

Prognostic significance of EIF4G1 in patients with pancreatic ductal adenocarcinoma

Tae Sik Goh¹, Mihyang Ha², Jung Sub Lee¹, Dae Cheon Jeong³, Eun Sang Jung⁴, Myoung-Eun Han², Yun Hak Kim^{5, 6*}, Sae-Ock Oh^{2, *}

¹Department of Orthopaedic Surgery and Biomedical Research Institute, Pusan National University Hospital, Busan, Republic of Korea.

²Department of Anatomy, School of Medicine, Pusan National University, Yangsan, Republic of Korea.

³Deloitte Analytics Group, Deloitte Consulting LLC, Republic of Korea

⁴Department of Bioenvironmental Energy, College of Life & Resources Science, Pusan National University, Miryang, Republic of Korea.

⁵Department of Anatomy, Department of Biomedical Informatics, School of Medicine, Pusan National University, Yangsan, Republic of Korea.

⁶Biomedical Research Institute, Pusan National University, Busan, Republic of Korea

Tae Sik Goh and Mihyang Ha contributed equally to this work

Short title: EIF4G1 is a prognostic gene for PDAC

*Correspondence

Sae-Ock Oh, MD, PhD

Department of Anatomy, School of Medicine, Pusan National University, Yangsan, 50612, Republic of Korea

Tel: 82-51-510-8045

Fax: 82-51-510-8049

E-mail: hedgehog@pusan.ac.kr

Yun Hak Kim, MD, PhD

Department of Anatomy and Department of Biomedical Informatics, School of Medicine,
Pusan National University, 49 Busandaehak-ro, Yangsan, 50612, Republic of Korea

Tel: 82-51-510-8091

Fax: 82-51-510-8049

E-mail: yunhak10510@pusan.ac.kr

Abstract

Background: The advances of genomics have greatly improved the survival rate cancer patients. However, due to genetic heterogeneity, pancreatic ductal adenocarcinoma (PDAC) is still difficult to diagnose early and the survival rate is extremely low. Therefore, we identified biomarkers that predict the prognosis of PDAC patients by using independent cohort data.

Methods: To develop a novel prognostic biomarker, we used gene expression and clinical data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO). In Kaplan-Meier survival curve using median values of genes as cut off, the only statistically significant gene in the three cohorts was EIF4G1. We analyzed prognostic significance of EIF4G1 using the time-dependent area under the curve (AUC) of the Uno's C-index, the AUC value of the receiver operating characteristics (ROC) at 3 years, and multivariate cox analysis. Also, we compare EIF4G1 levels between tumor and matched non-tumor.

Results: EIF4G1 is the only prognostic gene patients with PDAC which was selected by Kaplan-Meier survival analysis. Kaplan-Meier survival analysis showed that high expression of EIF4G1 was associated with poor prognosis of PDAC with good discriminative ability in 3 independent cohorts. Risk stratifying ability of EIF4G1 was demonstrated by analyzing C-indices and AUC values. Multivariate cox regression analysis confirmed its prognostic significance. EIF4G1 expression was significantly higher in the PDAC tissues than in the matched normal tissues.

Conclusions: Herein, the novel prognostic biomarker EIF4G1 could be used as prognostic maker for PDAC and determining suitable treatment options.

Keyword: EIF4G1; pancreatic ductal adenocarcinoma; prognosis; GEO; TCGA

Introduction

Pancreatic cancer has a very poor prognosis and is difficult to detect early [1]. Of these, 90% are pancreatic ductal adenocarcinoma (PDAC) [2]. Only surgical treatment is known to be effective in patients with PDAC. Surgical resection is performed only in 10 to 20% of cases [3]. Because most of the cases are belong to advanced stage at the time of diagnosis [4, 5]. Moreover, the 5-year survival rate is less than 10% because most relapse or metastasis even if they receive complete surgical resection [6]. Therefore, prognostic biomarkers that could predict the prognosis more accurately and help early diagnosis are indispensable.

As the importance of precision medicine has been emphasized recently, genomic research is not only activated, but its use is expanding from bench to bedside, the actual diagnosis and treatment process [7, 8]. Through these efforts, public databases including The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and etc. related to patients with various cancer types and their genomes have been made, and researches and papers have been actively made through them. Using data sets of these databases and our novel statistical methods, we have been reporting a single gene or set of genes in a specific cancer that can predict the prognosis [9, 10].

