Mathematical Biosciences and Engineering

MBE, 18 (5): 6941–6960.
DOI: 10.3934/mbe.2021345
Received: 14 June 2021
Accepted: 30 July 2021
Published: 20 August 2021

http://www.aimspress.com/journal/MBE

Research article

Upregulation of BIRC5 plays essential role in esophageal squamous cell carcinoma

Zu-Xuan Chen [1,#], Guo-Sheng Li [1,#], Li-Hua Yang [2], He-Chuan Liu [1], Guang-Mei Qin [1], Lang Shen [1], Wei-Ying He [1], Ting-Qing Gan [1,*] and Jian-Jun Li [3,*]

[1] Department of Medical Oncology, Second Affiliated Hospital of Guangxi Medical University, 166 DaxueXi Road, Nanning, Guangxi Zhuang Autonomous Region, 530021, P. R. China
[2] Department of Medical Oncology, First Affiliated Hospital of Guangxi Medical University, 6 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region, 530021, P. R. China
[3] Department of General Surgery, Second Affiliated Hospital of Guangxi Medical University, 166 DaxueXi Road, Nanning, Guangxi Zhuang Autonomous Region, 530021, P. R. China

[#]Contributed equally as co-first authors; [*]Contributed equally as co-corresponding authors

* **Correspondence:** Email: gantingqing_gxmu@163.com, lijianjunmail@163.com; Tel: 0771-3277224, Fax: 0771-3277277.

Abstract: *Background*: Esophageal squamous cell carcinoma (ESCC) is one of the most common cancers in the world, the detection and prognosis of which are still unsatisfactory. Thus, it is essential to explore the factors that may identify ESCC and evaluate the prognosis of ESCC patients. *Results*: Both protein and mRNA expression levels of BIRC5 are upregulated in ESCC group rather than non-ESCC group (standardized mean difference > 0). *BIRC5* mRNA expression is related to the age, tumor location, lymph node stage and clinical stage of ESCC patients (p <0.05). BIRC5 expression makes it feasible to distinguish ESCC from non-ESCC (area under the curve > 0.9), and its high expression is related to poor prognosis of ESCC patients (restrictive survival time difference = –0.036, p < 0.05). BIRC5 may play an important role in ESCC by influencing the cell cycle pathway, and *CDK1*, *MAD2L* and *CDC20* may be the hub genes of this pathway. The transcription factors—*MAZ* and *TFPD1* —are likely to regulate the transcription of *BIRC5*, which may be one of the factors for the high expression of BIRC5 in ESCC. *Conclusions*: The current study shows that upregulation of BIRC5 may have essential clinical value in ESCC, and contributes to the understanding of the pathogenesis of ESCC.

Keywords: esophageal squamous cell carcinoma; baculoviral inhibitor of apoptosis repeat containing 5; microarray; RNA sequencing; prognosis

1. Introduction

Esophageal cancer is the eighth most common type of cancer and the sixth most deadly in terms of mortality rates [1,2]. Worldwide, there are about 456,000 newly diagnosed cases of esophageal cancer every year [3]. Among several histological types of esophageal cancer, esophageal squamous cell carcinoma (ESCC) accounts for about 90% of all cases [3] . Currently, the main clinical treatment methods for ESCC include endoscopic resection, surgery, radiotherapy, and chemotherapy [4,5]. Early diagnosis and treatment of ESCC are beneficial for patients in a 5-year survival rate [6]. However, although it is easy to perform endoscopic examination in the esophagus, the early symptoms of ESCC are not obvious, causing a delay in diagnosis [7,8] . Quite a few ESCC patients were diagnosed at advanced clinical stages [9]; and worse still, ESCC patients with advanced clinical stages tend to have dropped significantly 5-year survival rate [6]. Therefore, to improve the prognosis of patients, it is critical to study the molecular mechanisms of ESCC. According to previous reports, several risk factors, such as drinking, smoking, diet, nutrition, heredity and body mass index, play essential roles in the occurrence and development of ESCC [1,10,11]. However, the molecular pathogenesis of ESCC is extremely complex, involving several prospects, such as host genetics, host-microbe interactions, and epigenetic imbalance [12–15]. Therefore, it is necessary to explore the molecular mechanisms of ESCC and find new clinical markers beneficial for identifying and predicting prognosis of the disease.

Belonging to the inhibitor of the apoptosis (IAP) family, baculoviral IAP repeat containing 5 (BIRC5, also known as API4, EPR-1, survivin) protein is an inhibitor of apoptosis encoded by its homonymous gene – *BIRC5*. In mammalian cells, it is involved in the control of mitosis, the regulation of apoptosis, and the cell stress response. Increased *BIRC5* expression can be seen in multiple types of cancer, such as non-small cell lung cancer [15] and breast cancer [16]. In ESCC, BIRC5 protein level is also upregulated, which is associated with poor clinical prognosis of ESCC [17,18]. Meanwhile, BIRC5 can reduce the sensitivity of tumor cells to radiation induction [19]. These studies indicate that BIRC5 plays an important role in ESCC.

However, the current reports on BIRC5 and ESCC are limited; most of them only focus on the protein level and are based on a small sample size. Some studies have shown that reduced BIRC5 protein level can affect the development of ESCC by inhibiting the migration and invasion of tumor cells [19], but the mechanism of BIRC5 in the occurrence and development of ESCC is not clear. Also, whether the overexpression of BIRC5 is related to the differentiation, invasion depth, stage, and metastasis of ESCC cells remains controversial [17,19], Therefore, it is necessary to further research the clinical significance and biological role of BIRC5 expression in ESCC and explore the mechanism of the occurrence and development of ESCC.

In this study, by comprehensively using ESCC clinical samples and information from in-house, literature, and multiple databases, we attempted to explore BIRC5 expression in ESCC at both mRNA and protein levels. Furthermore, we have explored the clinical significance of BIRC5 expression in the prognosis and distinction of ESCC and the underlying molecular mechanism of the disease. From this, the research contributes to understanding of the pathogenesis of ESCC, and is conducive to identify ESCC and prediction in prognosis of the disease.

2. Materials and methods

2.1. Tissue microarray and in-house BIRC5 *protein expression assay*

The current study with approval number 2018-KY-(0162), was approved by the Ethics Committee of the Second Affiliated Hospital of Guangxi Medical University, China. The inclusion criteria were: (1) patients diagnosed as ESCC by pathology; (2) patients with completely preserved paraffin specimens; (3) participants aged no less than 18 years; (4) patients with oral informed consent. Patients who did not agree to participate in or did not provide sufficient clinical information for subsequent research were excluded.

To understand the difference in BIRC5 protein expression between ESCC patients and non-ESCC patients, five formalin fixed paraffin embedded tissue microarray sections (ESC1504, ESC1503, ESC242, ESC241, and ESC481), purchased from Fanpu Biotech (Guilin, China), were used to explore BIRC5 protein expression. These sections were used for immunohistochemistry. First, the tissue sections were dewaxed and repaired. After that, 0.01 M citrate buffer (pH = 6.0) was used to soak tissue sections for antigen extraction. Then, the endogenous peroxidase was inactivated by 3% H_2O_2. Moreover, the tissue sections were placed in rabbit anti-human BIRC5 monoclonal antibody (ab126762, Abcam, UK, dilution ratio 1:100), while the negative control slides were placed in phosphate buffer, and then both were cultured overnight at 4°C. Furthermore, a second antibody labeled horseradish peroxidase (Changdao Biotechnology Co., Ltd., Shanghai, China) was added to the tissue sections, which would be heated to 25°C for 25 minutes and dyed with 3,3'-diaminobenzidine. Afterwards, tissue sections were dehydrated and sealed, and then evaluated under a microscope to observe the staining of the nucleus and/or cytoplasm, in which the positive staining was brown granules, and the negative staining was blue granules. In this study, the staining intensity score and the percentage record score were evaluated for each sample. Two senior pathologists individually and randomly selected 10 areas from each staining section for evaluation and recording. The staining intensity score was as follows: integer scores 0–3 respectively represented no, light, moderate, and strong staining; the percentage score was as follows: integer scores 0–4 respectively represented < 5%, 5%–25%, 26%–50%, 51%–75% and >75% positive cells. The total protein expression score was obtained by multiplying the intensity score and the recorded percentage score. Supplementary Appendix 1 shows the clinical parameters of samples in tissue microarray.

2.2. Collection of public microarrays and RNA sequencing data sets of BIRC5 mRNA expression

To further explore the significance of *BIRC5* expression in ESCC, through public databases and published articles, we collected microarrays, RNA sequencing data sets related to *BIRC5* mRNA expression, and clinical information. The databases consulted by us include Gene Expression Omnibus, Oncomine, ArrayExpress, and the GDC Data Portal. The search terms used for microarrays and RNA sequencing data sets was: "esophag* and (mRNA or gene)". The included criteria for data sets were: (1) *homo sapiens*-related research. (2) ESCC-related tissues or cells; (3) data with mRNA expression of *BIRC5*. Data sets with duplicate, incomplete, or unqualified sample numbers were excluded. The steps for selecting data sets are listed in Supplementary Appendix 2.

