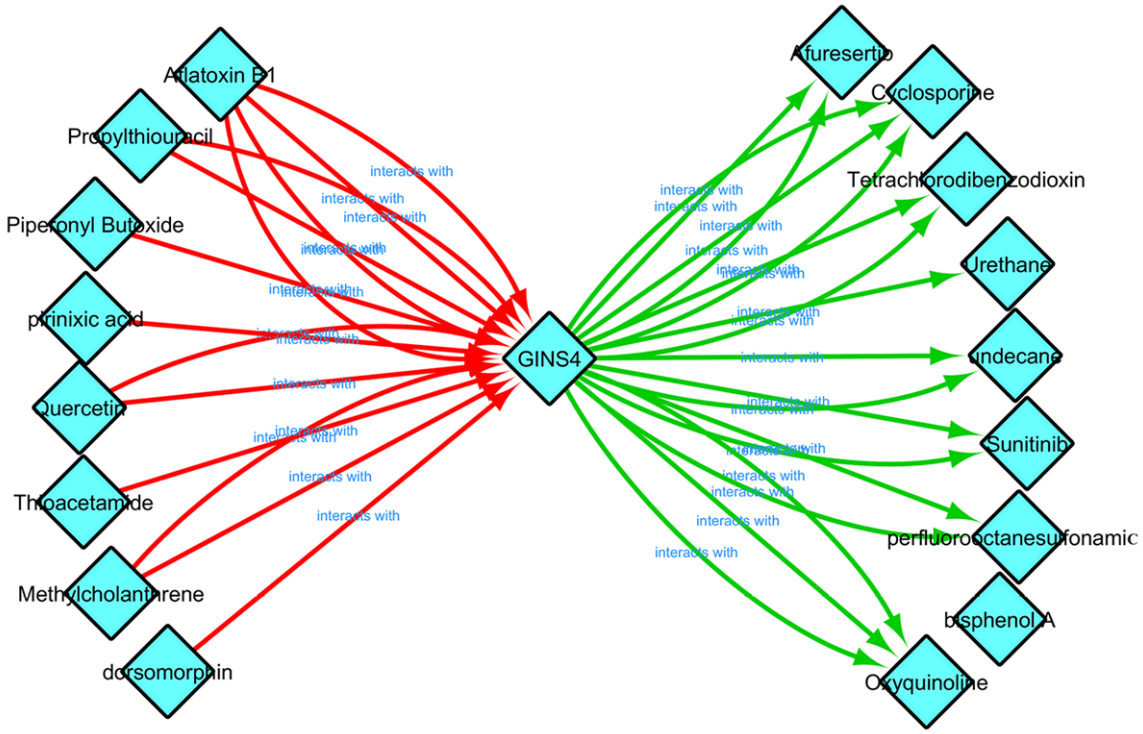


Figure 12. Gene-drug interaction network of the GINS4 and chemotherapeutic drugs. Red arrows: drugs that increase GINS4 expression; green arrows: drugs that decrease GINS4 expression. The numbers of arrows in this network represent the supported numbers of literatures by previous reports.

colorectal cancer (CRC) [19], non-small cell lung cancer (NSCLC) [20], bladder cancer [21], pancreatic cancer [22], and gastric cancer [23]. Nevertheless, the GINS4 effect on different other cancer subtypes is relatively unknown. Via detailed pan-cancer analysis, we analyzed the feasibility of utilizing GINS4 as an ideal diagnostic, prognostic biomarker, and therapeutic target for several cancer subtypes.