In the present study, we investigated whether a gene could be a biomarker to predict the prognosis of patients with PDAC based using three cohorts from TCGA and GEO. Finally, we found the only gene that could predict the prognosis in PDAC. It was also demonstrated that the prognosis could be stratified according to the expression level of the gene.

Results

To select the gene which could predict prognosis of PDAC using public databases, the clinical and genetic information of 316 patients with PDAC from 3 independent cohorts (TCGA, n = 172; GSE21501, n = 102 and GSE28735, n = 42) were downloaded and analyzed. Patient information used in the present study is detailed in Table 1. The patients in TCGA were almost diagnosed at low stage, but the patients in GSE21501 were diagnosed almost at high stage.

Prognostic performance of EIF4G1 in PDAC

We obtained the median value of gene expression in all genes in each cohort. Each cohort was divided into 2 groups based on the median value of each gene, and the survival of the 2 groups was compared using Kaplan-Meier survival analysis and statistically significant genes were extracted for each cohort. Among the commonly extracted genes in all 3 cohorts, only EIF4G1 has prognostic significance at the same time in the three cohorts.

We analyzed the Kaplan-Meier curves for survival according to the expression level of EIF4G1 to demonstrate the prognostic performance of EIF4G1 in PDAC. Intriguingly, lower expression of EIF4G1 was significantly associated with good prognosis in all 3 cohorts (TCGA, $P=0.00053$; GSE21501, $P=0.0077$ and GSE28735, $P=0.041$) (Figure 2A, 2D and 2G). Results of univariate analysis of overall survival in each cohort, the expression level of EIF4G1 was statistically significant in all cohorts (Table 2). Furthermore, the multivariate analysis of the TCGA, GSE21501 demonstrated significant prognostic performance of EIF4G1 in PDCA which was consistent with above survival analysis (TCGA, $P=0.00132$, GSE21501, $P=0.025$, Table 2). Hazard ratio (HR) of EIF4G1 is particularly high when compared to the other variables (Table 2). In addition, age in TCGA was significant variable which could stratify prognosis ($P=0.02121$, Table 2).

Biomarker ability of EIF4G1 in PDAC

We compared the Uno's C-index values and Area Under the Curve (AUC) values at 3 years of the expression level of EIF4G1 with other variables such as tumor staging, sex, age which could be obtained from clinical information of each cohorts to examine the ability of EIF4G1 as a biomarker. EIF4G1 showed most high C-index values for 3 years in 2 independent cohorts (TCGA, 0.686; and GSE21501, 0.526; Figure 2B and 2E). Consistent to the Uno's C-index results, the 3 years AUC value was slightly less than 0.6 in GSE21501 (0.594; Figure 2F), and nearly 0.7 in other 2 cohorts (TCGA, 0.696; and GSE28735, 0.655; Figure 2C and 2I).

Overexpression of EIF4G1 in PDAC

In order to confirm that EIF4G1 could predict prognosis as a tumor biomarker, the expression level of EIF4G1 was analyzed in matched tumor tissues and adjacent non-tumor normal tissues using GSE28735. Expression level of EIF4G1 was significantly higher in the PDAC tissues than in the matched normal tissues (Figure 3).

Discussion

EIF4G1 encodes a scaffold protein upon which ribosomes and the eukaryotic initiation factor (EIF) 4F complex assemble [11, 12]. EIF4F complex regulates the key step of initiation in translation of almost all genes in eukaryotes [12, 13]. Increased expression of EIF4G1 has been found in inflammatory breast cancer, lung cancer, hypopharyngeal cancer and nasopharyngeal cancer which are consistent with our findings [11, 14]. Also higher expression of EIF4G1 is associated with shorter overall survival in various cancers [11, 13-16]. EIF4G1 may play a tumorigenic role through enhancing translation of IRES-containing p120 mRNA, which contributed to survival of breast tumor cell [13]. However, the process by which EIF4G1 has been identified in tumorigenesis has not been fully elucidated in many cancers, including PDAC.

Despite the development of precision medicine, the only prognostic/diagnostic marker for PDAC is CA19-9 [17]. Although CA19-9 has been widely used, it is not useful for screening due to its low positive predictive value (<1%). And increased false positivity of CA19-9 have been shown in the presence of obstructive jaundices (10-60%) [18]. To complement the disadvantages of CA19-9, we investigated novel biomarker in patients with PDAC using 3 independent cohorts from TCGA and GEO databases. As described in Table 1, the patients' information of TCGA and GSE21501 is quite different. EIF4G1, which is associated with survival in both cohorts, is likely to be a universal prognostic predictor applied to all stage patients. Additionally, other clinical variables except for age were not statistically related to survival. These results may be due to the fact that the patient composition each cohort (TCGA, GSE21501) is biased toward one side.

