

## Research Article

# Systematic Analysis of the Expression and Prognostic Significance of P4HA1 in Pancreatic Cancer and Construction of a lncRNA-miRNA-P4HA1 Regulatory Axis

Zhili Hu,<sup>1,2</sup> Fang Song,<sup>3</sup> Yangzhi Hu,<sup>2</sup> and Tianyou Liao <sup>4</sup>

<sup>1</sup>Department of Gastrointestinal Surgery, Liuzhou People's Hospital, Liuzhou, Guangxi Province 545006, China

<sup>2</sup>Department of Gastrointestinal Surgery, Affiliated Hospital of Xiangnan University, Chenzhou, Hunan Province 423000, China

<sup>3</sup>Department of Gastroenterology, Zaoyang First People's Hospital, Zaoyang, Xiangyang, Hubei Province 441200, China

<sup>4</sup>Department of Gastrointestinal Surgery, Shunde Hospital, Southern Medical University (The First People's Hospital of Shunde Foshan), Shunde, Foshan, Guangdong Province, 528300, China

Correspondence should be addressed to Tianyou Liao; 790163236@qq.com

Received 16 September 2020; Revised 16 November 2020; Accepted 30 November 2020; Published 21 December 2020

Academic Editor: Musthafa M. Essa

Copyright © 2020 Zhili Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objectives.** Prolyl 4-hydroxylase subunit alpha 1 (P4HA1) plays a crucial role in modulating extracellular matrix component and promoting tumor progression by changing tumor adhesion, migration, and other biological behaviors in some cancers. However, its expression pattern, biological function, and underlying mechanism in pancreatic cancer remain largely unclear. **Materials and Methods.** In this study, a set of bioinformatics tools were used to analyze the expression of P4HA1 and its prognostic value in pancreatic cancer. In addition, the mechanism through which P4HA1 promotes the progression of pancreatic cancer was explored by constructing a competing endogenous RNA (ceRNA) regulatory axis. **Results.** It was found that the mRNA and protein expression of P4HA1 was significantly higher in pancreatic cancer tissues than in normal tissues. Its high P4HA1 expression correlated with poor clinicopathological features (T stage:  $P = 0.0078$ ; N stage:  $P = 0.0124$ ; TNM stage:  $P = 0.0013$ ; pathological grade:  $P = 0.0108$ ) and poor prognosis [OS: HR = 1, 95% CI (1-1.01),  $P = 0.00028$ ; DSS: HR = 1, 95% CI (1-1.01),  $P = 0.00049$ ; PFI: HR = 1.01, 95% CI (1.01-1.02),  $P = 0.0057$ ; and DFI: HR = 1, 95% CI (1-1.01),  $P = 0.0034$ ]. The LINC01503/miR-335-5p/P4HA1 axis might mediate the effects of P4HA1 in promoting the progression on pancreatic cancer. **Conclusions.** Collectively, our findings suggest that high expression of P4HA1 may be used as a promising prognostic biomarker and could be considered for the development of a novel therapeutic strategy for pancreatic cancer in the future.

## 1. Introduction

Pancreatic cancer is one of the most malignant tumors with the worst prognosis in the world, and it ranks seventh among the cancer-related deaths in the world and fourth amid all the cancer-related mortalities in the United States; annually, about 432,000 patients die of pancreatic cancer every year<sup>1</sup>. In the recent years, although some progress has been made in the field of combined and precision therapy for pancreatic cancer. However, overall prognosis remains very poor with a 5-year survival rate of less than 6%<sup>2</sup>. This makes the cancer one of the deadliest malignant tumors worldwide.

Prolyl 4-hydroxylase subunit  $\alpha 1$  (P4HA1), located at 10q22.1, is the active catalytic subunit of prolyl 4-hydroxylase (P4H). P4HA1 is responsible for newly synthesized collagen chains and plays a significant role in the formation and stability of the triple helix region of newly synthesized collagen chains. Additionally, it is the key catalytic subunit in regulating collagen biosynthesis<sup>3</sup>. Notably, recent studies found that P4HA1 not only could regulate extracellular collagen synthesis and secretion but also were closely related to the progression of some tumors, such as breast cancer<sup>4</sup>, prostate cancer<sup>5</sup>, and glioma<sup>6, 7</sup> by changing tumor adhesion, migration, and other biological

behaviors. However, the association between P4HA1 and pancreatic cancer as well as its role in the progression of the tumor was not clearly shown.

Human genome-wide and transcriptome research data shows that protein-coding genes account for only a tiny proportion of the genome (1-2%), while most of the genes in the genome are transcribed; they encode large amount of nonprotein coding RNA (ncRNA) much more than we expected<sup>8, 9</sup>. Interestingly, at least 80% of the ncRNA in the human genome is bioactive as opposed to the previous belief that it was simply “junk product”<sup>10</sup>. Based on the length of the transcript, ncRNA is mainly divided into two categories: short-stranded noncoding RNA, such as microRNA (miRNA) (<200 bp), and long noncoding RNA (lncRNA) (>200 bp). In 2011, Salmena et al.<sup>11</sup> proposed a competing endogenous RNA (ceRNA) theory. They postulated that the 3'UTR of lncRNA and mRNA are both rich in miRNA recognition elements (MRE). Moreover, lncRNA could adsorb miRNA through MRE and competitively bind miRNA to inhibit its activity. This happened in order to partially remove the inhibitory effect of miRNA on its target genes and upregulate their expression. Therefore, miRNA and lncRNA play a molecular regulatory role in many processes of gene expression, participate in the pathophysiological process of tumors, and are closely related to the occurrence and progression of cancer<sup>12</sup>. Moreover, the mechanism of the ceRNA of P4HA1 affecting the malignant progression of pancreatic cancer has not been reported.

Therefore, the expression and clinical value of P4HA1 in pancreatic cancer were comprehensively analyzed in this study. Additionally, a ceRNA regulatory axis was constructed to explore the potential mechanism of P4HA1 in the progression of pancreatic cancer.

## 2. Materials and Method

**2.1. Pan-Cancer Analysis of P4HA1 in the Oncomine Database.** Oncomine (<http://www.oncomine.org/resource/login.html>) is an open-access cancer microarray database and an integrated data mining platform that facilitates the discovery of valuable clues from genome-wide expression analysis. So far, the database has 19 cancer types and contains 92 datasets<sup>13</sup>. In this study, Oncomine was used to compare the 19 cancer samples with normal samples and identify the expression of P4HA1 in various tumors. At the same time, the expression levels of P4HA1 mRNA (log<sub>2</sub> transformed) in pancreatic cancer and normal tissues were examined in order to determine the transcription level of P4HA1 in the disease ( $P < 0.001$ , |fold change| >2, and the top 10% of the gene ranking as the threshold condition).