In this study, our results revealed that the levels of GINS4 expression in all the 24 major cancers tissue, including LIHC, CHOL, KIRP, ESCA, COAD, CESC, BRCA, and STAD was significantly ($P<0.05$) elevated relative to normal tissues. We further revealed that the up-regulation of GINS4 is generally associated with the reduced OS, RFS durations and advanced metastasis of ESCA, KIRC, LIHC, LUAD, and UCEC patients. Taken together, these findings suggested that GINS4 may play an important role in the initiation, development, and progression of ESCA, KIRC, LIHC, LUAD, and UCEC, therefore, the current investigation focuses on these five cancer subtypes. Following OS, RFS, and metastasis analyses, we further explored the correlation between GINS4 overexpression and different clinicopathological features of ESCA, KIRC, LIHC, LUAD, and UCEC. In view of the results of this analysis, we have also observed a notable overexpression of GINS4 in different clinicopathological features of ESCA, KIRC, LIHC, LUAD, and UCEC including different cancer stages, patient's races, patient's genders, and patients ages as compared to the normal controls.

The GINS4 expression can be influenced by different factors such as promoter methylation, genetic alteration, and CNVs [41]. Therefore, in our study, we utilized MEXPRESS and cBioPortal online resources to analyze the correlation between GNS4 expression and its promoter methylation and genetic alterations in ESCA, KIRC, LIHC, LUAD, and UCEC samples. Our results revealed a significant negative correlation between GINS4 expression and its promoter methylation levels in ESCA, KIRC, LIHC, LUAD, and UCEC patients. We further revealed low percentages (7%, 1.1%, 6%, 6%, and 6%) of the GINS4 genetic alterations and CNVs in ESCA, KIRC, LIHC, LUAD, and UCEC, respective-

ly. Additionally, it was also observed that mutations in GINS4 could change amino acids at different sites of the encoded protein. Taken together these results, we speculated that promoter hypermethylation may have a solid impact on the expression regulation of GINS4 while genetic alterations and CNVs may have very little or possibly no impact on the expression regulation of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC.

Although, a growing number of studies have discovered numerous expression-based biomarkers in ESCA, KIRC, LIHC, LUAD, and UCEC including different genes, such as EGFR, VEGF, ER, E-cadherin, α-catenin, and β-catenin, p53, MAP3K3, and ASPM, in ESCA [42, 43], MYC, VHL, PBRM1, BAP1, PTGS2, ALB, TOP2A, CDK1, AKT1, VEGFA, CASR, MMP9, PTPRC, and EGFR in KIRC [44, 45], FOS, EPHA2, IGFBP3, ID1, DUSP6, MT1G, SNRPD2, MT1H, FGA, SOCS2, LMNB1, ITIH2, KNG1, EGR1, PRR11, FGG, APOA1, AHSG, F2, FOS, DUSP1, APOA2, APOB, and PROC, in LIHC [46, 47], CDH1, PECAM1, SPP1, IL6, THBS1, SNCA, HGF, CAV1, DLC1, and CDH5 in LUAD [48, 49], RNF183, FGFs, FGFRs, ADCY7, and ZBTB7A in UCEC [50-52]. However, best to our knowledge, none of these or any other biomarkers have been generalized so far in ESCA, KIRC, LIHC, LUAD, and UCEC patients of different clinicopathological features. Therefore, the heterogeneity-specific behavior of these markers leads to the high ESCA, KIRC, LIHC, LUAD, and UCEC-associated mortality rates and remains a major therapeutic obstacle for clinicians and doctors. In the current study, a notable overexpression of GINS4 was observed in ESCA, KIRC, LIHC, LUAD, and UCEC patients with different clinicopathological parameters, including different cancer stages, patient's races, genders, and age groups relative to controls. Furthermore, GINS4 prognostic values and promoter methylation levels have also proven its useful significance as a novel potential biomarker of these cancers. Therefore, our study is the first to report a shared clinicopathological features-specific diagnostic and prognostic potential of GINS4 in five different cancers including ESCA, KIRC, LIHC, LUAD, and UCEC, which may open up new therapeutic avenues for these cancer patients.

Furthermore, to know the possible roles of miRNAs and TFs in the dysregulation of GINS4, we predicted the potential miRNAs and TFs of GINS4 using Enrichr from TRRUST 2019 and miRTarBase 2017 sources. Our results revealed the ten most significant miRNAs and TFs that can potentially regulate GINS4 expression, including hsa-miR-193b-3p, hsa-miR-215-5p, hsa-miR-192-5p, hsa-miR-3613-3p, hsa-miR-372-5p, hsa-miR-373-5p, hsa-miR-371b-5p hsa-miR-616-5p, hsa-miR-371a-5p, and hsa-miR-6849-5p miRNAs and E2F4, RBL2, MEN1, E2F1, E2F3, BRCA1, E2F4, OTX2, FOXM1, and FOXO3 TFs. This important piece of information might also help to understand the GINS4 oncogenic roles in more detail.