2.3. Statistical analysis

2.3.1. Data processing

All data processing steps were executed in R software (v4.2.0). Expression of genes was first transformed by $\log_2(x + 1)$. Then, for those platforms (e.g., GPL20795) containing more than one data set, the Surrogate Variable Analysis software package [20] was used to reduce the batch effects. An example of before and after removing batch effects can be seen in Supplementary Appendix 3, and the data was clearly clustered together after removing the batch effects. The limma [21] or edge [22] software packages were used to standardize the gene expression value of each data set.

2.3.2. BIRC5 expression in ESCC and non-ESCC

In R (v4.2.0), the *t* test and the standardized mean difference (SMD) were used to analyze whether there was a significant difference in BIRC5 expression between the ESCC group and the non-ESCC control group. The fixed effect model and random effect model were used to calculate SMD. When I^2 value (I^2 test) < 50% or *p* value (chi-square test) > 0.1, which indicates that the heterogeneity of SMD is not obvious, the fixed effect model should be used. Otherwise, the random effect model was used. For SMD, that the 95% confidence interval (CI) does not include zero represents its statistical significance. The funnel plot of *Begg's* test was applied to assess publication bias, and there was no significant publication bias when $p > 0.1$. Violin plots and forests plots were drawn, respectively, based on ggplot2 and meta [23] packages.

2.3.3. Clinical significance of BIRC5 expression in ESCC

The *Wilcoxon* test was used to explore the correlation between the expression of BIRC5 and the clinical parameters of ESCC patients, including gender, age, survival status, tumor stage, lymph node stage, clinical stage, tumor location, alcohol use, and tobacco use. Receiver operating characteristic curves (ROCs) and summary receiver operating characteristic curve (sROC) were used to identify ESCC samples and non-ESCC samples based on area under the curve (AUC, ranging from zero to one). The closer the AUC is to one, the better the discrimination effect of BIRC5 expression. Also, the range and criteria of sensitivity and specificity are the same as AUC. For likelihood ratios, the bigger the positive likelihood ratio or the smaller the negative likelihood ratio, the increasing superior the discrimination effect of BIRC5 expression. The effect of BIRC5 expression on the prognosis of ESCC patients was explored by Kaplan-Meier survival curve and restricted mean survival time (RMST). All above plots were drawn in R (v4.2.0).

2.3.4. Potential mechanisms of *BIRC5* in ESCC

The limma [21] software package was used to screen differential expression genes. A gene, with the absolute value of $\log_2$ (fold change) ≥ 1 and SMD > 0 (the 95% CI of SMD does not contain zero), was identified as a differential high-expression gene (DHEG). At the same time, positive correlation genes (*Pearson* coefficient ≥ 3, $p < 0.05$) of *BIRC5* expression were screened. The positive correlation genes and DHEGs were crossed to obtain the *BIRC5*-related DHEGs (RDHEGs). To explore the potential molecular mechanism of *BIRC5* in ESCC, Gene Ontology (GO) terms, Disease Ontology (DO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and Reactome pathway

analyses were carried out based on RDHEGs. A term or a pathway with an adjusted p value < 0.05 was selected for the study. Also, the identification of hub genes in ESCC was accomplished by using a protein-protein interaction (PPI) network through RDHEGs.

To explore the mechanism of regulating *BIRC5* expression, we predicted the transcription factors regulating *BIRC5* expression using the following methods: First, we obtained the base sequence of the underlying promoter region of *BIRC5* (1 kb upstream of the *BIRC5* transcription start site) from the National Center for Biotechnology Information. Second, transcription factors (TFs) that may regulate *BIRC5* expression were predicted from animal transcription factor (through the base sequences) [24] and Cistrome Data Browser [25] databases. Third, the initial TFs were obtained by the intersection of predicted TFs and RDHEGs. Fourth, matched sequences between the underlying promoter region of *BIRC5* and motifs of initial TFs (obtained from the JASPAR database [26,27]) were explored via MEME-Suite [28]. Finally, the potential TFs regulating *BIRC5* were identified from the initial TFs with chromatin immunoprecipitation sequencing (CHIP Seq) data in Cistrome Data Browser.

In this study, when a p value < 0.05, there was statistical significance. Supplementary Appendix 4 shows the main design of this study. To avoid confusion, in this study, the *BIRC5* gene was presented in an italic form; the "BIRC5 protein" just represented the protein product of the gene *BIRC5*, while the "BIRC5" without "protein" represented both *BIRC5* gene and BIRC5 protein.

3. Results

3.1. Samples and their information of data sets included in the study

A total of 60 data sets (e.g., GSE70409) for *BIRC5* mRNA expression and one in-house data set for BIRC5 protein expression were included for further research, containing 1,119 samples of the ESCC group and 1,004 samples of the non-ESCC group (Table 1). Then, the data sets with the same platform were reorganized into a new data set, and the batch effects were removed. For example, the new data set called "GPL13497" consists of data sets GSE97558 and GSE45168. A total of 16 new data sets were used for further research (Table 1).

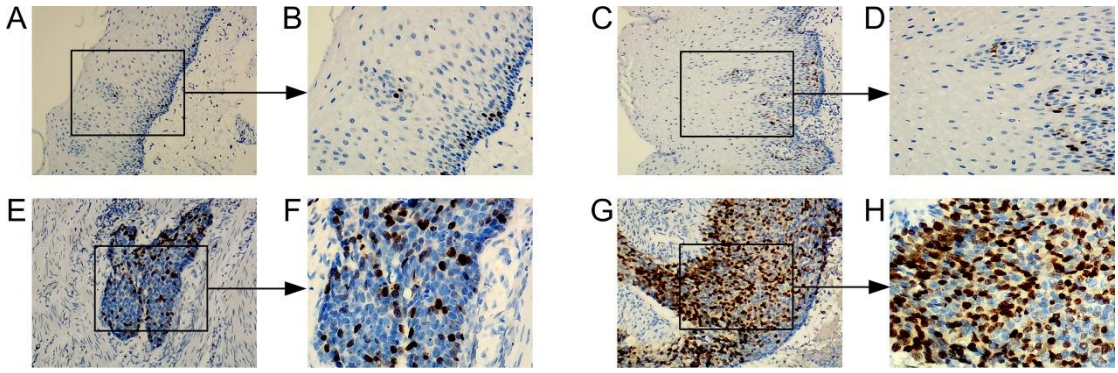


Figure 1. The expression of BIRC5 in tissues by in-house tissue microarrays. The expression of BIRC5 in non-esophageal squamous cell carcinoma (A–D) and esophageal squamous cell carcinoma (E–H) tissues under the microscope. The left figure of each two combined figures is 100x, and the right figure is 200x.

Table 1. The situation of the samples included in this study.

Platform	Data set	*n* of ESCC [a] vs *n* of non-ESCC	Platform	Data set	*n* of ESCC vs *n* of non-ESCC
GPL13287	GSE70409	17 vs 17	GPL570	GSE3526	0 vs 4
GPL13607	GSE45350	4 vs 4		GSE161533	28 vs 56
GPL16956	GSE89102	5 vs 5		GSE77861	7 vs 7
GPL18109	GSE53624	119 vs 119		GSE100942	4 vs 5
GPL9052	GSE119436	4 vs 4		GSE26886	9 vs 19
GPL13497	GSE97558	3 vs 0		GSE33810	2 vs 1
	GSE45168	5 vs 5		GSE17351	5 vs 5
GPL16791	GSE111011	7 vs 7		GSE7307	0 vs 4
	GSE156651	0 vs 4		GSE19472	0 vs 2
	GSE103356	0 vs 6		GSE148247	0 vs 3
	GSE116272	0 vs 4		GSE146808	3 vs 0
	GSE113341	0 vs 10		GSE45670	0 vs 10
	GSE113777	0 vs 1		GSE17353	0 vs 4
GPL20301	GSE149609	17 vs 10		GSE63941	22 vs 0
	GSE142556	2 vs 0		GSE13378	8 vs 0
GPL20795	GSE164158	8 vs 8		GSE67508	8 vs 0
	GSE128914	3 vs 0		GSE86013	2 vs 0
	GSE124514	2 vs 0		GSE27424	0 vs 6
	GSE128913	3 vs 0		GSE69925	274 vs 0
GPL5175	GSE75241	15 vs 15		GSE44021	6 vs 6
	GSE49292	0 vs 3		GSE32701	20 vs 0
	GSE65013	0 vs 6		GSE35975	1 vs 0
	GSE75243	2 vs 0		GSE11373	1 vs 0
GPL571	GSE36223	0 vs 23	GPL96	GSE1420	0 vs 8
	GSE39491	0 vs 40		GSE13083	0 vs 7
	GSE53892	3 vs 0		GSE52138	1 vs 2
	GSE38129	30 vs 30		GSE44021	34 vs 34
	GSE29001	21 vs 24		GSE23400	53 vs 53
	GSE20347	17 vs 17	TCGA-GTEx	TCGA [c]	81 vs 2
	GSE44021	73 vs 73		GTEx [d]	0 vs 291
-	In-house TMA [b]	190 vs 40	-	Total	1119 vs 1004

Notes: [a]: esophageal squamous cell carcinoma; [b]: in-house tissue microarrays; [c]: the Cancer Genome Atlas; [d]: the Genotype-Tissue Expression.