**2.2. Pan-Cancer Analysis of P4HA1 in the Sangerbox Database.** The Sangerbox database (<http://sangerbox.com/Tool>) is a comprehensive Chinese bioinformatics analysis platform. The pan-cancer analysis tool of this database was used to explore the expression of P4HA1 in pancreatic cancer and normal samples (\* $P < 0.05$ , \*\* $P < 0.01$ , and \*\*\* $P < 0.001$ ). In addition, this analysis platform was also used to investigate the effect of P4HA1 on the survival and prognosis of

pancreatic cancer. This included an analysis of overall survival (OS) and disease-specific survival (DSS). A log-rank  $P$  value < 0.05 was considered to be statistically significant.

**2.3. Expression, Prognosis, Coexpression, and Correlation Analysis of P4HA1 in the GEPIA Database.** Gene Expression Profiling and Interactive Analysis (GEPIA) (<http://gepia.cancer-pku.cn/index.html>) is an interactive platform including 9,736 tumors and 8,587 normal samples from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) projects. It helps in the analysis of RNA sequencing expression data<sup>14</sup>. GEPIA provides customizable functions, such as tumor/normal differential expression analysis, patient survival analysis, similar gene detection, and correlation analysis. The database uses the log-rank (Mantel-Cox) test in the prognosis of 33 different types of cancer based on gene expression. Moreover, the box plot was used to display the ordinate of gene expression in tumor/normal differential expression analysis. The expression data was converted by log<sub>2</sub> (TPM+1), |log<sub>2</sub> FC| >1, and  $P$  value < 0.05 was set as the truncation criterion. Finally, gene expression correlation analysis of the given TCGA expression dataset was used to determine the correlation coefficient using the Pearson method.

**2.4. The Expression of P4HA1 in the HPA Database.** The Human Protein Atlas (HPA) (<http://www.proteinatlas.org/>) provides tissue and cellular distribution information for 26,000 types of human proteins<sup>15</sup>. In this database, the researchers used highly specific antibodies and immunodetection techniques (immunoblotting, immunofluorescence, and immunohistochemistry) to identify in detail the expression of each protein in 64 cell lines, 48 normal human tissues, and 20 tumor tissues. The immunohistochemical staining data from the Tissue and Pathology section of the HPA was used to detect the expression level of the P4HA1 protein in human pancreatic cancer and normal pancreatic tissues.

**2.5. TCGA-PAAD Data Download and Data Analysis in UCSC Xena.** The University of California Santa Cruz (UCSC) Xena browser (<https://xenabrowser.net/>) was used to obtain detailed information on mRNA expression of genes. In addition, the browser was employed to source clinical information on pancreatic adenocarcinoma (PAAD) patients in the TCGA cohort (Project ID: TCGA-PAAD)<sup>16, 17</sup>. Moreover, the relationship between the expression of P4HA1 and clinicopathological parameters (gender, T stage, N stage, M stage, TNM stage, and pathological grade) in pancreatic cancer was analyzed using the TCGA-PAAD data downloaded from Xena.

**2.6. Screening for Genes Coexpressed with P4HA1.** Screening for genes coexpressed with P4HA1 was done using the GEPIA and UALCAN databases. Afterwards, the same coexpression genes in the two databases were screened and a Venn diagram was drawn using the intersection method.

The UALCAN database is a user-friendly and interactive network resource, which provides a convenient way of publicly obtaining cancer transcriptome data from TCGA<sup>18</sup>. In this study, the database was used to find genes coexpressed

with P4HA1 in PAAD. Additionally, expression of some genes/miRNAs was analyzed.

**2.7. GO Analysis and KEGG Pathway Analysis of Coexpressed Genes.** Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>) is a comprehensive gene set enrichment analysis web server<sup>19, 20</sup>. It was used to perform Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis on P4HA1 and its coexpressed genes. The top 10 rich enrichment projects and paths were displayed on the webpage and could be downloaded directly.

**2.8. PPI Analysis of Coexpressed Genes and Hub Genes Were Identified.** In this study, the String database (<http://string-db.org>)<sup>21</sup> and Cytoscape (version 3.4.0) software<sup>22</sup> were used to construct a protein-protein interaction (PPI) network of coexpressed genes. The Hub gene identifying plug-in cytoHubba (version 0.1) is a Cytoscape application that uses several topological algorithms to predict and identify important nodes and subnets in a given network. Moreover, Molecular Complex Detection (MCODE) (version 1.6.1) is an application plug-in for Cytoscape that is used to cluster a given network according to topology, in order to find closely connected areas. The cytoHubba and MCODE were used to identify the most important genes and module in the PPI network, respectively.

**2.9. Expression, Correlation, and Pathway Enrichment Analysis of P4HA1 and Top 15 Hub Genes.** The TCGA data downloaded from the Xena browser was used to show the expression patterns of P4HA1 and obtained Hub genes in a heat map. Additionally, correlation analysis and KEGG enrichment analyses were performed on these genes. The results of gene enrichment analysis could intuitively show which genes were enriched in each pathway in a circos diagram.

**2.10. Analysis of Three Hub Genes Directly Interacting with P4HA1.** The String database was used to construct a PPI network between P4HA1 and its top 15 coexpressed Hub genes. The three Hub genes directly interacting with P4HA1 were screened out, and a PPI network was performed. Afterwards, the three screened Hub genes were analyzed for expression and correlation using the GEPIA database.

**2.11. NetworkAnalyst Database.** The NetworkAnalyst (<http://www.networkanalyst.ca/faces/home.xhtml>) is an online website for comprehensive gene expression analysis, meta-analysis, and biological network regulation analysis<sup>23, 24</sup>. To further explore the mechanism of P4HA1 underlying the prognosis of pancreatic cancer, the Gene-miRNA Interactome module of the NetworkAnalyst database was used to predict the upstream miRNAs that could regulate P4HA1. The predicted data originated from comprehensive experimental verifications collected in TarBase and miRTarBase miRNA-Gene interaction information.

**2.12. Predictive miRNA Expression, Correlation, and Prognostic Analysis.** The UALCAN database was used to compare the

expression of 5 miRNAs between normal and pancreatic cancer tissues. Moreover, the LinkedOmics database (<http://www.linkedomics.org/login.php>)<sup>25</sup> was used to analyze the correlation between 5 miRNAs and the expression level of P4HA1 in pancreatic cancer. It is well known that a negative regulatory relationship exists between miRNA and targeted mRNA. Therefore, miRNA that negatively correlated to P4HA1 expression was considered to have good research prospects.

The Kaplan-Meier Plotter (<http://kmplot.com/analysis/>) is an online database established from gene expression data and survival information of patients with cancer<sup>26</sup>. Using TCGA data, this database was utilized to evaluate the prognostic value of the predicted miRNAs in PAAD. The miRNA name was keyed into the database, after which its Kaplan-Meier survival plots, hazard ratio (HR), 95% confidence interval (CI), and log-rank *P* values were displayed directly on the web page.