Next, we have also identified different mutant genes that can alter GINS4 expression via MuTarget. The top 3 mutant genes that we selected in each ESCA, KIRC, LIHC, LUAD, and UCEC, respectively, are TP53, NELL2, and RUNX1 in ESCA, SLC22A4, PTPRZ1, and VARS2 in KIRC, TP53, CSMD3, and CDH10 in LIHC, TP53, KIF19, and RB1 in LUAD, and TP53, TCOF1, and ZNF780A in UCEC. By connecting these mutant genes with GINS4 expression, it is easier for clinicians to identify potential multigene-based therapies for ESCA, KIRC, LIHC, LUAD, and UCEC patients.

Previous studies have revealed that accessing the relationships between immune cells infiltration, tumor purity, and biomarker gene expression is quite valuable for developing the appropriate immunotherapy [53]. Based on the markers gene expression, different studies have explored correlations between tumor purity and marker gene expression to predict the clinical outcomes in different cancers [54]. Moreover, in a recent study, *Nataliya et al.* have accessed the immune cells in the normal and cancerous human HNSC tissues using the CIBERSORT algorithm. In view of their results, it was observed that different immune cells, including neutrophil, B cells, CD4+ cells, and CD8+ T cells were increased in the cancerous tissues relative to normal controls [55]. However, little information is already available regarding tumor purity and the prognostic roles of immune cells infiltration in ESCA, KIRC, LIHC, LUAD, and UCEC patients. Interestingly, in our study, we revealed that GINS4 was noticeably correlated with the tumor purity and immune cells infiltration, which may help clinicians to gain deeper insights into the tumor microenvironment landscape of ESCA, KIRC, LIHC, LUAD, and UCEC. However, we lack direct evidence on

how GINS4 regulates tumor purity and immune cells infiltration in these cancers, therefore, the precise pathways and mechanisms need further studies.

The PPI network of GINS4 has shown that it directly interacts with 23 different other genes, and correlation analysis between GINS4 and these genes expression has revealed a strong positive correlation. Moreover, GINS4 associated genes were found significantly ($P<0.05$) enriched in DNA replication BP, DNA helicase activity MF, MCM complex CC, and different KEGG terms including DNA replication, Cell cycle, Glucagon signaling pathway, Biosynthesis of antibiotics, and Cysteine and methionine metabolism. These results have shown that GINS4 might be involved in a variety of BP, MF, and CC by interacting with its associated genes that participate in caner development. In addition, the two most significant KEGG terms including DNA replication and cell cycle are important processes involved in duplication, growth, and division of the genome [56, 57]. Dysregulation of DNA replication and the cell cycle are one of the most common events in cancer development [58-60]. Moreover, defects in these pathways have also been reported to have an adverse effect on cancer prognosis [61, 62]. Our study suggested that GINS4, via its associated genes, may play a critical role in tumorigenesis by regulating DNA replication and cell cycle processes. Additionally, by querying CTD, we have excavated several available compounds that could enhance or inhibit GINS4 expression, implying their significance in the treatment of ESCA, KIRC, LIHC, LUAD, and UCEC.

Conclusion

This detailed in silico study has effectively uncovered the diagnostic and prognostic roles of GINS4 in ESCA, KIRC, LIHC, LUAD, and UCEC by analyzing its expression and correlations of its expression with different parameters. However, prior to clinical implication, we strongly recommend oncology researchers around the globe to further investigate GINS4 roles on a larger scale in ESCA, KIRC, LIHC, LUAD, and UCEC, and to deeply explore the biology of GINS4 in the immune microenvironment of these cancers, which will aid in successful immunotherapy.