We demonstrated the prognostic significance of EIF4G1 in patients with PDAC using public databases. EIF4G1 known to contribute not only to tumorigenesis, but also to tumor progression in several cancers. EIF4G1 is more expressed in cancer tissues and associated with

poor prognosis as expression increases. We suggest that EIF4G1 could act as a prognostic biomarker to help determine the precise treatment strategy for PDAC.

Materials and Methods

Patients' data and study design

The RNA-seq and microarray data and clinical data of PDAC were downloaded from TCGA [19, 20], GSE21501 [21], and GSE28735 [22] in Mar 2018. Patients lacking clinical information were excluded. We identified the prognostic significances of mRNAs in 3 independent cohorts. After that we performed paired T-test whether the statistically significant genes in all cohorts were increased in cancer tissues compared with normal tissues using GSE28735 cohort. The overall process is described in Figure 1. These processes were performed in R software version 3.5.0 (The R Foundation for Statistical Computing, 2018) with the “*cgdsr*” and “*GEOquery*” R packages.

Statistical analysis

Kaplan-Meier survival curves were used to identify the discriminatory power of EIF4G1. We determined the optimal cut-off value of survival curve as described before [4, 23, 24]. Furthermore, we used 2 methods to evaluate biomarker performance: (1) Uno's C-index in the time-dependent area under the curve (AUC) analysis and (2) AUC values in receiver operating characteristic (ROC) curves at the 3-year mark as we described previously [25, 26]. These values were calculated using the R packages “*survival*” and “*survAUC*”. The paired T-test was performed to analyse the EIF4G1 expression values between matched tumor and adjacent non-tumor tissue samples using the “*coin*” package. We used uni- and multi-variate Cox regression analyses to compare the effect of EIF4G1 on prognosis along with several clinical variables.

Acknowledgements

This study was supported by Biomedical Research Institute Grant (2018B032), Pusan National University Hospital.

Author Contributions

TSG and MH: formal analysis and writing original draft; JSL, ESJ, and MEH: data curation; DCJ: formal analysis; SOO and YHK: conceptualization, project administration, supervision.

Conflicts of Interest

The authors declare no potential conflict of interests.

References

1. Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Thun, M. J., Cancer statistics, 2009. *CA Cancer J Clin* **2009**, *59*, (4), 225-49.
2. Kwon, H. J.; Kim, S. G., Surgical treatment for advanced pancreatic cancer. *Korean J Hepatobiliary Pancreat Surg* **2012**, *16*, (3), 89-92.
3. Li, C. P.; Chao, Y.; Chi, K. H.; Chan, W. K.; Teng, H. C.; Lee, R. C.; Chang, F. Y.; Lee, S. D.; Yen, S. H., Concurrent chemoradiotherapy treatment of locally advanced pancreatic cancer: gemcitabine versus 5-fluorouracil, a randomized controlled study. *Int J Radiat Oncol Biol Phys* **2003**, *57*, (1), 98-104.
4. Kim, Y. H.; Jeong, D. C.; Pak, K.; Goh, T. S.; Lee, C. S.; Han, M. E.; Kim, J. Y.; Liangwen, L.; Kim, C. D.; Jang, J. Y.; Cha, W.; Oh, S. O., Gene network inherent in genomic big data improves the accuracy of prognostic prediction for cancer patients. *Oncotarget* **2017**, *8*, (44), 77515-77526.
5. Eto, S.; Ishikawa, M.; Asanoma, M.; Tashiro, Y.; Matsuyama, K.; Oshio, T., A long-term survival case of advanced biliary cancer with repeated resection due to recurrence in the pancreaticogastrostomy site after pancreaticoduodenectomy. *Ann Hepatobiliary Pancreat Surg* **2018**, *22*, (2), 173-177.
6. Zammarchi, F.; Morelli, M.; Menicagli, M.; Di Cristofano, C.; Zavaglia, K.; Paolucci, A.; Campani, D.; Aretini, P.; Boggi, U.; Mosca, F.; Cavazzana, A.; Cartegni, L.; Bevilacqua, G.; Mazzanti, C. M., KLF4 is a novel candidate tumor suppressor gene in pancreatic ductal carcinoma. *Am J Pathol* **2011**, *178*, (1), 361-72.
7. Friedman, A. A.; Letai, A.; Fisher, D. E.; Flaherty, K. T., Precision medicine for cancer with next-generation functional diagnostics. *Nat Rev Cancer* **2015**, *15*, (12), 747-56.
8. Arnedos, M.; Vicier, C.; Loi, S.; Lefebvre, C.; Michiels, S.; Bonnefoi, H.; Andre, F., Precision medicine for metastatic breast cancer--limitations and solutions. *Nat Rev Clin*