**2.13. Analysis of Upstream lncRNA Targeted to miRNA.** Here, the StarBase (<http://starbase.sysu.edu.cn/index.php>)<sup>27, 28</sup> and the miRNet (<https://www.mirnet.ca/>) databases were introduced to predict lncRNAs that could bind miR-335-5p. Subsequently, the lncRNAs that predicted from the StarBase were intersected with the predicted from miRNet for miR-335-5p. According to the regulatory pattern of ceRNA, there is a negative correlation between the expression of lncRNA and miRNA and a positive correlation with the expression of target genes. Based on this feature, the study further screened for the most likely regulatory lncRNA targeted to miR-335-5p by correlation analysis. The StarBase was used to analyze the correlation between lncRNA-miRNA and lncRNA-P4HA1.  $R > |0.1|$  and  $P < 0.05$  were set as the critical criteria for identifying important miRNA-lncRNA and P4HA1-lncRNA pairs.

**2.14. A ceRNA Regulation Axis and Brief Mechanism Diagram of P4HA1.** The study combined the coexpressed Hub genes directly interacting with P4HA1 obtained from above process, upstream miRNA for P4HA1 and upstream lncRNA for miRNA to construct a ceRNA regulatory lncRNA-miRNA-P4HA1 network diagram. Additionally, previous relevant reports from the literature were combined to draw a brief model diagram of the mechanism by which P4HA1 promotes the progression in pancreatic cancer.

**2.15. Statistical Analysis.** By default, all statistical analyses were performed as described in the web resources. Briefly, the *t*-test was performed to compare the expression of P4HA1 mRNA in the Oncomine database. The *t*-test was also used to perform comparative analysis of P4HA1 expression among different subgroups for each clinicopathological feature. The log-rank test was used to calculate the *P* value for survival data in the Sangerbox, GEPIA, and Kaplan-Meier Plotter databases. Furthermore, the Fisher's exact test was used to measure gene enrichment in annotated items of the Enrichr annotation system. Finally, linear correlation between mRNA-mRNA, miRNA-mRNA, and lncRNA-

miRNA was measured using Pearson's correlation coefficient in the GEPIA, LinkedOmics, and StarBase databases.  $P < 0.05$  was considered statistically significant (\* $P < 0.05$ , \*\* $P < 0.01$ , and \*\*\* $P < 0.001$ ).

### 3. Results

#### 3.1. The Expression and Prognosis of P4HA1 in Pan-Cancer.

The overall expression profile of P4HA1 in various cancers was first analyzed using the OncoPrint database by making comparisons between cancer and normal tissue samples. P4HA1 was markedly overexpressed in most tumors, especially in kidney and pancreatic cancers (Figure 1(a)). Afterwards, TCGA and GTEx data were used to analyze the expression of P4HA1 in pan-cancer, in order to further quantitatively verify its expression in various tumors. Figure 1(b) shows that there was a significant difference in the expression of P4HA1 between normal and tumor samples of most tissues except for the Kidney Chromophobe (KICH). Furthermore, P4HA1 was overexpressed in most tumors. Data from TCGA was also used to analyze the impact of high P4HA1 expression on the OS (Figure 1(c)) and DSS (Figure 1(d)) of patients. This was done in order to understand further the impact of P4HA1 on the prognosis of survival from each tumor. Figures 1(c) and 1(d) show that the expression of P4HA1 affects the OS and DSS of patients with PAAD.

In the above preliminary investigation, P4HA1 was significantly overexpressed in pancreatic cancer. Additionally, it was related to poor prognosis of survival in patients with PAAD warranting further study.

**3.2. Expression of P4HA1 in Pancreatic Cancer.** The analysis of tumor vs. normal samples in different datasets from OncoPrint showed that P4HA1 was significantly overexpressed in multiple pancreatic cancer datasets (Figures 2(a)–2(e)). The GEPIA database was also used to further confirm that P4HA1 was significantly highly expressed in the TCGA-PAAD dataset (Figure 2(f)). In addition, the HPA database showed that immunohistochemical staining of cancer tissues was deeper than that of normal pancreatic tissues. This suggests that there was remarkable expression of the P4HA1 protein in pancreatic cancer tissues.

The study further explored the association between P4HA1 expression and different clinical-pathological features in patients with PAAD (Figure 3). The results showed that there was no significant difference in the expression of P4HA1 between male and female patients as shown in Figure 3(a) ( $P = 0.5868$ ). Additionally, there was no significant difference between the positive and the negative metastasis groups as highlighted in Figure 3(d) ( $P = 0.9642$ ). In terms of T stage (T1-2 vs. T3-4), N stage (N- vs. N+), TNM stage (I vs. II-III), and pathological grade (G1 vs. G2-4), expression of P4HA1 was significantly high in the subgroups with high stage/grade (T stage:  $P = 0.0078$ ; N stage:  $P = 0.0124$ ; TNM stage:  $P = 0.0013$ ; pathological grade:  $P = 0.0108$ ). Therefore, these results indicate that P4HA1 expression was associated with unfavorable clinicopathological

characteristics in patients with pancreatic cancer and might play a detrimental role in pancreatic cancer.

All in all, the findings show that P4HA1 is a reliable and promising biomarker of pancreatic cancer.

**3.3. Prognostic Significance of P4HA1 Gene in Pancreatic Cancer.** Three online databases including Sangerbox and GEPIA were used to further evaluate the prognostic value of P4HA1 in human pancreatic cancer (Figure 4). Patients with high P4HA1 expression showed poor OS [OS: HR = 1, 95% CI (1-1.01),  $P = 0.00028$ ], DSS [DSS: HR = 1, 95% CI (1-1.01),  $P = 0.00049$ ], progress-free interval (PFI) [PFI: HR = 1.01, 95% CI(1.01-1.02),  $P = 0.0057$ ], and disease-free interval (DFI) [DFI: HR = 1, 95% CI (1-1.01),  $P = 0.0034$ ] as shown in Figures 4(a)–4(d) (SangerBox). Results of the GEPIA database analysis similarly indicated that the high expression of P4HA1 indicated poor prognosis in pancreatic cancer patients [OS: HR = 1.6,  $P = 0.031$ ; DFS: HR = 1.7,  $P = 0.014$ ] (Figures 4(e) and 4(f)). All these findings showed that P4HA1 was associated with poor prognosis in patients with pancreatic cancer. It could therefore be used as a predictor of poor prognosis in the disease.

**3.4. GO and KEGG Pathway Analysis of Genes Coexpressed with P4HA1.** Identifying genes coexpressed with P4HA1 would help to better understand the potential function of the protein in pancreatic cancer. Therefore, the UALCAN and GEPIA databases were used to identify these coexpressed genes, as shown in Figure 5(a). A total of 139 coexpressed genes were predicted by both databases. Afterwards, the GO enrichment (Figures 5(b)–5(d)) and KEGG pathway enrichment analyses were performed on the genes (Figure 5(e)). Additionally, the Enrichr database was used to perform the KEGG pathway enrichment and functional annotation analyses including three types of GO analysis: biological processes (BP), cellular components (CC), and molecular functions (MF).