3.2. BIRC5 expression in ESCC

Under the microscope, it can be seen that BIRC5 protein is mainly expressed in ESCC tissues (Figure 1), and upregulated expression of BIRC5 protein is detected in the ESCC group instead of the non-ESCC group via tissue microarrays (Figure 2, $p < 0.05$). Much more interestingly, consistent with the expression of protein level in ESCC, the results of almost every data set (except GSE45350) consistently show that the *BIRC5* expression in the ESCC group was significantly higher than that in

the non-ESCC group (Figure 2, $p < 0.05$).

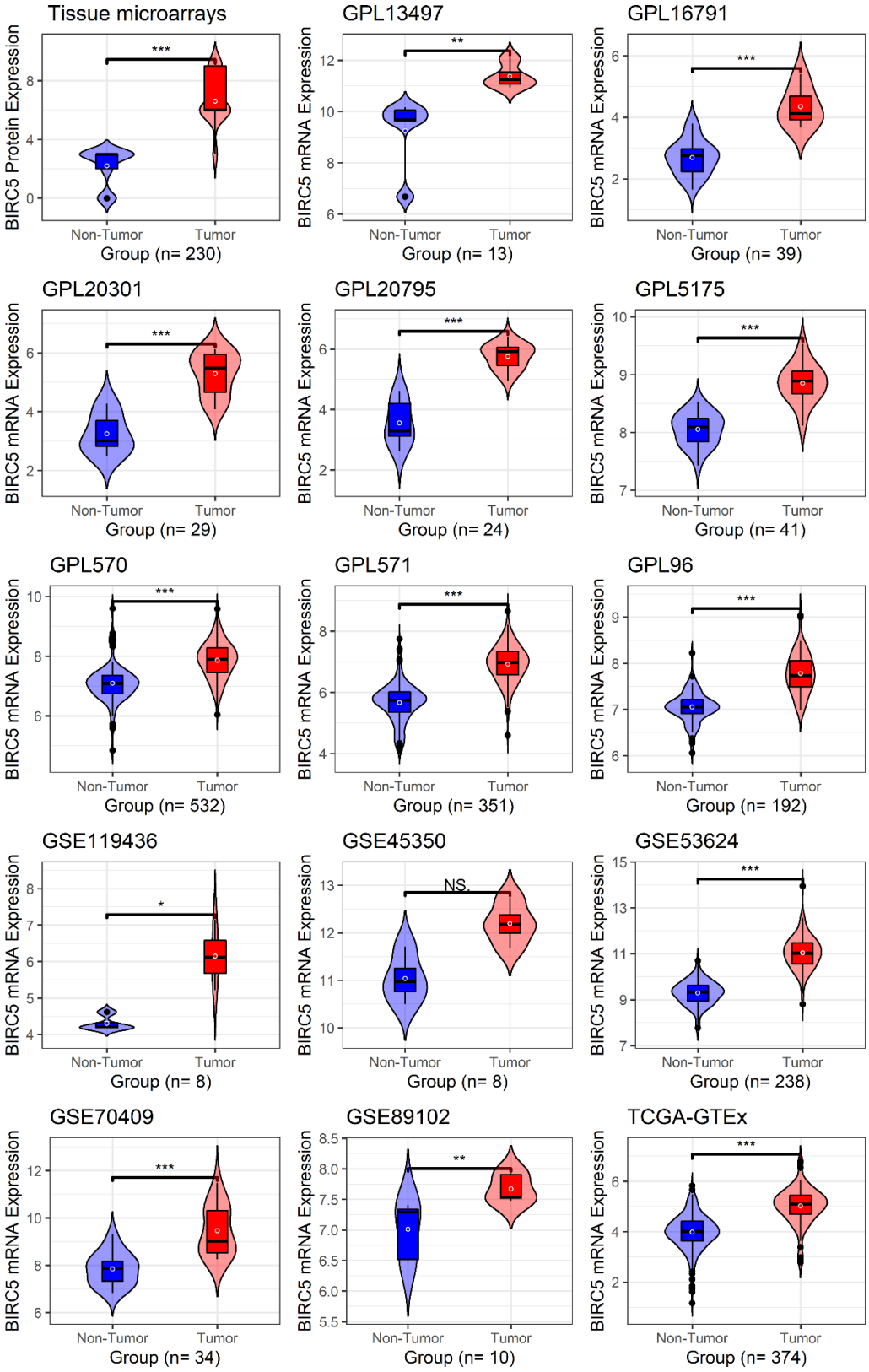


Figure 2. Violin plots of BIRC5 expression. Tissue microarrays are based on BIRC5 protein expression, while others are based on *BIRC5* mRNA expression.

Fifteen data sets were tested for heterogeneity, and the I^2 = 86% (p value < 0.01) indicated significant heterogeneity, for which a random effect model was selected for calculating SMD. It can be seen from Figure 3A that BIRC5 mRNA and protein levels are both highly expressed in the ESCC group rather than the non-ESCC group (SMD = 2.21, 95% CI: 1.84 to 2.58), and *Begg's* funnel plot shows no significant publication bias in the SMD results, p = 0.729 (Figure 3B). These results indicate that *BIRC5* is highly expressed in ESCC.

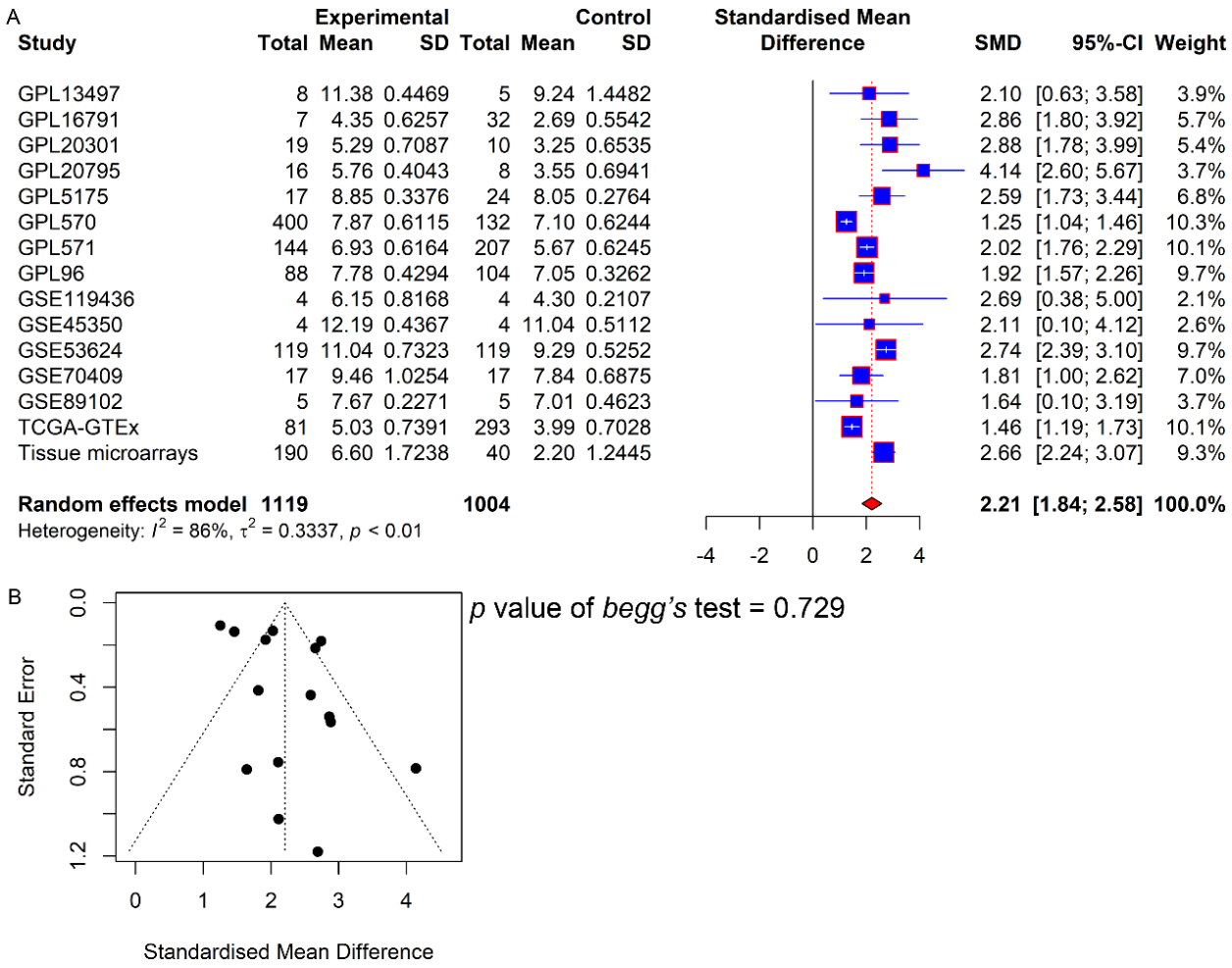


Figure 3. Expression of BIRC5 in esophageal squamous cell carcinoma (ESCC) tissues. A: Forest plot of evaluating standard mean difference (SMD) of BIRC5 expression between ESCC group and non-ESCC group. B: Funnel plot with *Begg's* test for publication bias test.