Oncol **2015**, *12*, (12), 693-704.

9. Kim, Y. H.; Jeong, D. C.; Pak, K.; Han, M. E.; Kim, J. Y.; Liangwen, L.; Kim, H. J.; Kim, T. W.; Kim, T. H.; Hyun, D. W.; Oh, S. O., SLC2A2 (GLUT2) as a novel prognostic factor for hepatocellular carcinoma. *Oncotarget* **2017**, *8*, (40), 68381-68392.
10. Suh, S.; Kim, Y. H.; Goh, T. S.; Jeong, D. C.; Lee, C. S.; Jang, J. Y.; Cha, W.; Han, M. E.; Kim, S. J.; Kim, I. J.; Pak, K., mRNA Expression of SLC5A5 and SLC2A Family Genes in Papillary Thyroid Cancer: An Analysis of The Cancer Genome Atlas. *Yonsei Med J* **2018**, *59*, (6), 746-753.
11. Tu, L.; Liu, Z.; He, X.; He, Y.; Yang, H.; Jiang, Q.; Xie, S.; Xiao, G.; Li, X.; Yao, K.; Fang, W., Over-expression of eukaryotic translation initiation factor 4 gamma 1 correlates with tumor progression and poor prognosis in nasopharyngeal carcinoma. *Mol Cancer* **2010**, *9*, 78.
12. Badura, M.; Braunstein, S.; Zavadil, J.; Schneider, R. J., DNA damage and eIF4G1 in breast cancer cells reprogram translation for survival and DNA repair mRNAs. *Proc Natl Acad Sci U S A* **2012**, *109*, (46), 18767-72.
13. Silvera, D.; Arju, R.; Darvishian, F.; Levine, P. H.; Zolfaghari, L.; Goldberg, J.; Hochman, T.; Formenti, S. C.; Schneider, R. J., Essential role for eIF4G1 overexpression in the pathogenesis of inflammatory breast cancer. *Nat Cell Biol* **2009**, *11*, (7), 903-8.
14. Cao, Y.; Wei, M.; Li, B.; Liu, Y.; Lu, Y.; Tang, Z.; Lu, T.; Yin, Y.; Qin, Z.; Xu, Z., Functional role of eukaryotic translation initiation factor 4 gamma 1 (EIF4G1) in NSCLC. *Oncotarget* **2016**, *7*, (17), 24242-51.
15. Jaiswal, P. K.; Koul, S.; Shanmugam, P. S. T.; Koul, H. K., Eukaryotic Translation Initiation Factor 4 Gamma 1 (eIF4G1) is upregulated during Prostate cancer progression and modulates cell growth and metastasis. *Sci Rep* **2018**, *8*, (1), 7459.

16. Liang, S.; Zhou, Y.; Chen, Y.; Ke, G.; Wen, H.; Wu, X., Decreased expression of EIF4A1 after preoperative brachytherapy predicts better tumor-specific survival in cervical cancer. *Int J Gynecol Cancer* **2014**, 24, (5), 908-15.
17. Goonetilleke, K. S.; Siriwardena, A. K., Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur J Surg Oncol* **2007**, 33, (3), 266-70.
18. Ballehaninna, U. K.; Chamberlain, R. S., The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J Gastrointest Oncol* **2012**, 3, (2), 105-19.
19. Cancer Genome Atlas Research, N.; Weinstein, J. N.; Collisson, E. A.; Mills, G. B.; Shaw, K. R.; Ozenberger, B. A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J. M., The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **2013**, 45, (10), 1113-20.
20. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N., The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2012**, 2, (5), 401-4.
21. Stratford, J. K.; Bentrem, D. J.; Anderson, J. M.; Fan, C.; Volmar, K. A.; Marron, J. S.; Routh, E. D.; Caskey, L. S.; Samuel, J. C.; Der, C. J.; Thorne, L. B.; Calvo, B. F.; Kim, H. J.; Talamonti, M. S.; Iacobuzio-Donahue, C. A.; Hollingsworth, M. A.; Perou, C. M.; Yeh, J. J., A six-gene signature predicts survival of patients with localized pancreatic ductal adenocarcinoma. *PLoS Med* **2010**, 7, (7), e1000307.
22. Zhang, G.; Schetter, A.; He, P.; Funamizu, N.; Gaedcke, J.; Ghadimi, B. M.; Ried, T.; Hassan, R.; Yfantis, H. G.; Lee, D. H.; Lacy, C.; Maitra, A.; Hanna, N.; Alexander, H. R.; Hussain, S. P., DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity

and predicts clinical outcome in pancreatic ductal adenocarcinoma. *PLoS One* **2012**, 7, (2), e31507.

23. Ha, M.; Han, M. E.; Kim, J. Y.; Jeong, D. C.; Oh, S. O.; Kim, Y. H., Prognostic role of TPD52 in acute myeloid leukemia: A retrospective multicohort analysis. *J Cell Biochem* **2018**.

24. Han, M. E.; Kim, J. Y.; Kim, G. H.; Park, S. Y.; Kim, Y. H.; Oh, S. O., SAC3D1: a novel prognostic marker in hepatocellular carcinoma. *Sci Rep* **2018**, 8, (1), 15608.

25. Cho, S. H.; Pak, K.; Jeong, D. C.; Han, M. E.; Oh, S. O.; Kim, Y. H., The AP2M1 gene expression is a promising biomarker for predicting survival of patients with hepatocellular carcinoma. *J Cell Biochem* **2018**.

26. Goh, T. S.; Lee, J. S.; Il Kim, J.; Park, Y. G.; Pak, K.; Jeong, D. C.; Oh, S. O.; Kim, Y. H., Prognostic scoring system for osteosarcoma using network-regularized high-dimensional Cox-regression analysis and potential therapeutic targets. *J Cell Physiol* **2019**.