The GO functional enrichment of these coexpressed genes in the BP category included the following: the glycolytic process through glucose-6-phosphate (GO:0061620), canonical glycolysis (GO:0061621), glucose catabolic process to pyruvate (GO:0061718), glycolytic process (GO:0006096), and generation of ATP from ADP (GO:0006757). In the CC category, it included focal adhesion (GO:0005925), endoplasmic reticulum lumen (GO:0005788), platelet alpha granule (GO:0031091), actin cytoskeleton (GO:0015629), and platelet alpha granule membrane (GO:0031092). Finally, the MF group included the following: cadherin binding (GO:0045296), protein-lysine 6-oxidase activity (GO:0004720), dolichyl-diphosphooligosaccharide-protein glycotransferase activity (GO:0004579), oligosaccharyl transferase activity (GO:0004576), and oxidoreductase activity, acting on the CH-NH2 group of donors and oxygen as an acceptor (GO:0016641). The top 10 KEGG pathways were as follows: glycolysis/gluconeogenesis, the HIF-1 signaling pathway, central carbon metabolism in cancer, fructose and mannose metabolism, the PI3K-Akt signaling pathway, the Pentose phosphate pathway, proteoglycans in cancer, starch and sucrose metabolism, protein processing in the

Cancer Summary for P4HA1

Analysis Type by cancer	Cancer vs. Normal	
Bladder cancer		
Brain and CNS cancer	3	
Breast cancer	1	
Cervical cancer	1	
Colorectal cancer	2	
Esophageal cancer	2	
Gastric cancer		
Head and Neck cancer	3	
Kidney cancer	4	
Leukemia		
Liver cancer		
Lung cancer		
Lymphoma	1	
Melanoma		
Myeloma		
Other cancer	1	
Ovarian cancer		
Pancreatic cancer	4	
Prostate cancer		
Sarcoma	2	
Significant unique analyses	23	
Total unique analyses	448	

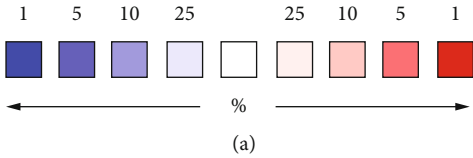
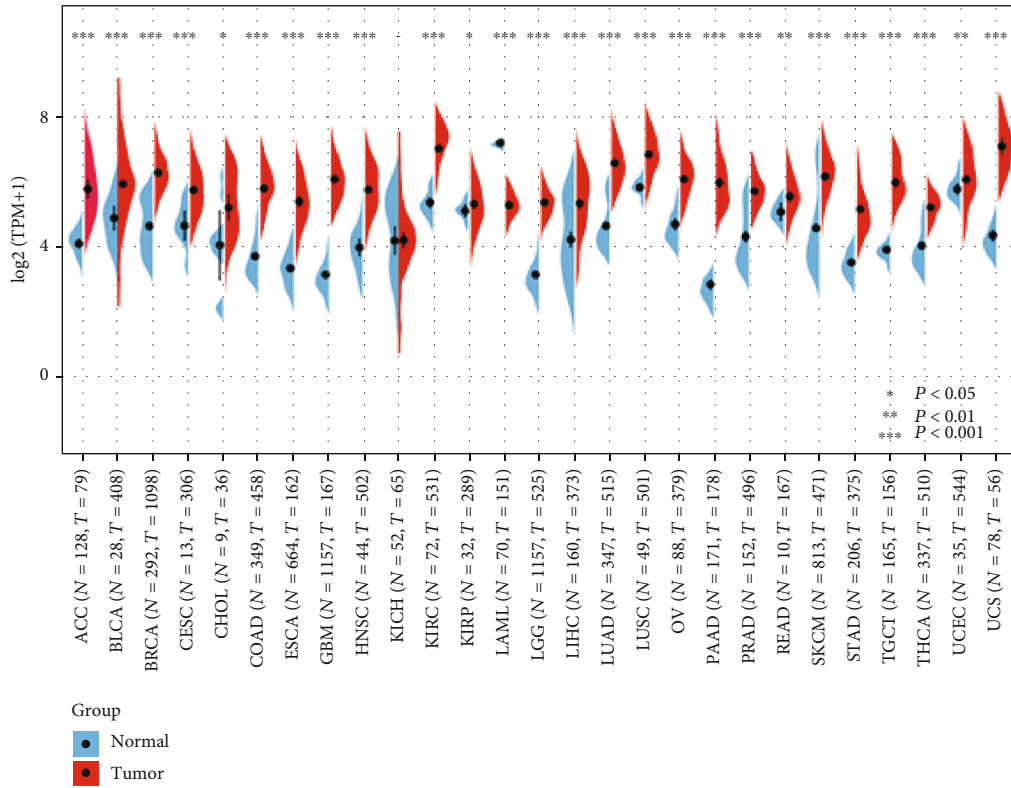
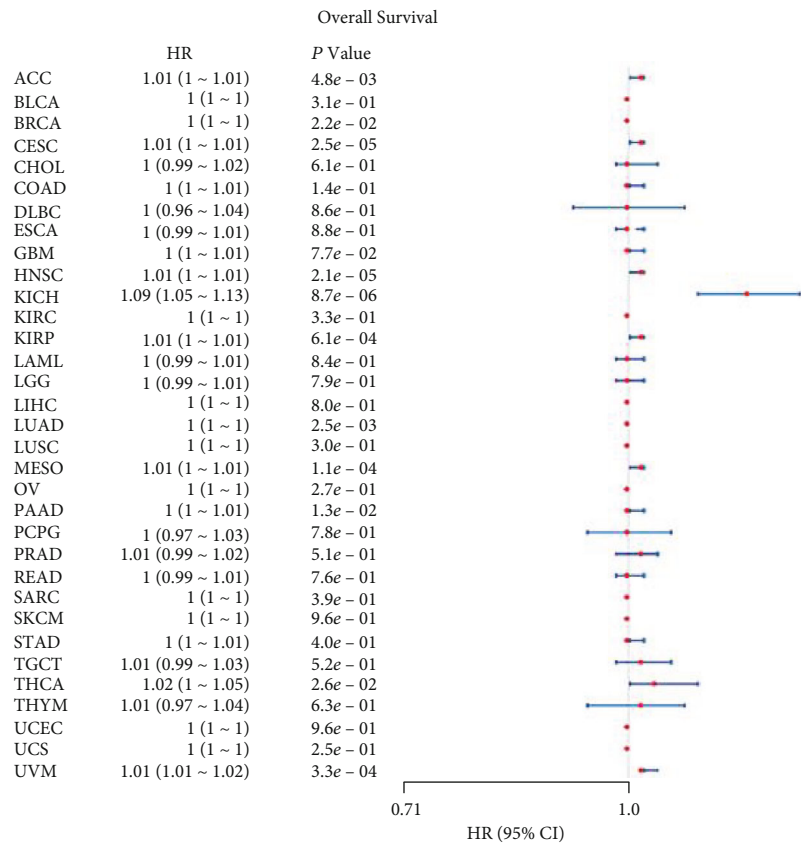


FIGURE 1: Continued.



(b)



(c)

FIGURE 1: Continued.