3.3. The significance of BIRC5 *expression in ESCC*

3.3.1. Relationship between clinical parameters and BIRC5 expression

There exist three data sets—in-house tissue microarrays, GSE53624, and the Cancer Genome Atlas (TCGA)—with clinical parameters. Thus, the relationship between clinical parameters and BIRC5 expression was explored through the three data sets. As shown in Supplementary Appendix 5,

although no clinical parameter was found to be related to BIRC5 expression at both protein and mRNA levels, the mRNA expression of *BIRC5* was higher in patients aged 65 or older and in non-lower esophageal tumor locations ($p < 0.05$, Supplementary Appendixes 5A–5B); also, BIRC5 protein expression was increased in higher lymph node stages (N1–N3) and clinical stages (III–IV) ($p < 0.05$, Supplementary Appendix 5C).

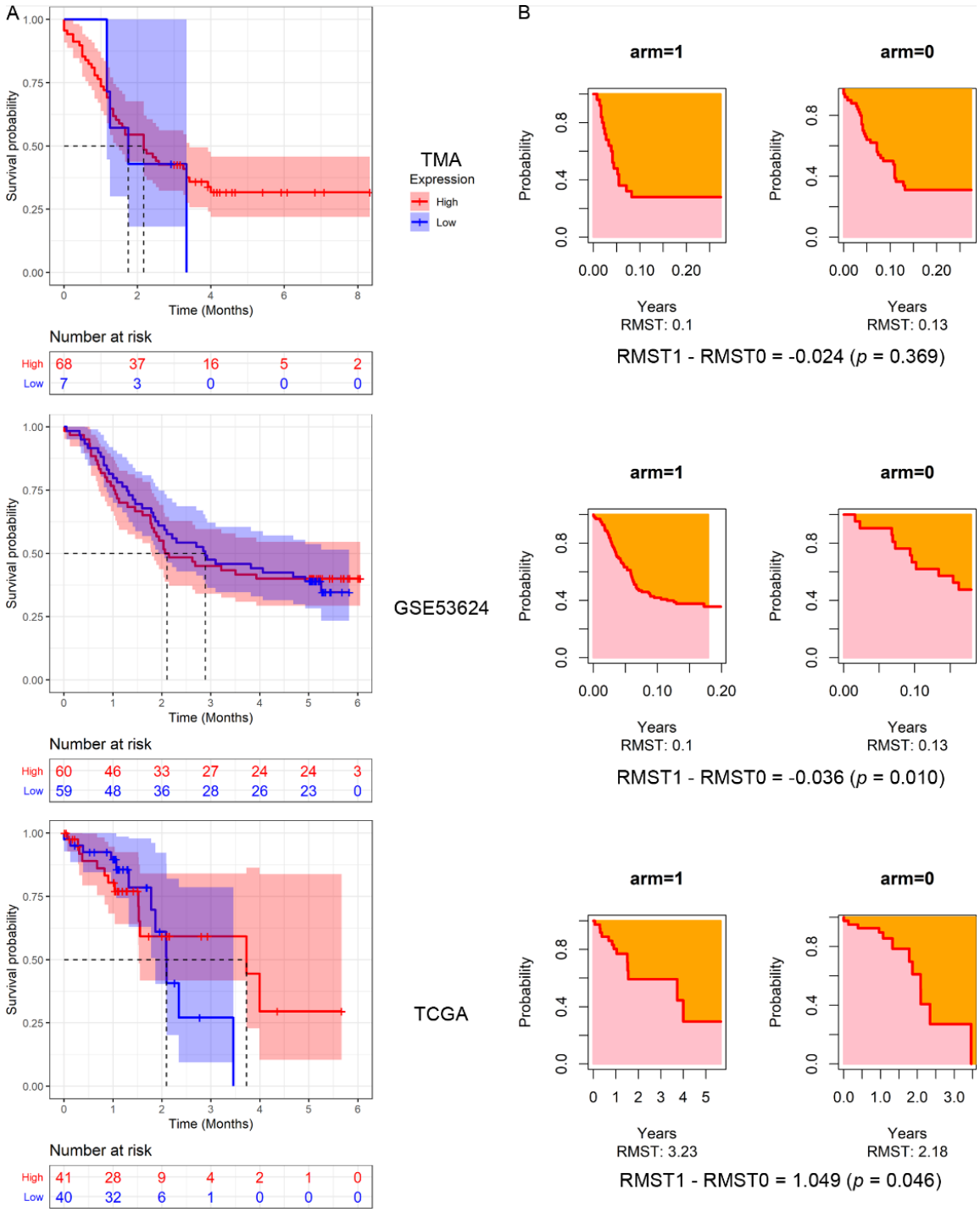


Figure 4. The Kaplan–Meier curves and restricted mean survival time of BIRC5 expression. TMA: tissue microarrays, TCGA: the Cancer Genome Atlas.

3.3.2. Relationship between BIRC5 expression and prognosis of ESCC patients

Based on prognosis information from ESCC samples in the three data sets—in-house tissue microarrays, GSE53624, and TCGA—the median of the expression value of BIRC5 was used as the cut-off value for survival analyses. Through the Kaplan–Meier curves (Figure 4A), there is a cross between the high expression group and the low expression group for each data set. Thus, the RMST instead of the *log-rank* test was suitable for evaluating the relationship between BIRC5 expression and the prognosis of ESCC patients [29]. For in-house tissue microarrays, although patients in the high BIRC5 protein level group tended to have a shorter RMST, the difference was not statistically significant ($p > 0.05$, Figure 4B). For GSE53624, ESCC patients from the high *BIRC5* expression group tended to have a shorter RMST than those from the low *BIRC5* expression group ($p < 0.05$, Figure 4B). However, this was not consistent with the TCGA data set, with RMST of the high *BIRC5* expression group = 3.23 years and that of the low expression of *BIRC5* group = 2.18 years ($p < 0.05$, Figure 4B).

3.3.3. Effect of BIRC5 expression in distinguishing ESCC

Via ROCs and sROC, BIRC5 expression shows significant identifying ability between ESCC and non-ESCC samples, with all AUCs of ROCs > 0.82 (Supplementary Appendix 6) and AUC of sROC was 0.96 (95% CI: 0.93–0.97, Figure 5). The sensitivity and specificity were 0.92 (95% CI: 0.84–0.96) and 0.91 (95% CI: 0.87–0.94), respectively. The positive likelihood ratio and negative likelihood ratio were 10.71 (95% CI: 7.11–16.14) and 0.09 (95% CI: 0.05–0.17), respectively (Supplementary Appendix 7). These indicate that BIRC5 expression has an outstanding ability to distinguish ESCC samples from non-ESCC samples.