Figures

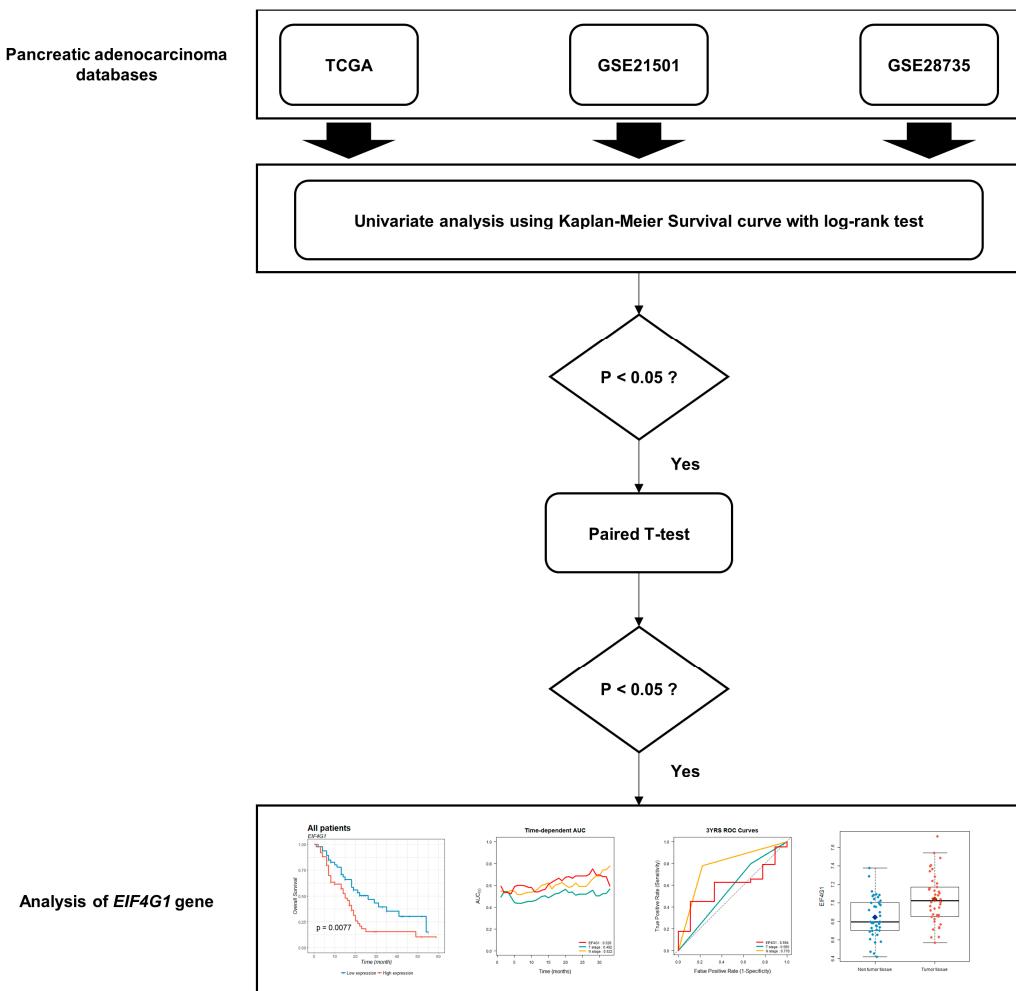


Figure 1. Study protocol of this study.