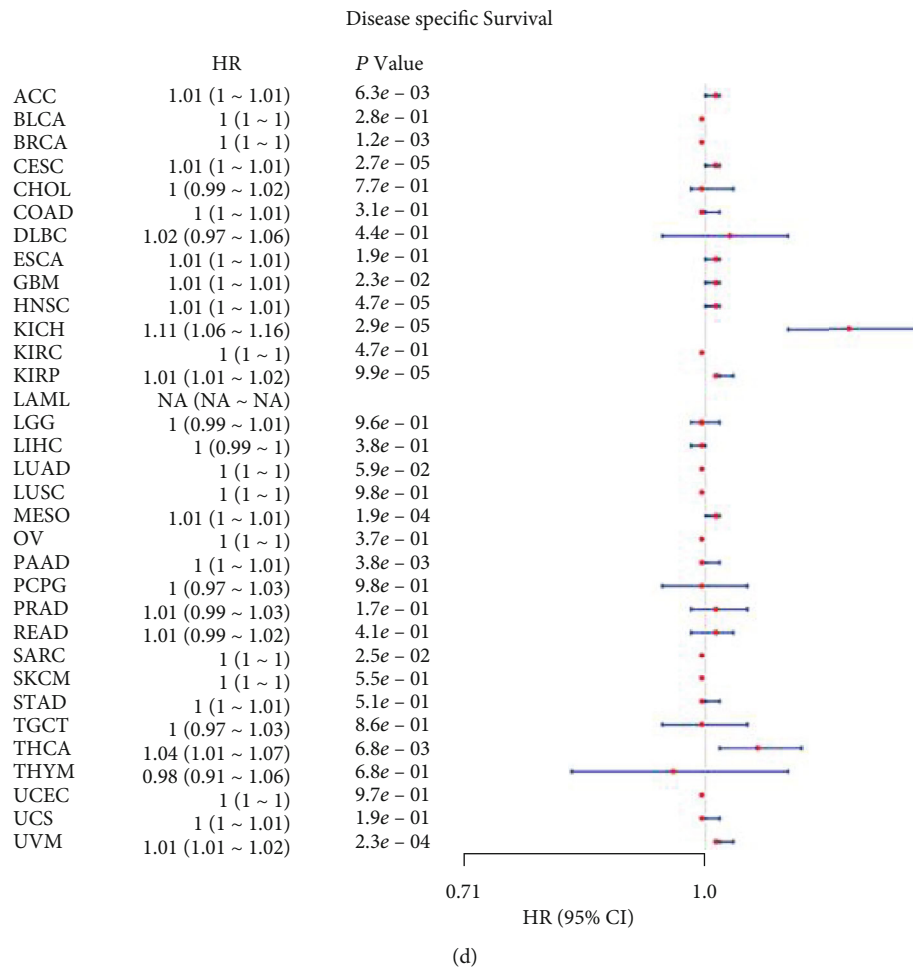


FIGURE 1: Expression pattern and prognostic significance of P4HA1 in human pan-cancer. (a) Expression of P4HA1 in various cancers by Oncomine analysis of cancer samples versus normal samples. (b) Comparison of P4HA1 expression between pan-cancer samples and paired normal samples based on TCGA and GETx data. (c, d) The effect of P4HA1 expression on the overall survival (OS) and disease-specific survival (DSS) of pan-cancer based on TCGA data, respectively.

endoplasmic reticulum, and neomycin, kanamycin, and gentamicin biosynthesis (Figure 5).

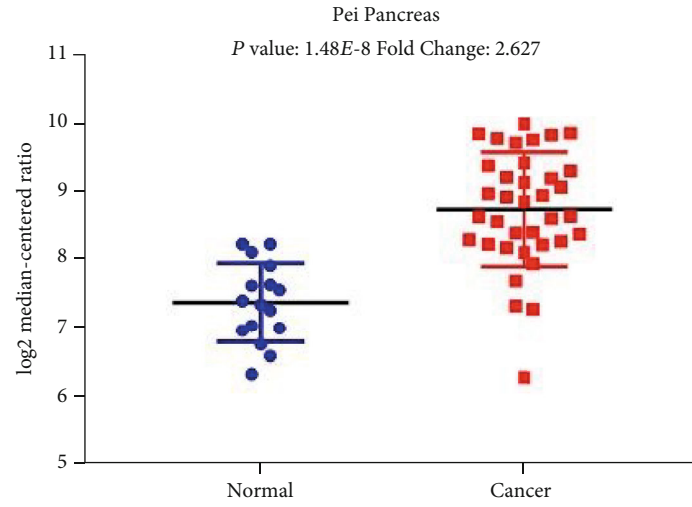
**3.5. Identification and Analysis Core Genes Coexpressed with P4HA1.** A PPI network of the coexpressed genes was constructed using both the STRING database and Cytoscape software. Initially, the CytoHubba plug-in in the Cytoscape software was used to identify the top 15 node genes coexpressed with P4HA1 and were highlighted as potential key genes (Hub genes), based on the score of the degree among nodes. The genes included GAPDH, FN1, VEGFA, HSPA5, ENO1, PGK1, HSP90B1, TPI1, LDHA, HK1, HK2, ALDOA, PFKP, SLC2A1, and HSP90AA1 (Figure 6(a)). Afterwards, the MCODE plug-in was used to obtain the most important interaction module (Figure 6(b)). The results showed that Hub genes identified by the two methods were generally the same.

The expression heat map of P4HA1 and the Hub genes was displayed using the UCSC Cancer Genomics Browser. The findings showed that the expression pattern of 10 Hub genes was consistent with that of P4HA1 in pancreatic cancer (Figure 6C). Additionally, correlation heat maps showed that

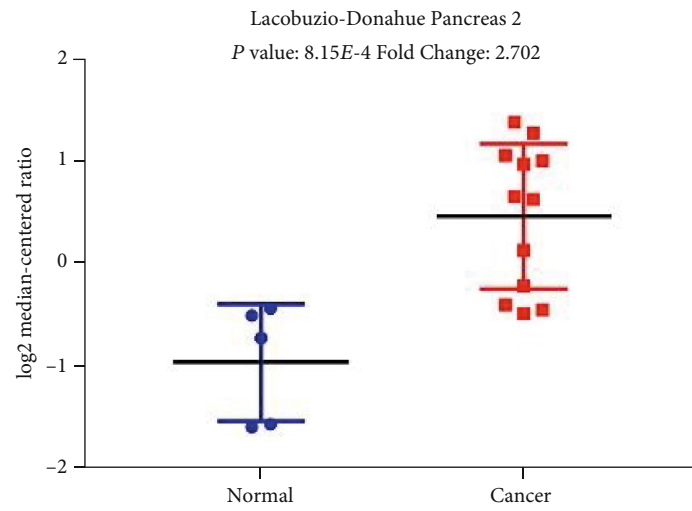
the expression of P4HA1 was positively associated with the expression patterns of these Hub genes (Figure 6(d)). Furthermore, the KEGG circus diagram of the top 15 Hub genes showed that they were mainly involved in the HIF-1 signaling pathway, the PI3K-Akt signaling pathway, glycolysis/gluconeogenesis, central carbon metabolism in cancer, proteoglycans in cancer, protein processing in the endoplasmic reticulum, fructose and mannose metabolism, and the pentose phosphate pathway (Figure 6(e)).