Acknowledgements

The authors extended their appreciation to the Researchers Supporting Project number (RSP-2021/374) King Saud University, Riyadh, Saud Arabia.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Yasir Hameed, Department of Biochemistry and Biotechnology, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan. E-mail: Yasirhameed2011@gmail.com

References

[1] Andre F, Mardis E, Salm M, Soria JC, Siu L and Swanton C. Prioritizing targets for precision cancer medicine. Ann Oncol 2014; 25: 2295-2303.

[2] Schmitz KH, Campbell AM, Stuiver MM, Pinto BM, Schwartz AL, Morris GS, Ligibel JA, Cheville A, Galvão DA, Alfano CM, Patel AV, Hue T, Gerber LH, Sallis R, Gusani NJ, Stout NL, Chan L, Flowers F, Doyle C, Helmrich S, Bain W, Sokolof J, Winters-Stone KM, Campbell KL and Matthews CE. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. CA Cancer J Clin 2019; 69: 468-484.

[3] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[4] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018; 68: 7-30.

[5] Colzani E, Johansson AL, Liljegren A, Foukakis T, Clements M, Adolfsson J, Hall P and Czene K. Time-dependent risk of developing distant metastasis in breast cancer patients according to treatment, age and tumour characteristics. Br J Cancer 2014; 110: 1378-1384.

[6] Laas E, Hamy AS, Michel AS, Panchbhaya N, Faron M, Lam T, Carrez S, Pierga JY, Rouzier R, Lerebours F, Feron JG and Reyal F. Impact of time to local recurrence on the occurrence of metastasis in breast cancer patients treated with neoadjuvant chemotherapy: a random forest survival approach. PLoS One 2019; 14: e0208807.

[7] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 7-33.

[8] Bengtsson A, Andersson R and Ansari D. The actual 5-year survivors of pancreatic ductal adenocarcinoma based on real-world data. Sci Rep 2020; 10: 16425.
[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] Boskovic J, Coloma J, Aparicio T, Zhou M, Robinson CV, Mendez J and Montoya G. Molecular architecture of the human GINS complex. EMBO Rep 2007; 8: 678-684.
[11] Seo YS and Kang YH. The human replicative helicase, the CMG complex, as a target for anti-cancer therapy. Front Mol Biosci 2018; 5: 26.
[12] Seki T, Akita M, Kamimura Y, Muramatsu S, Araki H and Sugino A. GINS is a DNA polymerase ϵ accessory factor during chromosomal DNA replication in budding yeast. J Biol Chem 2006; 281: 21422-21432.
[13] De Falco M, Ferrari E, De Felice M, Rossi M, Hübscher U and Pisani FM. The human GINS complex binds to and specifically stimulates human DNA polymerase α-primase. EMBO Rep 2007; 8: 99-103.
[14] Joshi K, Shah VJ and Maddika S. GINS complex protein Sld5 recruits SIK1 to activate MCM helicase during DNA replication. Cell Signal 2016; 28: 1852-1862.
[15] Labib K and Gambus A. A key role for the GINS complex at DNA replication forks. Trends Cell Biol 2007; 17: 271-278.
[16] Gouge CA and Christensen TW. Drosophila Sld5 is essential for normal cell cycle progression and maintenance of genomic integrity. Biochem Biophys Res Commun 2010; 400: 145-150.
[17] Toda H, Seki N, Kurozumi S, Shinden Y, Yamada Y, Nohata N, Moriya S, Idichi T, Maemura K, Fujii T, Horiguchi J, Kijima Y and Natsugoe S. RNA-sequence-based microRNA expression signature in breast cancer: tumor-suppressive miR-101-5p regulates molecular pathogenesis. Mol Oncol 2020; 14: 426-446.
[18] Xing Z, Luo Z, Yang H, Huang Z and Liang X. Screening and identification of key biomarkers in adrenocortical carcinoma based on bioinformatics analysis. Oncol Lett 2019; 18: 4667-4676.
[19] 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 Krüppel-like factor 4, promotes the growth of colorectal cancer. Cancer Sci 2020; 111: 1203-1217.
[20] 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.
[21] Yamane K, Naito H, Wakabayashi T, Yoshida H, Muramatsu F, Iba T, Kidoya H and Takakura N. Regulation of SLD5 gene expression by miR-370 during acute growth of cancer cells. Sci Rep 2016; 6: 30941.
[22] Bu F, Zhu X, Yi X, Luo C, Lin K, Zhu J, Hu C, Liu Z, Zhao J, Huang C, Zhang W and Huang J. Expression profile of GINS complex predicts the prognosis of pancreatic cancer patients. Onco Targets Ther 2020; 13: 11433.
[23] Zhu Z, Yu Z, Rong Z, Luo Z, Zhang J, Qiu Z and Huang C. The novel GINS4 axis promotes gastric cancer growth and progression by activating Rac1 and CDC42. Theranostics 2019; 9: 8294.
[24] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[25] Chen C, Wang X, Huang S, Wang L, Han L and Yu S. Prognostic roles of Notch receptor mRNA expression in human ovarian cancer. Oncotarget 2017; 8: 32731-32740.
[26] Bartha Á and Győrffy B. TNMplot.com: a web tool for the comparison of gene expression in normal, tumor and metastatic tissues. Int J Mol Sci 2021; 22: 2622.
[27] Park SJ, Yoon BH, Kim SK and Kim SY. GENT2: an updated gene expression database for normal and tumor tissues. BMC Med Genet 2019; 12 Suppl 5: 101.
[28] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.
[29] Cheng WC, Chung IF, Chen CY, Sun HJ, Fen JJ, Tang WC, Chang TY, Wong TT and Wang HW. DriverDB: an exome sequencing database for cancer driver gene identification. Nucleic Acids Res 2014; 42: D1048-D1054.
[30] Koch A, De Meyer T, Jeschke J and Van Criekinge W. MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genom 2015; 16: 636.
[31] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Dis 2012; 2: 401-404.
[32] von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B. STRING: a database of pre-