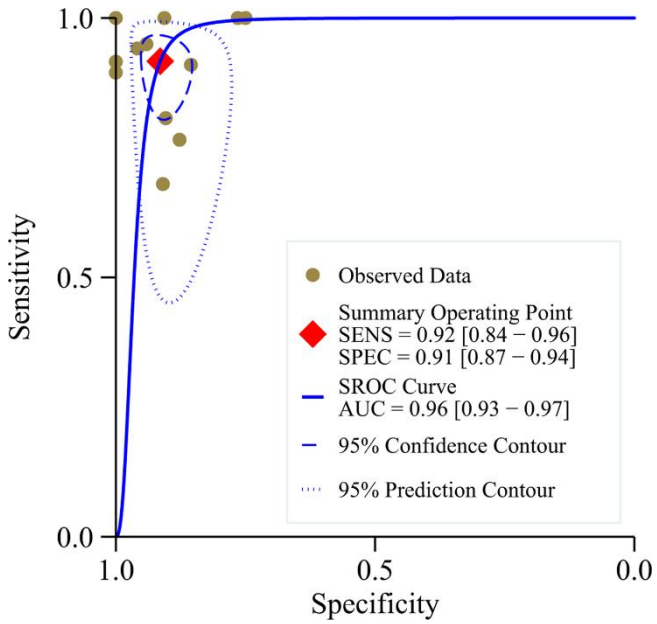
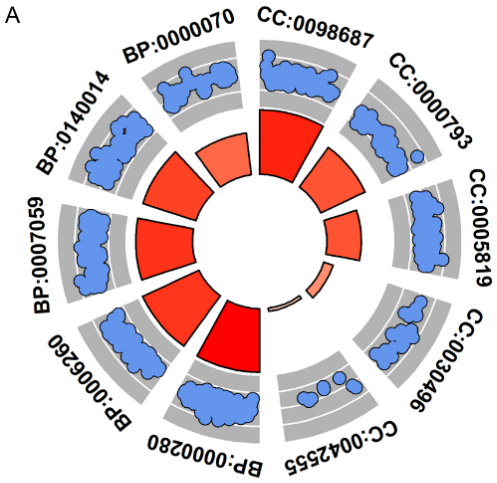


Figure 5. Summary receiver operating characteristic curve for identifying esophageal squamous cell carcinoma based on BIRC5 expression.

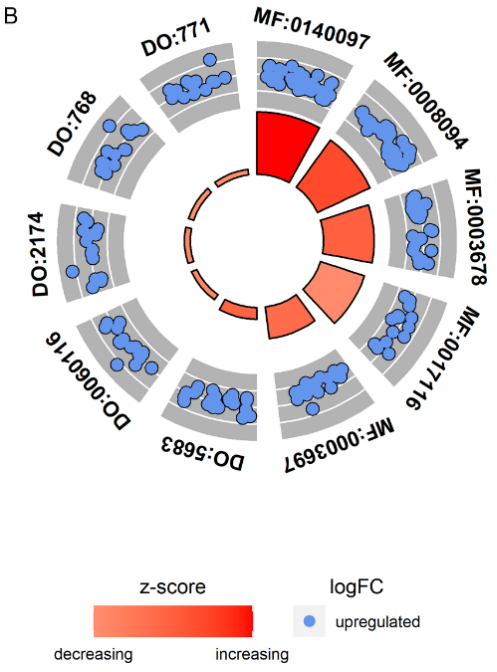
3.4. The potential mechanism of BIRC5 in ESCC

3.4.1. Identification of RDHEGs for exploring mechanism of *BIRC5* in ESCC

A total of 8,857 genes with the absolute value of $\log_2$ (fold change) $\geqslant$ 1 were screened (Supplementary Appendix 8). Among them, 4,484 were selected with SMD > 0 and the 95% CI of SMD did not contain zero (data not shown). In the genes related to *BIRC5* expression, 647 were identified to be positively correlated with *BIRC5* expression in at least five data sets. Then, 445 RDHEGs were obtained by crossing DHEGs with *BIRC5*-positive correlation genes (data not shown).



ID	Description
CC:0098687	Chromosomal region
CC:0000793	Condensed chromosome
CC:0005819	Spindle
CC:0030496	Midbody
CC:0042555	MCM complex
BP:0000280	Nuclear division
BP:0006260	DNA replication
BP:0007059	Chromosome segregation
BP:0140014	Mitotic nuclear division
BP:0000070	Mitotic sister chromatid segregation



ID	Description
MF:0140097	Catalytic activity, acting on DNA
MF:0008094	DNA-dependent ATPase activity
MF:0003678	DNA helicase activity
MF:0017116	Single-stranded DNA helicase activity
MF:0003697	Single-stranded DNA binding
DO:5683	Hereditary breast ovarian cancer
DO:0060116	Sensory system cancer
DO:2174	Ocular cancer
DO:768	Retinoblastoma
DO:771	Retinal cell cancer

Figure 6. Gene Ontology terms and Disease Ontology terms of *BIRC5*-related differential high-expression genes. CC: cellular component; BP: biological process; MF: molecular function; DO: Disease Ontology.

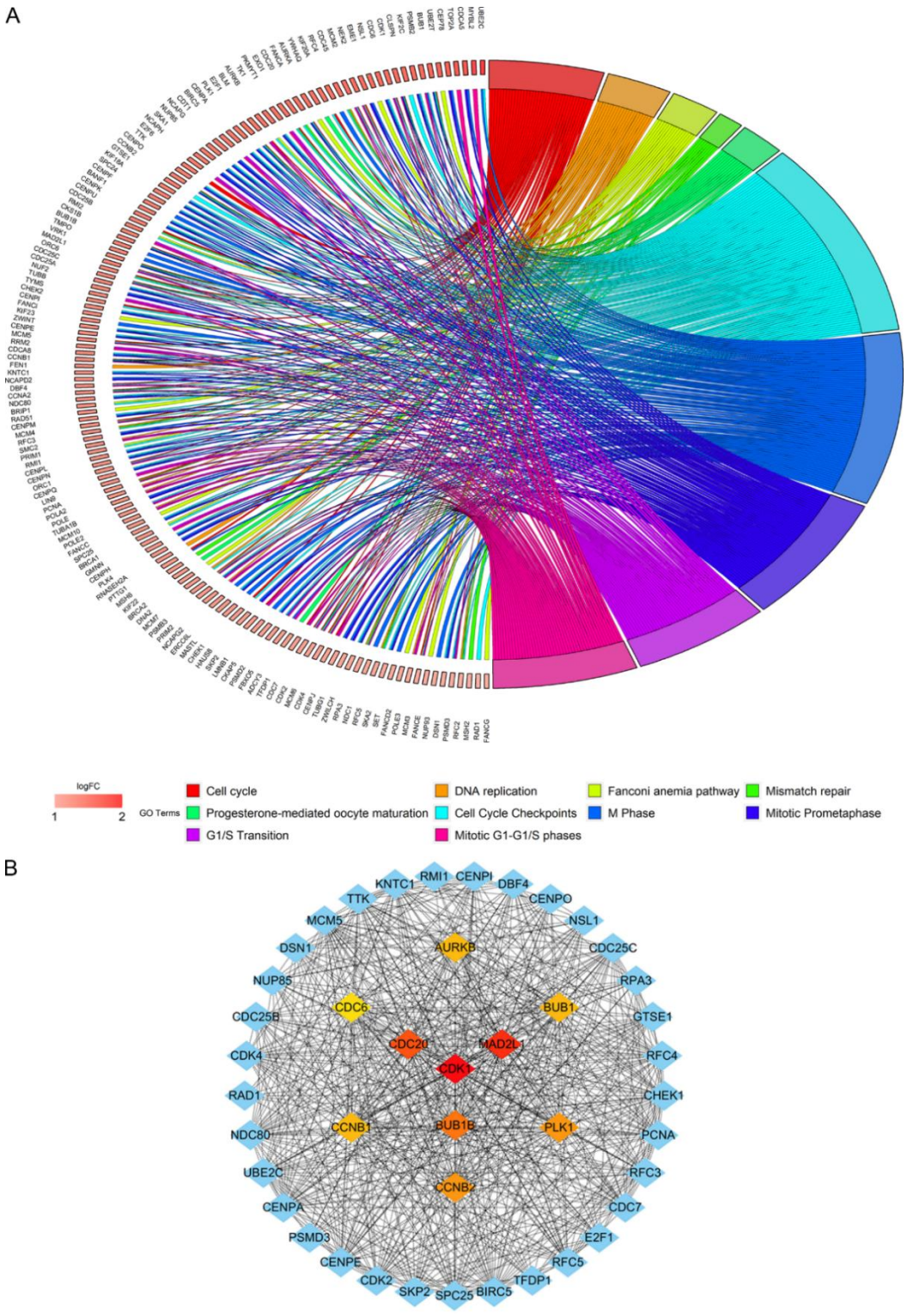


Figure 7. Kyoto Encyclopedia of Genes and Genomes, Reactome pathways (A), and protein-protein interaction (B) of *BIRC5*-related differential high-expression genes. For panel A, Cell cycle, DNA replication, Fanconi anemia pathway, Mismatch repair and Progesterone-mediated oocyte maturation are based on Kyoto Encyclopedia of Genes and Genomes, and the rest are the results of Reactome pathway analysis.