Moreover, P4HA1 and the top 15 Hub genes were subjected to PPI analysis (Figure 7(a)). Hub genes that directly interacted with P4HA1 were selected, including PGK1, ALDOA, and LDHA (Figure 7(b)). Expression level and correlation analyses showed that these three genes were significantly overexpressed in pancreatic cancer (Figure 7(c)) and had a positive correlation with P4HA1 (Figure 7(d)). Therefore, the results suggested that PGK1, ALDOA, and LDHA may play important synergistic roles for P4HA1 in pancreatic cancer.

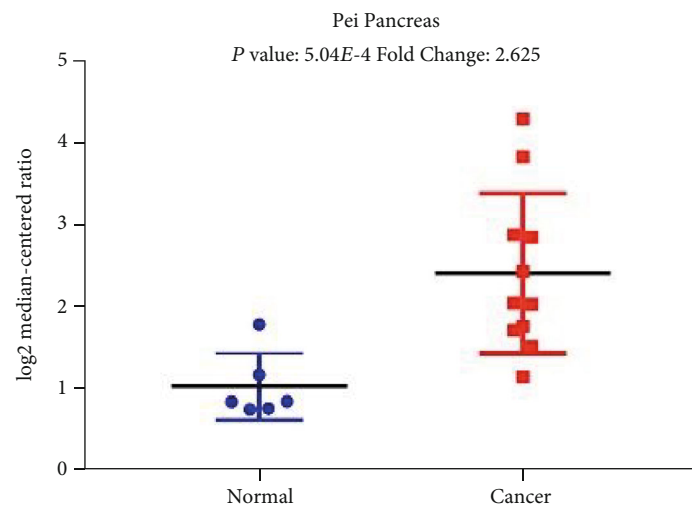
**3.6. Prediction miRNA, lncRNA, and Building a ceRNA Regulatory Network.** In order to further explore the



(a)



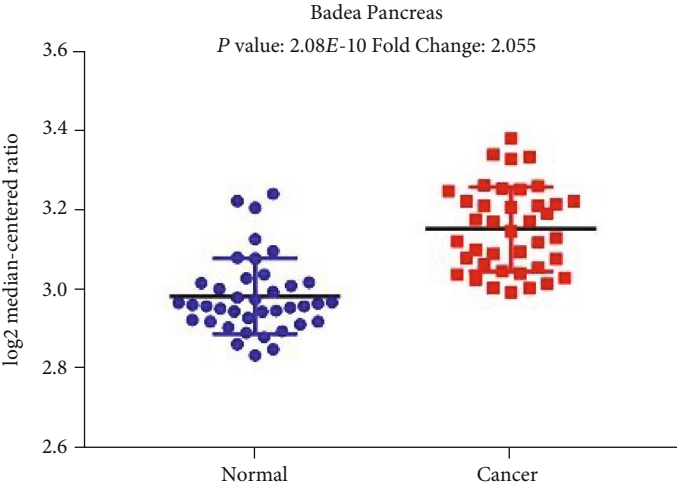
(b)



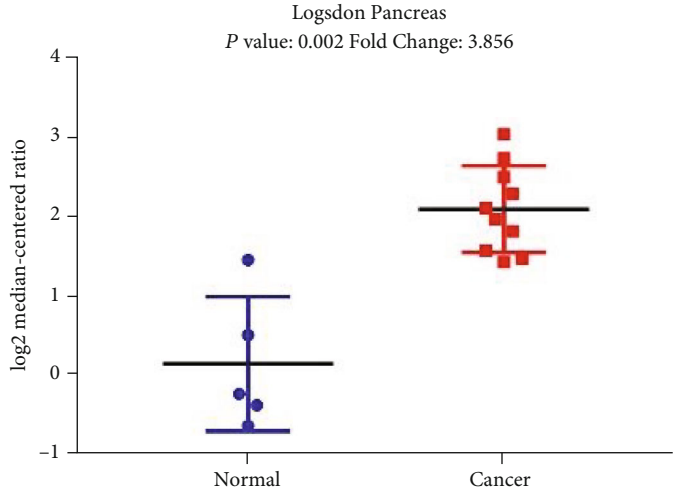
(c)

FIGURE 2: Continued.

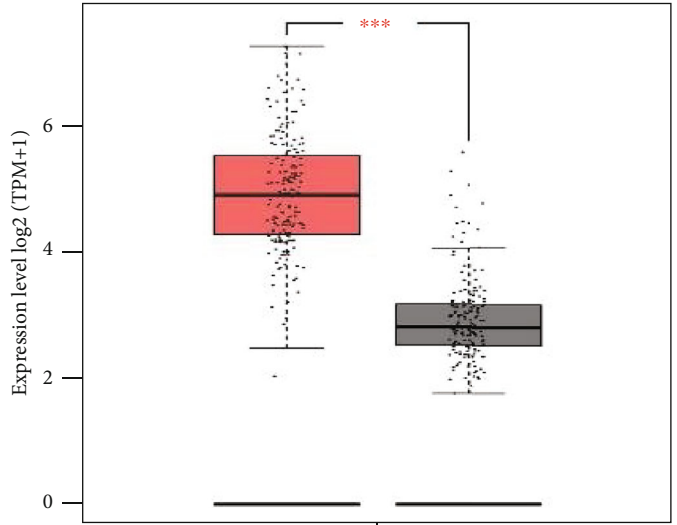




(d)



(e)



(num (T) = 179; num (N) = 171)

(f)

FIGURE 2: Continued.