dicted functional associations between proteins. Nucleic Acids Res 2003; 31: 258-261.
[33] Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007; 8: R183.
[34] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW and Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44: W90-W97.
[35] Nagy Á and Győrffy B. muTarget: a platform linking gene expression changes and mutation status in solid tumors. Int J Cancer 2021; 148: 502-511.
[36] Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B and Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 2020; 48: W509-W514.
[37] Mattingly CJ, Colby GT, Forrest JN and Boyer JL. The comparative toxicogenomics database (CTD). Environ Health Perspect 2003; 111: 793-795.
[38] Karlow JA, Miao B, Xing X, Wang T and Zhang B. Common DNA methylation dynamics in endometriod adenocarcinoma and glioblastoma suggest universal epigenomic alterations in tumorigenesis. Commun Biol 2021; 4: 607.
[39] Sarkar S, Horn G, Moulton K, Oza A, Byler S, Kokolus S and Longacre M. Cancer development, progression, and therapy: an epigenetic overview. Int J Mol Sci 2013; 14: 21087-21113.
[40] Boskovic J, Coloma J, Aparicio T, Zhou M, Robinson CV, Méndez J and Montoya G. Molecular architecture of the human GINS complex. EMBO Rep 2007; 8: 678-684.
[41] Shi X, Radhakrishnan S, Wen J, Chen JY, Chen J, Lam BA, Mills RE, Stranger BE, Lee C and Setlur SR. Association of CNVs with methylation variation. NPJ Genom Med 2020; 5: 41.
[42] Tan C, Qian X, Guan Z, Yang B, Ge Y, Wang F and Cai J. Potential biomarkers for esophageal cancer. Springerplus 2016; 5: 467-467.
[43] Chen X, Huang L, Yang Y, Chen S, Sun J, Ma C, Xie J, Song Y and Yang J. ASPM promotes glioblastoma growth by regulating G1 restriction point progression and Wnt-β-catenin signaling. Aging 2020; 12: 224.
[44] Bailey ST, Smith AM, Kardos J, Wobker SE, Wilson HL, Krishnan B, Saito R, Lee HJ, Zhang J, Eaton SC, Williams LA, Manocha U, Peters DJ, Pan X, Carroll TJ, Felsher DW, Walter V, Zhang Q, Parker JS, Yeh JJ, Moffitt RA, Leung JY and Kim WY. MYC activation cooperates with Vhl and Ink4a/Arf loss to induce clear cell renal cell carcinoma. Nat Commun 2017; 8: 15770.
[45] Han G, Zhao W, Song X, Kwok-Shing Ng P, Karam JA, Jonasch E, Mills GB, Zhao Z, Ding Z and Jia P. Unique protein expression signatures of survival time in kidney renal clear cell carcinoma through a pan-cancer screening. BMC Genom 2017; 18 Suppl 6: 678.
[46] Bai Q, Liu H, Guo H, Lin H, Song X, Jin Y, Liu Y, Guo H, Liang S, Song R, Wang J, Qu Z, Guo H, Jiang H, Liu L and Yang H. Identification of hub genes associated with development and microenvironment of hepatocellular carcinoma by weighted gene co-expression network analysis and differential gene expression analysis. Front Genet 2020; 11: 615308.
[47] Gu Y, Li J, Guo D, Chen B, Liu P, Xiao Y, Yang K, Liu Z and Liu Q. Identification of 13 key genes correlated with progression and prognosis in hepatocellular carcinoma by weighted gene co-expression network Analysis. Front Genet 2020; 11: 153.
[48] Luo X, Feng L, Xu W, Bai X and Wu M. Weighted gene co-expression network analysis of hub genes in lung adenocarcinoma. Evol Bioinform Online 2021; 17: 11769343211009898.
[49] Niu N, Ma X, Liu H, Zhao J, Lu C, Yang F and Qi W. DLC1 inhibits lung adenocarcinoma cell proliferation, migration and invasion via regulating MAPK signaling pathway. Exp Lung Res 2021; 47: 173-182.
[50] Geng R, Zheng Y, Zhao L, Huang X, Qiang R, Zhang R, Guo X and Li R. RNF183 is a prognostic biomarker and correlates with tumor purity, immune infiltrates in uterine corpus endometrial carcinoma. Front Genet 2020; 11: 595733.
[51] Li Y, Wu L, Tao W, Wu D, Ma F and Li N. Expression atlas of FGF and FGFR genes in pancancer uncovered predictive biomarkers for clinical trials of selective FGFR inhibitors. Biomed Res Int 2020; 2020: 5658904.
[52] Zeng Y, Li N, Zheng Z, Chen R, Liu W, Zhu J, Zeng M, Cheng J, Peng M and Hong C. A pan-cancer analysis of the prognostic value and expression of adenylate cyclase 7 (ADCY7) in Human Tumors. Int J Gen Med 2021; 14: 5415-5429.
[53] Hato T, Goyal L, Greten TF, Duda DG and Zhu AX. Immune checkpoint blockade in hepatocellular carcinoma: current progress and future directions. Hepatology 2014; 60: 1776-1782.
[54] Aran D, Sirota M and Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun 2015; 6: 8971.
[55] Liang B, Tao Y and Wang T. Profiles of immune cell infiltration in head and neck squamous