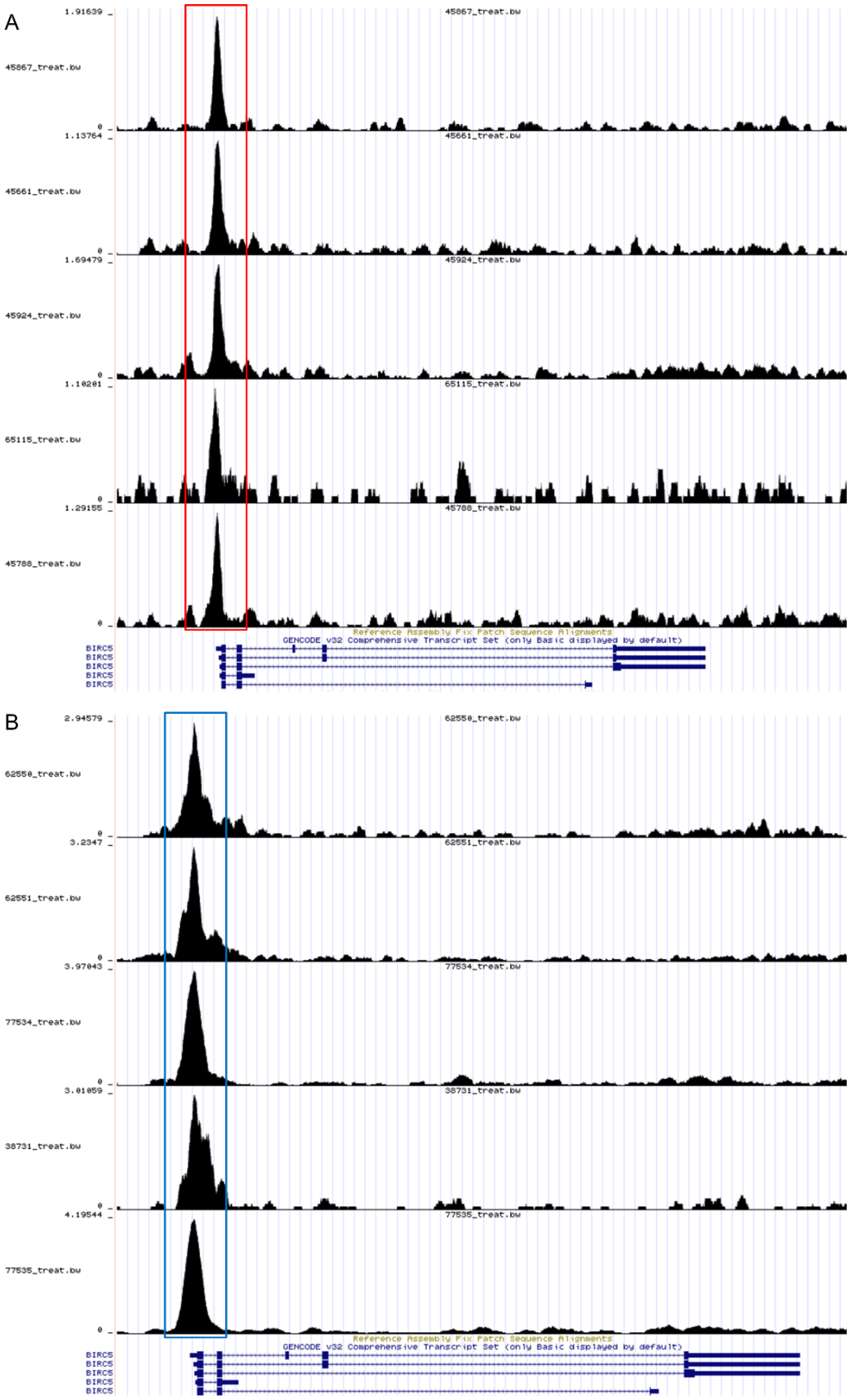


Figure 8. For both *MAZ* (A) and *TFDP1* (B), there exist chromatin immunoprecipitation sequencing binding sites with the potential promoter region of *BIRC5*.

3.4.2. Enrichment analyses for exploring mechanism of *BIRC5* in ESCC

In GO analysis, RDHEGs significantly clustered in GO terms related to DNA replication, including "Chromosomal region" (cell components), "Nuclear division" (biological processes), and "Catalytic activity acting on DNA" (molecular functions) (Figure 6). In DO enrichment analysis, RDHEGs significantly involved in a variety of tumors, including "Hereditary breast ovarian cancer" and "Sensory system cancer" (Figure 6B).

In the main KEGG analysis, the top pathway of RDHEGs is the "Cell cycle," while the Reactome enrichment pathway of RDHEGs is the "Cell cycle checkpoint" (Figure 7A). Both KEGG and Reactome analyses suggest that RDHEGs play an essential role in the cell cycle. Thus, we further performed PPI analysis based on 88 RDHEGs in the cell cycle checkpoint and cell cycle pathways. Then, *CDK1* (cyclin dependent kinase 1), *MAD2L*, and *CDC20* (cell cycle 20) are identified as hub genes in these RDHEGs (Figure 7B), indicating their important role in the cell cycle.

To explore the mechanism of high expression of *BIRC5* in ESCC, TFs that may regulate *BIRC5* expression were predicted. 422 and 100 TFs were predicted from Animal TFDB and Cistrome Data Browse, respectively. After the intersection of these TFs and RDHEGs, four TFs—*E2F1*, *EZH2*, *MAZ,* and *TFDP1*—were screened (Supplementary Appendix 9A). The four TFs were highly expressed in ESCC (SMDs > 0, 95% CI excluding 0) (Supplementary Appendix 9B) and positively correlated with the expression of *BIRC5* (Supplementary Appendix 10). Among the four TFs, *E2F1* has been reported to regulate *BIRC5* expression [30], while no motif of *EZH2* can be found in JASPAR. Thus, *MAZ* and *TFDP1* were further analyzed. With the MEME-Suite analysis tool, among the upstream *BIRC5* transcription initiation site (chr78213252–7821425) locate the binding sequences of *MAZ* and *TFDP1* (Supplementary Appendix 11). Moreover, there were ChIP-Seq binding peaks of *MAZ* and *TFDP1* in the potential promoter region upstream of the transcription start site of *BIRC5*, suggesting that these two TFs regulate the transcription of *BIRC5* and affect its expression (Figure 8). Therefore, *MAZ* and *TFDP1* may regulate *BIRC5* expression.

4. Discussion

In this study, to promote understanding of the pathogenesis of ESCC and provide a new idea for the clinical distinction and treatment of ESCC, we pay attention to the changes in the expression level of BIRC5 in ESCC, the clinical significance of BIRC5 expression level, and the underlying molecular mechanism of it in ESCC. By analyzing numerous samples (n = 2,123) from multiple research centers, with a t test and calculating SMD, we found that BIRC5 mRNA and protein are highly expressed in the ESCC group. The high expression of BIRC5 in ESCC is related to the clinical characteristics of ESCC patients, including age, lymph node stages, and clinical stages. Upregulated BIRC5 expression makes it feasible to distinguish ESCC from non-ESCC, suggesting its important clinical significance in ESCC. Also, with several analytical methods, this study also reveals that *BIRC5* plays an essential role in ESCC and that it is mainly related to the cell cycle. The two TFs, *MAZ* and *TFDP1*, may contribute to the upregulated expression of *BIRC5* in ESCC; and as far as we know, no relevant reports have been found before this study, demonstrating the novelty of our research. In conclusion, this study demonstrated that elevated BIRC5 expression may have important clinical significance in ESCC. Its molecular mechanism in ESCC was initially discussed, which contributes to the understanding of ESCC in pathogenesis.

ESCC is one of the most common types of malignant tumors in humans. Most ESCC patients were diagnosed at an advanced stage, and their prognosis is not ideal [31]. Therefore, exploring new markers for distinguishing ESCC from non-ESCC and predicting prognosis of ESCC patients contributes to improving the survival rate of patients [32]. As a member of the apoptosis-suppressing gene family, *BIRC5* is of great significance in a variety of malignant tumors. For instance, upregulated *BIRC5* expression can be seen in multiple cancers, such as non-small cell lung cancer [15] and breast cancer [16], and it may become a marker for early detection and/or treatment for these two cancers. Elevated BIRC5 expression has been revealed in ESCC based on previous reports [15,33], but some limitations are generally common in these reports, such as small sample size and focusing on a single level (protein) of BIRC5 expression. In order to further verify the expression of BIRC5 in ESCC, this study, first based on more than 2,000 samples from multiple research centers and various statistical methods, disclosed that the mRNA and protein expression levels of BIRC5 were upregulated in ESCC.