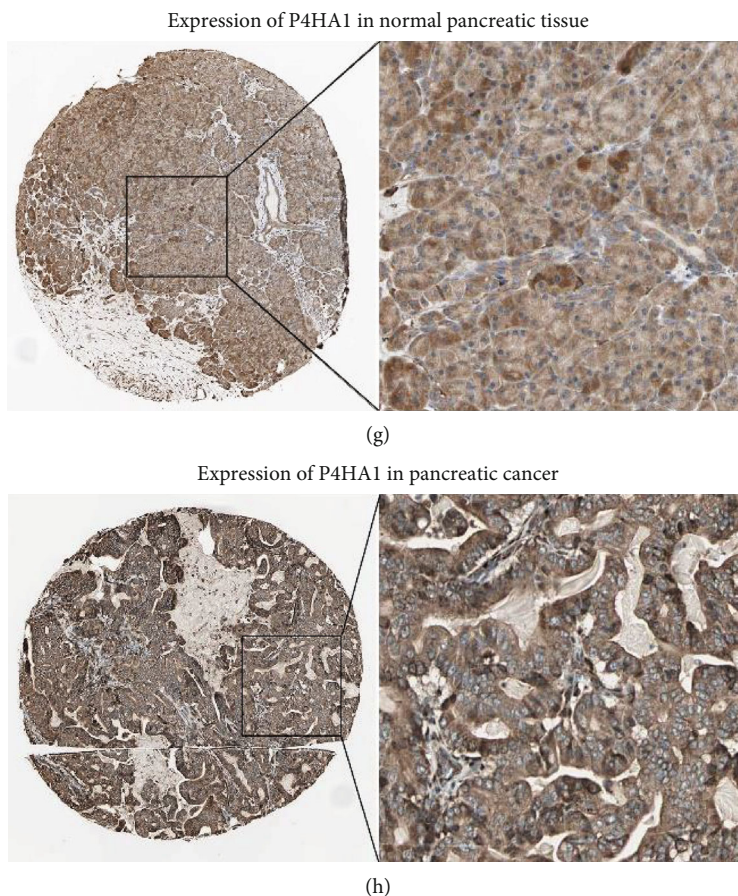


FIGURE 2: P4HA1 expression in normal pancreatic tissues and pancreatic adenocarcinoma (PAAD) tissues. (a–e) P4HA1 mRNA expression in four pancreatic cancer studies based on analysis of Oncomine database. (f) Expression of P4HA1 in PAAD tissues and normal pancreatic tissues based on TCGA and GETx datasets using the GEPIA database. Error bars represent SD. \*\*\* $P < 0.001$ . (g, h) Immunohistochemical staining of P4HA1 protein, respectively, showed the protein expression levels in normal pancreatic tissues and pancreatic cancer tissues based on data from the HPA database.

mechanism of P4HA1 in the prognosis of pancreatic cancer, the potential mechanism of noncoding RNA regulating P4HA1 was explored. First, NetworkAnalyst was used to predict the upstream miRNAs that could target and regulate P4HA1, including hsa-mir-145-5p, hsa-mir-320a, hsa-mir-335-5p, hsa-mir-877-3p, hsa-mir-122-5p, and hsa-mir-941 (Figure 8(e)). Afterwards, the UALCAN database, TCGA-based data, was used to analyze the expression of these miRNAs in pancreatic cancer (hsa-mir-941 was not reported in TCGA). The results showed that hsa-mir-320a and hsa-mir-335-5p had significantly low expression in pancreatic cancer (Figures 8(b) and 8(c)). Additionally, the LinkedOmics database was used to analyze the correlation between the five predicted miRNAs and P4HA1. The results showed that only miR-335-5p had a significant negative correlation with P4HA1 (Figure 8(h)). Furthermore, the Kaplan-Meier Plotter was used to analyze the effects of these miRNAs on the OS of patients with pancreatic cancer. The findings revealed that miR-122-5p was a poor prognostic marker [HR = 1.64, 95% CI (1.00-2.68),  $P = 0.046$ ] while miR-335-5p was a protective factor in patients with pancreatic cancer [HR = 0.53, 95% CI (0.34-0.83),  $P = 0.0049$ ]. However, the other three miRNAs had no effect on the overall survival of

patients (Figures 8(k) and 9(l)). These results suggest that miR-335-5p was the most promising upstream miRNA regulating P4HA1.

Moreover, StarBase and miRNet were used to predict upstream lncRNA targeting to bind miR-335-5p. 31 lncRNAs were predicted from the StarBase database (Figure 9(a)) and 34 lncRNAs from the miRNet database (Figure 9(b)). 18 lncRNAs were predicted by both the two databases including SLFNL1-AS1, GAS5, MIR29B2CHG, SNHG8, TMEM161B-AS1, LINC01503, KCNQ1OT1, LINC00294, NEAT1, LINC00943, OIP5-AS1, SNHG20, URB1-AS1, TUG1, XIST, FTX, LINC00630, and LINC00893 (Figure 9(c)). The StarBase was then used to analyze the correlation between the predicted lncRNAs and miR-335-5p and the correlation between lncRNAs and P4HA1. As shown in Table 1, according to the regulatory characteristics of ceRNA theory, LINC01503 was screened out because it was negatively correlated with P4HA1 ( $r = 0.152$ ,  $P = 4.30E - 02$ ) (Figure 9(d)) and positively correlated with miR-335-5p ( $r = 0.307$ ,  $P = 3.13E - 05$ ) (Figure 9(e)). In addition, the expression of LINC01503 in PAAD was further verified on the GEPIA database. It was confirmed that LINC01503 was significantly overexpressed in PAAD (Figure 9(f)).