carcinoma. Bioscience Rep 2020; 40: BSR20192724.
[56] Poon RY. Cell cycle control: a system of interlinking oscillators. Cell Cycle Oscillators 2016; 3-19.
[57] Alberts B. DNA replication and recombination. Nature 2003; 421: 431-435.
[58] Preston BD, Albertson TM and Herr AJ. DNA replication fidelity and cancer. Semin Cancer Biol 2010; 20: 281-293.
[59] Visconti R, Della Monica R and Grieco D. Cell cycle checkpoint in cancer: a therapeutically targetable double-edged sword. J Exp Clin Cancer Res 2016; 35: 153.
[60] Otto T and Sicinski P. Cell cycle proteins as promising targets in cancer therapy. Nat Rev 2017; 17: 93-115.
[61] Pillaire MJ, Selves J, Gordien K, Gouraud PA, Gentil C, Danjoux M, Do C, Negre V, Bieth A, Guimbaud R, Trouche D, Pasero P, Méchali M, Hoffmann JS and Cazaux C. A 'DNA replication' signature of progression and negative outcome in colorectal cancer. Oncogene 2010; 29: 876-887.
[62] Bai J, Li Y and Zhang G. Cell cycle regulation and anticancer drug discovery. Cancer Biol Med 2017; 14: 348-362.