The high expression of BIRC5 may be related to the poor prognosis of ESCC patients, and it makes it possible to distinguish ESCC from non-ESCC. Regarding the clinical significance of BIRC5 in ESCC, Shang et al. [19] found that as ESCC progressed, the expression level of BIRC5 increased. However, whether BIRC5 expression is correlated with the tumor size, lymph node metastasis, and advanced clinical stage of ESCC patients is still controversial [15,17]. In this study, we have not yet observed that BIRC5 is related to several clinical parameters—gender, survival status, tumor size, distant metastasis, alcohol use, and tobacco consumption —of ESCC patients at both the mRNA and protein levels. However, high mRNA expression of *BIRC5* can be detected in patients with the location of ESCC lesions (lower esophagus) and an older age (≥ 65 years old). The protein expression of BIRC5 is increased in the higher lymph node stage (N1–N3) and late clinical stages (III–IV). These results suggest that BIRC5 may be associated with the prognosis of ESCC. However, according to previous reports, with BIRC5 expression, ESCC patients has better [34] or worse [35] prognosis is still controversial. Interestingly, such a contradiction can be seen between mRNA and protein expression levels of BIRC5 in this study. Based on samples from TCGA, ESCC patients with high expression of *BIRC5* had longer RMST, which is contrary to the result of GSE53624. However, we tended to believe that BIRC5 expression was associated with a worse prognosis of ESCC patients for these reasons: (1) The sample in the GSE53624 data set (n = 119) is larger than the sample in the TCGA data set (n = 81), and the result of a larger sample size is more reliable. (2) The prognostic data of the GSE53624 data set is more orderly than that of TCGA (in the TCGA data set, before the 2-year survival time, the prognosis of patients with high expression of *BIRC5* is worse, while after 2-year survival time, patients with upregulated expression of *BIRC5* show better prognosis), so results based on GSE53624 are more credible. (3) The high expression of BIRC5 protein is related to advanced lymph node and clinical stages. Therefore, current studies suggest that elevated BIRC5 expression may be a risk factor for poor prognosis in ESCC patients. It is also worth noting that, based on the expression data of BIRC5, the samples of the ESCC group can be well distinguished from the samples of the non-ESCC group, suggesting the potential of BIRC5 as an ESCC discrimination indicator; a finding that, as far as we know, has not been reported before. In summary, the current study indicates that *BIRC5* may have potential significance in the clinical application of ESCC, such as treatment and early discrimination of the disease.

To improve the understanding of the role of BIRC5 in ESCC, we also analyzed the molecular mechanism of BIRC5 in ESCC. The expression products of RDHEGs screened based on *BIRC5* were mainly distributed in the cell chromosomal region, mainly involved in the process of cell nuclear

division and DNA replication and participate in catalyzing DNA activity. The important period of DNA replication is the S phase of the interphase (interphase is one phase of the cell cycle, and the other is the division phase). Interestingly, although the BIRC5 (also known as survivin) protein encoded by *BIRC5* mainly aggregates in G2 during cell division, it has been localized in the nucleus during the S phase [36], suggesting the possibility of BIRC5 participating in DNA replication. Also, BIRC5 expression may not only be related to cell proliferation (DNA replication), but also affect the progression of a variety of tumors (such as ovarian cancer and sensory system cancer) based on DO analysis, implying the key role of BIRC5 in ESCC. Further analysis revealed that RDHEGs significantly aggregated in the cell cycle pathways of KEGG and Reactome, indicating that the mechanism of BIRC5 in ESCC may be related to the cell cycle. According to previous studies, BIRC5 protein is essential for cell division and can inhibit cell death [37]. The mechanism remains unclear, and the cell cycle may be one of the factors involved. In addition to the interphase shown above, the division phase may also be the stage where BIRC5 protein participates in the cell cycle process. During the division phase, BIRC5 protein helps the chromosomes to be arranged correctly in the later stage by positioning the chromosomal passenger complex at the centromere of the ante-mid period of mitosis [37]. Thus, BIRC5 may participate in cell proliferation by maintaining the cell cycle. In short, RDHEGs screened based on *BIRC5* are mainly involved in the cell cycle, especially DNA replication, suggesting that BIRC5 may participate in the occurrence and development of ESCC through these pathways.

With PPI network analysis of RDHEGs enriched in the cell cycle pathway, *CDK1*, *CDC20*, and *MAD2L* were found to be the hub genes of these RDHEGs in the cell cycle pathway. *CDK1*, a serine/threonine protein kinase that acts on the G2/M point of the cell cycle, is a necessary condition for eukaryotic cell division and thus participates in cell proliferation. Abnormal regulation of *CDK1* leads to abnormal cell differentiation and cell cycles, and eventually leads to malignant tumor formation [38]. Like *CDK1*, *CDC20* is overexpressed in a variety of human tumors and is considered to show a carcinogenic effect in human tumorigenesis [39]. Its expression was associated with poor prognosis in patients with breast cancer [40] and colorectal cancer [41]. Another hub gene, *MAD2L*, was thought to be associated with gastric cancer progression by participating in the cell cycle [42], although it is rarely reported in tumors. As RDHEGs screened based on *BIRC5*, the three genes – *CDK1*, *CDC20*, and *MAD2L* all play a role in promoting cancer and are also closely related to the cell cycle. Therefore, it is worth further exploring whether *BIRC5* plays a role in ESCC by interacting with these three genes.

In this study, we also found that *MAZ* and *TFDP1* may be transcription factors regulating *BIRC5* expression. Studies have shown that *MAZ* is overexpressed in prostate cancer tissues with bone metastasis and further enhanced in metastatic bone tissues [43]. The *TFDP1* gene was amplified in non-small cell lung cancer [44], and was involved in the amplification of hepatocellular carcinoma by up-regulating the expression of *CCNE1* [45]. These findings suggest that *MAZ* and *TFDP1* are involved in the expression regulation of cancer-related genes. In this study, these two TFs were not only upregulated in ESCC (similar to *BIRC5*), but also significantly positively correlated with the expression of *BIRC5* in ESCC. Furthermore, for *MAZ* and *TFDP1*, there were ChIP-Seq binding peaks in the potential promoter region of *BIRC5*. These results suggest that the transcription and expression of *BIRC5* in ESCC may be caused by the regulation of *MAZ* and *TFDP1*, but this needs further study.

Although we have revealed some noteworthy findings, this study still has limitations: First, all the samples we collected are from tissues or cells, and we failed to collect serum samples from ESCC

patients to study the clinical significance of BIRC5 in ESCC, such as whether it is possible to quickly screen ESCC by detecting serum BIRC5 levels. At the same time, this study lacks samples with detailed prognostic parameters, so it is impossible to carry out prognostic-related research through large samples. Also, in the future, the regulation of *MAZ* and *TFDP1* on *BIRC5* should be verified through in vivo and in vitro experiments.

5. Conclusions

In summary, the current studies have revealed that the mRNA and protein of BIRC5 are highly expressed in the ESCC group, and the highly expressed BIRC5 may be used as a potential marker for discrimination and treatment of ESCC. Also, this study also explored the potential mechanism of BIRC5 in ESCC and found that the cell cycle may be an important way for BIRC5 to participate in the occurrence and development of ESCC, and the regulation of *MAZ* and *TFDP1* may be a mechanism for the high expression of *BIRC5* in ESCC.

Acknowledgments

This work was supported by the College Student Innovation and Entrepreneurship Training Program Project of the First Clinical Medical University of Guangxi Medical University (2020YFYA05, 2020YFYA06), the College Student Innovation and Entrepreneurship Training Program Project (202010598052) and the Guangxi Medical University Future Academic Stars Project (WLXSZX20081). The results shown in the study are in part based upon data generated by the TCGA Research Network: www.cancer.gov/tcga.

Conflict of interest

The authors have no conflicts of interest to declare.

References

1. M. J. Domper Arnal, A. Ferrandez Arenas, A. Lanas Arbeloa, Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries, *World J. Gastroenterol.*, **21** (2015), 7933–7943.
2. F. Kamangar, G. M. Dores, W. F. Anderson, Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world, *J. Clin. Onco.l*, **24** (2006), 2137–2150.
3. J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, et al., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, *Int. J. Cancer*, **136** (2015), E359–386.
4. P. C. Enzinger, R. J. Mayer, Esophageal cancer, *N Engl. J. Med.*, **349** (2003), 2241–2252.
5. A. Hesari, M. Azizian, A. Sheikhi, A. Nesaei, S. Sanaei, N. Mahinparvar, et al., Chemopreventive and therapeutic potential of curcumin in esophageal cancer: Current and future status, *Int. J. Cancer*, **144** (2019), 1215–1226.
6. M. Watanabe, Y. Tachimori, T. Oyama, Y. Toh, H. Matsubara, M. Ueno, et al., Comprehensive registry of esophageal cancer in Japan, 2013, *Esophagus*, **18** (2021), 1–24.