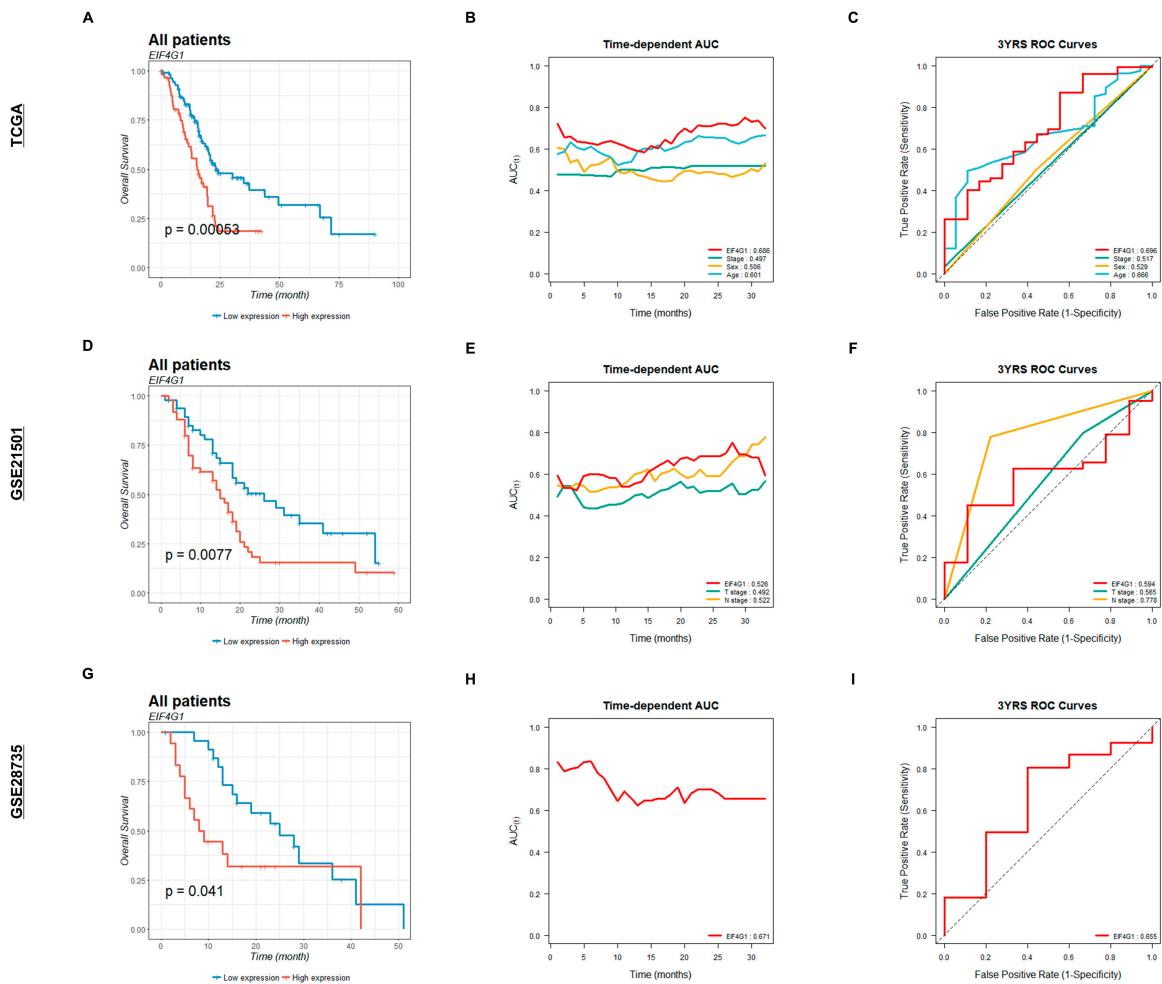


Figure 2. Comparison of EIF4G1 gene expression between matched non-tumor (blue) and tumor tissue (red) using GSE28735.

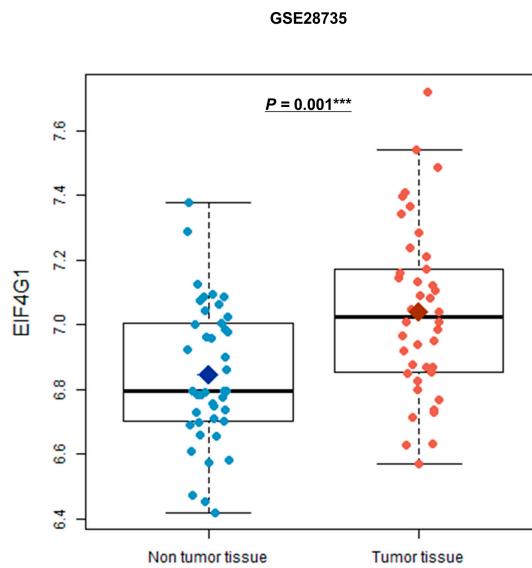


Figure 3. Survival analyses of EIF4G1 in 3 independent cohorts. Kaplan-Meier estimates of all patients in TCGA (A), GSE21501 (D), and GSE28735 (G) according to EIF4G1 expression. Time-dependent Area Under the Curve (AUC) of EIF4G1 with clinical variables in TCGA (red, EIF4G1; green, Stage; yellow, Sex; blue, Age) (B), GSE21501 (red, EIF4G1; green, T stage; yellow, N stage) (E), and GSE28735 (red, EIF4G1) (H). 3 years receiver operating characteristics (ROC) of EIF4G1 with clinical variables in TCGA (red, EIF4G1; green, Stage; yellow, Sex; blue, Age) (C), GSE21501 (red, EIF4G1; green, T stage; yellow, N stage) (F), and GSE28735 (red, EIF4G1) (I).

Tables

		Group	TCGA	GSE21501	GSE28735
<i>EIF4G1</i>	All patients		172	102	42
	High expression (event)		58 (40)	54 (39)	24 (16)
	Low expression (event)		114 (52)	48 (27)	18 (13)
Patients' information	Male		94	-	-
	Female		78	-	-
	Stage I & II		164	-	-
	Stage III & IV		8	-	-
	T1 & T2		-	18	-
	T3 & T4		-	80	-
	N0		-	28	-
	N1		-	73	-

Table 1. Patients' information used in current research in the TCGA, GSE21501 and GSE28735 cohorts.

Variables	Univariate Cox analysis				Multivariate Cox analysis			
	P value	Hazard Ratio	95% Confidence Interval		P value	Hazard Ratio	95% Confidence Interval	
TCGA								
EIF4G1	< 0.0001	2.072	1.359	3.157	0.00132	1.9974	1.3093	3.047
Stage	0.716	0.8072	0.2545	2.561	0.7005	0.7968	0.2504	2.535
Sex	0.39	0.8354	0.5543	1.259	0.2903	0.8010	0.5308	1.209
Age	0.0136	1.0261	1.005	1.047	0.02121	1.024	1.0036	1.045
GSE21501								
EIF4G1	0.00896	1.964	1.184	3.258	0.025	1.8076	1.077	3.034
T stage (T1 & T2 vs T3 & T4)	0.73	0.8977	0.4862	1.657	0.6047	0.8424	0.4400	1.613
N stage (N0 vs N1)	0.0425	1.8399	1.021	3.316	0.0709	1.7773	0.9523	3.317
GSE28735								
EIF4G1	0.045*	2.185	1.018	4.691				

Table 2. Univariate and multivariate analysis of overall survival in each cohort.