Supplementary Table 1. Clinicopathological parameters-based classification of ESCA and KIRC patients

Clinicopathological features of the ESCA cohort					
Sr. No	Clinicopathological Feature	No. Samples	Total no. of ESCA samples	No. Excluded Samples with Missing Information	Total No. of Included Samples
	Cancer stage distribution				
1	Stage 1	13			
	Stage 2	78		29	155
	Stage 3	55			
	Stage 4	09			
	Geographical distribution				
2	Caucasian	113			
	African-American	05	184	20	164
	Asian	46			
	Gender distribution				
3	Male	157		01	183
	Female	26			
	Age distribution				
4	21-40 years	03			
	41-60 years	89		01	183
	61-80 years	76			
	81-100 years	15			
Clinicopathological features of the KIRC cohort					
Sr. No	Clinicopathological Feature	No. Samples	Total no. of KIRC samples	No. Excluded Samples with Missing Information	Total No. of Included Samples
	Cancer stage distribution				
1	Stage 1	267			
	Stage 2	57		02	531
	Stage 3	123			
	Stage 4	84			
	Geographical distribution				
2	Caucasian	462			
	African-American	56		07	526
	Asian	08	533		
	Gender distribution				
3	Male	345		00	533
	Female	188			
	Age distribution				
4	21-40 years	26			
	41-60 years	238		00	533
	61-80 years	246			
	81-100 years	23			

Supplementary Table 2. Clinicopathological parameters-based classification of LIHC and LUAD patients

Clinicopathological features of the LIHC cohort					
Sr. No	Clinicopathological Feature	No. Samples	Total no. of LIHC samples	No. Excluded Samples with Missing Information	Total No. of Included Samples
	Cancer stage distribution				
1	Stage 1	168			
	Stage 2	84		31	340
	Stage 3	82			
	Stage 4	06			
	Geographical distribution				
2	Caucasian	177			
	African-American	17		21	350
	Asian	156	371		
	Gender distribution				
3	Male	245		09	362
	Female	117			
	Age distribution				
4	21-40 years	27			
	41-60 years	140		13	358
	61-80 years	181			
	81-100 years	10			
Clinicopathological features of the LUAD cohort					
Sr. No	Clinicopathological Feature	No. Samples	Total no. of LUAD samples	No. Excluded Samples with Missing Information	Total No. of Included Samples
	Cancer stage distribution				
1	Stage 1	277			
	Stage 2	125		0	515
	Stage 3	85			
	Stage 4	28			
	Geographical distribution				
2	Caucasian	387			
	African-American	51	515	69	446
	Asian	08			
	Gender distribution				
3	Male	238		01	514
	Female	276			
	Age distribution				
4	21-40 years	12			
	41-60 years	90		277	283
	61-80 years	149			
	81-100 years	32			

Supplementary Table 3. Clinicopathological parameters-based classification of UCEC patients

Clinicopathological features of the UCEC cohort					
Sr. No	Clinicopathological Feature	No. Samples	Total no. of UCEC samples	No. Excluded Samples with Missing Information	Total No. of Included Samples
	Cancer stage distribution				
1	Stage 1	341			
	Stage 2	52		0	546
	Stage 3	124			
	Stage 4	29			
	Geographical distribution				
2	Caucasian	374			
	African-American	107		45	501
	Asian	20	546		
	Gender distribution				
3	Male	297		60	486
	Female	189			
	Age distribution				
4	21-40 years	18			
	41-60 years	189		02	544
	61-80 years	292			
	81-100 years	45			