7. M. di Pietro, M. I. Canto, R. C. Fitzgerald, Endoscopic management of early adenocarcinoma and squamous cell carcinoma of the esophagus: Screening, diagnosis, and therapy, *Gastroenterology*, **154** (2018), 421–436.
8. S. W. Chen, H. F. Zhou, H. J. Zhang, R. Q. He, Z. G. Huang, Y. W. Dang, et al., The clinical significance and potential molecular mechanism of PTTG1 in esophageal squamous cell carcinoma, *Front. Genet.*, **11** (2020), 583085.
9. E. O. Then, M. Lopez, S. Saleem, V. Gayam, T. Sunkara, A. Culliford, et al., Esophageal cancer: An updated surveillance epidemiology and end results database analysis, *World J. Oncol.*, **11** (2020), 55–64.
10. C. C. Abnet, M. Arnold, W. Q. Wei, Epidemiology of esophageal squamous cell carcinoma, *Gastroenterology*, **154** (2018), 360–373.
11. M. Recalde, V. Davila-Batista, Y. Diaz, M. Leitzmann, I. Romieu, H. Freisling, et al., Body mass index and waist circumference in relation to the risk of 26 types of cancer: A prospective cohort study of 3.5 million adults in Spain, *BMC Med.*, **19** (2021), 10.
12. Z. W. Reichenbach, M. G. Murray, R. Saxena, D. Farkas, E. G. Karassik, A. Klochkova, et al., Clinical and translational advances in esophageal squamous cell carcinoma, *Adv. Cancer Res.*, **144** (2019), 95–135.
13. W. Cao, H. Lee, W. Wu, A. Zaman, S. McCorkle, M. Yan, et al., Multi-faceted epigenetic dysregulation of gene expression promotes esophageal squamous cell carcinoma, *Nat. Commun.*, **11** (2020), 3675.
14. Z. C. Xie, H. Y. Wu, F. C. Ma, Y. W. Dang, Z. G. Peng, H. F. Zhou, et al., Prognostic alternative splicing signatures and underlying regulatory network in esophageal carcinoma, *Am. J. Transl. Res.*, **11** (2019), 4010–4028.
15. Y. Bi, S. Guo, X. Xu, P. Kong, H. Cui, T. Yan, et al., Decreased ZNF750 promotes angiogenesis in a paracrine manner via activating DANCR/miR-4707-3p/FOXC2 axis in esophageal squamous cell carcinoma, *Cell Death Dis.*, **11** (2020), 296.
16. K. Ghaffari, M. Hashemi, E. Ebrahimi, R. Shirkoohi, BIRC5 genomic copy number variation in early-onset breast cancer, *Iran Biomed. J.*, **20** (2016), 241–245.
17. C. Li, Z. Li, M. Zhu, T. Zhao, L. Chen, W. Ji, et al., Clinicopathological and prognostic significance of survivin over-expression in patients with esophageal squamous cell carcinoma: a meta-analysis, *PLoS One*, **7** (2012), e44764.
18. H. Xia, S. Chen, H. Huang, H. Ma, Survivin over-expression is correlated with a poor prognosis in esophageal cancer patients, *Clin. Chim. Acta*, **446** (2015), 82–85.
19. X. Shang, G. Liu, Y. Zhang, P. Tang, H. Zhang, H. Jiang, et al., Downregulation of BIRC5 inhibits the migration and invasion of esophageal cancer cells by interacting with the PI3K/Akt signaling pathway, *Oncol. Lett.*, **16** (2018), 3373–3379.
20. J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments, *Bioinformatics*, **28** (2012), 882–883.
21. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, et al., limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Res.*, **43** (2015), e47.
22. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics*, **26** (2010), 139–140.

23. S. Balduzzi, G. Rucker, G. Schwarzer, How to perform a meta-analysis with R: A practical tutorial, *Evid. Based Ment. Health*, **22** (2019), 153–160.
24. H. Hu, Y. R. Miao, L. H. Jia, Q. Y. Yu, Q. Zhang, A. Y. Guo, AnimalTFDB 3.0: A comprehensive resource for annotation and prediction of animal transcription factors, *Nucleic Acids Res.*, **47** (2019), D33–D38.
25. R. Zheng, C. Wan, S. Mei, Q. Qin, Q. Wu, H. Sun, et al., Cistrome Data Browser: Expanded datasets and new tools for gene regulatory analysis, *Nucleic Acids Res.*, **47** (2019), D729–D735.
26. G. D. Stormo, Modeling the specificity of protein-DNA interactions, *Quant. Biol.*, **1** (2013), 115–130.
27. W. W. Wasserman, A. Sandelin, Applied bioinformatics for the identification of regulatory elements, *Nat. Rev. Genet.*, **5** (2004), 276–287.
28. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: Scanning for occurrences of a given motif, *Bioinformatics*, **27** (2011), 1017–1018.
29. L. Zhao, B. Claggett, L. Tian, H. Uno, M. A. Pfeffer, S. D. Solomon, et al., On the restricted mean survival time curve in survival analysis, *Biometrics*, **72** (2016), 215–221.
30. S. Tian, L. Zhang, Y. Li, D. Cao, S. Quan, Y. Guo, et al., Human papillomavirus E7 oncoprotein promotes proliferation and migration through the transcription factor E2F1 in cervical cancer cells, *Anticancer Agents Med. Chem.*, (2020).
31. M. Arnold, I. Soerjomataram, J. Ferlay, D. Forman, Global incidence of oesophageal cancer by histological subtype in 2012, *Gut*, **64** (2015), 381–387.
32. L. Jamali, R. Tofigh, S. Tutunchi, G. Panahi, F. Borhani, S. Akhavan, et al., Circulating microRNAs as diagnostic and therapeutic biomarkers in gastric and esophageal cancers, *J. Cell Physiol.*, **233** (2018), 8538–8550.
33. U. Malhotra, A. H. Zaidi, J. E. Kosovec, P. M. Kasi, Y. Komatsu, C. L. Rotoloni, et al., Prognostic value and targeted inhibition of survivin expression in esophageal adenocarcinoma and cancer-adjacent squamous epithelium, *PLoS One*, **8** (2013), e78343.
34. U. Warnecke-Eberz, S. Hokita, H. Xi, H. Higashi, S. E. Baldus, R. Metzger, et al., Overexpression of survivin mRNA is associated with a favorable prognosis following neoadjuvant radiochemotherapy in esophageal cancer, *Oncol. Rep.*, **13** (2005), 1241–1246.
35. Y. X. Zhou, Q. Liu, H. Wang, F. Ding, Y. Q. Ma, The expression and prognostic value of SOX2, beta-catenin and survivin in esophageal squamous cell carcinoma, *Future Oncol.*, **15** (2019), 4181–4195.
36. V. A. Beardmore, L. J. Ahonen, G. J. Gorbsky, M. J. Kallio, Survivin dynamics increases at centromeres during G2/M phase transition and is regulated by microtubule-attachment and Aurora B kinase activity, *J. Cell Sci.*, **117** (2004), 4033–4042.
37. S. P. Wheatley, D. C. Altieri, Survivin at a glance, *J. Cell Sci.*, **132** (2019).
38. L. Sisinni, F. Maddalena, V. Condelli, G. Pannone, V. Simeon, V. Li Bergolis, et al., TRAP1 controls cell cycle G2-M transition through the regulation of CDK1 and MAD2 expression/ubiquitination, *J. Pathol.*, **243** (2017), 123–134.
39. L. Wang, J. Zhang, L. Wan, X. Zhou, Z. Wang, W. Wei, Targeting Cdc20 as a novel cancer therapeutic strategy, *Pharmacol. Ther.*, **151** (2015), 141–151.
40. H. Karra, H. Repo, I. Ahonen, E. Loyttyniemi, R. Pitkanen, M. Lintunen, et al., CDC20 and securin overexpression predict short-term breast cancer survival, *B.r J. Cancer*, **110** (2014), 2905–2913.
41. W. J. Wu, K. S. Hu, D. S. Wang, Z. L. Zeng, D. S. Zhang, D. L. Chen, et al., CDC20 overexpression

predicts a poor prognosis for patients with colorectal cancer, *J. Transl. Med.*, **11** (2013), 142.
42. Z. Wang, C. Dang, R. Yan, H. Zhang, D. Yuan, K. Li, Screening of cell cycle-related genes regulated by KIAA0101 in gastric cancer, *Nan Fang Yi Ke Da Xue Xue Bao*, **38** (2018), 1151–1158.
43. Q. Yang, C. Lang, Z. Wu, Y. Dai, S. He, W. Guo, et al., MAZ promotes prostate cancer bone metastasis through transcriptionally activating the KRas-dependent RalGEFs pathway, *J. Exp. Clin. Cancer Res.*, **38** (2019), 391.
44. S. D. Castillo, B. Angulo, A. Suarez-Gauthier, L. Melchor, P. P. Medina, L. Sanchez-Verde, et al., Gene amplification of the transcription factor DP1 and CTNND1 in human lung cancer, *J. Pathol.*, **222** (2010), 89–98.
45. K. Yasui, S. Arii, C. Zhao, I. Imoto, M. Ueda, H. Nagai, et al., TFDP1, CUL4A, and CDC16 identified as targets for amplification at 13q34 in hepatocellular carcinomas, *Hepatology*, **35** (2002), 1476–1484.

AIMS Press