Supplementary Table 4. Detail of Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis

Detail of BP analysis					
ID	Description	Enriched Genes	Gene count	*P*-value	FDR
GO:0006260	DNA replication	RECQL4, POLA1, TIPIN, MCM7, GINS3, GINS4, MCM3, MCM5, SSRP1, MCM2	10	3.290495373208641E-14	3.652449864261592E-12
GO:0006270	DNA replication initiation	POLA1, MCM7, GINS4, MCM3, MCM5, MCM2	6	2.0667364613691976E-10	1.1470387360599047E-8
GO:0000082	G1/S transition of mitotic cell cycle	POLA1, MCM7, MCM3, MCM5, MCM2	5	4.63685843541141E-6	1.715637621102222E-4
GO:0032508	DNA duplex unwinding	RECQL4, GINS4, MCM3, MCM5	4	1.5795488676361072E-5	4.3832481076901977E-4
GO:0006271	DNA strand elongation involved in DNA replication	POLA1, GINS3, GINS4	3	1.2624816699795423E-4	0.0028027093073545836
Detail of MF analysis					
ID	Description	Enriched Genes	Gene count	*P*-value	FDR
GO:0003678	DNA helicase activity	MCM7, MCM3, MCM5, MCM2	4	2.0316708521633356E-6	1.1783690942547346E-4
GO:0005524	ATP binding	RECQL4, MCM7, KIF16B, MCM3, SIK1, MCM5, PAICS, MCM2	8	5.583934468849316E-4	0.016193409959663018
GO:0005515	protein binding	WDHD1, TIPIN, AHCYL1, MCM7, GINS4, SSRP1, PAICS, DUSP13, RECQL4, LDHB, POLA1, MCM3, SIK1, MCM5, NCDN, PPIL3, MCM2	17	0.001150587932684919	0.022244700031908433
GO:0003677	DNA binding	WDHD1, POLA1, TIPIN, MCM7, MCM3, SSRP1, MCM2	7	0.006126471340372367	0.08883383443539933
Detail of CC analysis					
ID	Description	Enriched Genes	Gene count	*P*-value	FDR
GO:0042555	MCM complex	MCM7, MCM3, MCM5, MCM2	4	9.45478752778547E-8	385.69312169312167
GO:0000784	nuclear chromosome, telomeric region	MCM7, MCM3, MCM5, MCM2	4	3.699886106115502E-4	0.006474800685702129
GO:0005634	nucleus	TIPIN, PDHA1, MCM7, GINS3, GINS4, DONSON, SSRP1, RECQL4, POLA1, MCM3, SIK1, MCM5, NCDN, MCM2	14	0.0011508043934053166	0.01319854923256369
GO:0005654	nucleoplasm	WDHD1, POLA1, TIPIN, MCM7, GINS3, GINS4, MCM3, MCM5, SSRP1, MCM2	10	0.0015084056265787073	0.01319854923256369
GO:0005737	cytoplasm	WDHD1, TIPIN, AHCYL1, MCM7, GINS4, SSRP1, PAICS, DUSP13, RECQL4, LDHB, POLA1, SIK1, MCM2	13	0.003367360728792998	0.023571525101550986
GO:0005658	alpha DNA polymerase:primase complex	POLA1, MCM3	3	0.006567582317544384	0.03283791158772192
Detail of KEGG analysis					
ID	Description	Enriched Genes	Gene count	*P*-value	FDR
hsa03030	DNA replication	POLA1, MCM7, MCM3, MCM5, MCM2	5	<0.05	95.5416
hsa04110	Cell cycle	MCM7, MCM3, MCM5, MCM2	4	<0.05	0.0066
hsa04922	Glucagon signaling pathway	LDHB, PDHA1, SIK1	3	<0.05	20.8454
hsa01130	Biosynthesis of antibiotics	LDHB, PDHA1, PAICS	3	<0.05	9.7344
hsa00270	Cysteine and methionine metabolism	LDHB, AHCYL1	2	<0.05	36.2052