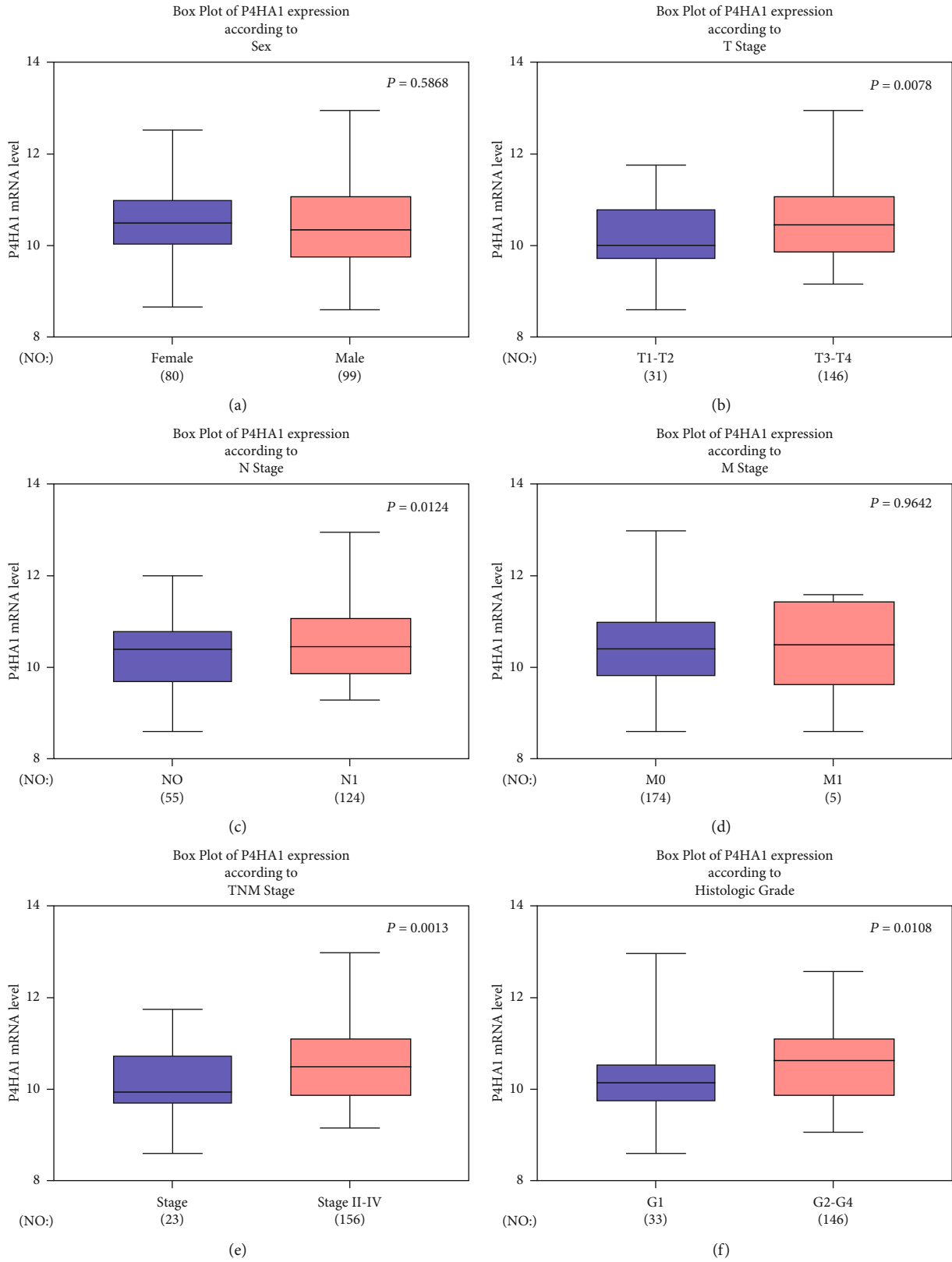
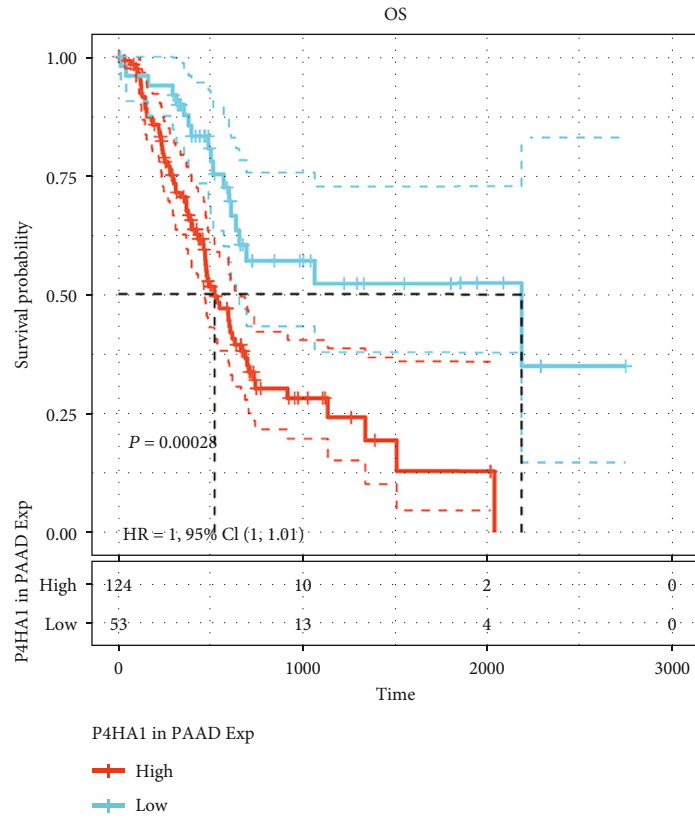
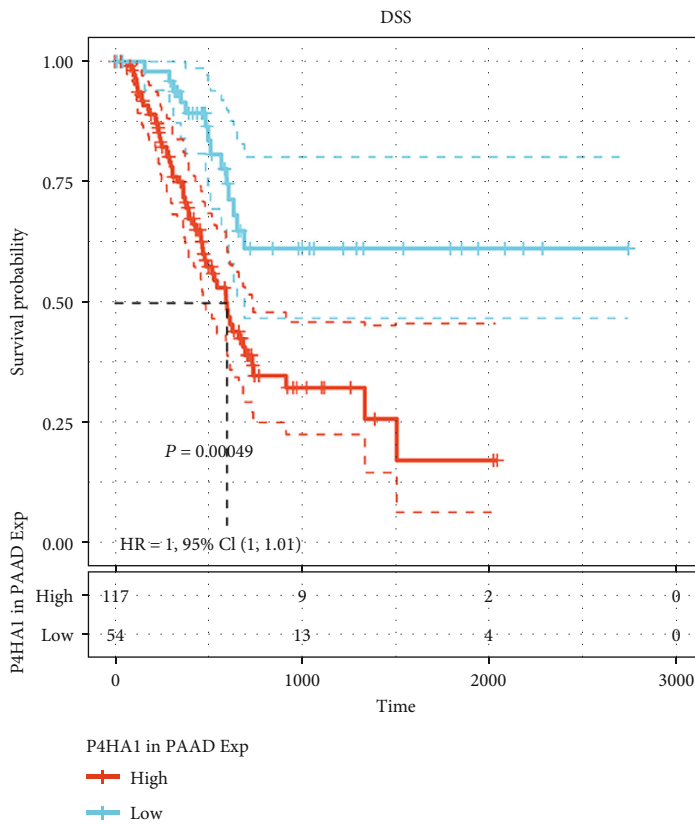


FIGURE 3: Association between P4HA1 expression and various clinical-pathological parameters in pancreatic cancer based on data from the TCGA. (a) A boxplot showing expression of P4HA1 based on gender. (b) A boxplot showing P4HA1 expression according to T stage. (c) A boxplot showing expression of P4HA1 according to nodal status. (d) A boxplot showing expression of P4HA1 according to metastatic status. (e) A boxplot showing expression of P4HA1 according to TNM stage. (f) A boxplot showing expression of P4HA1 according to histological grade.

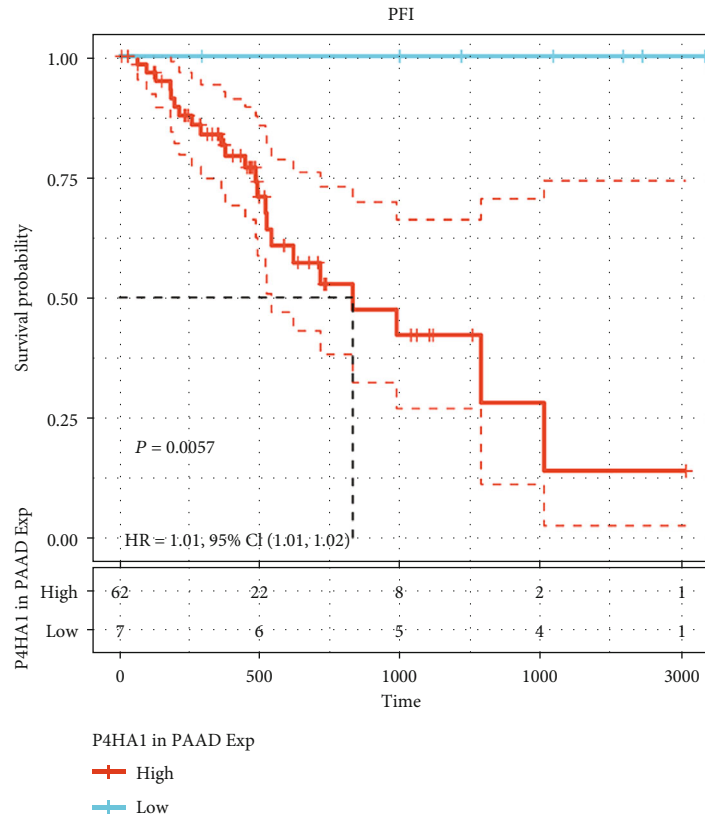


(a)

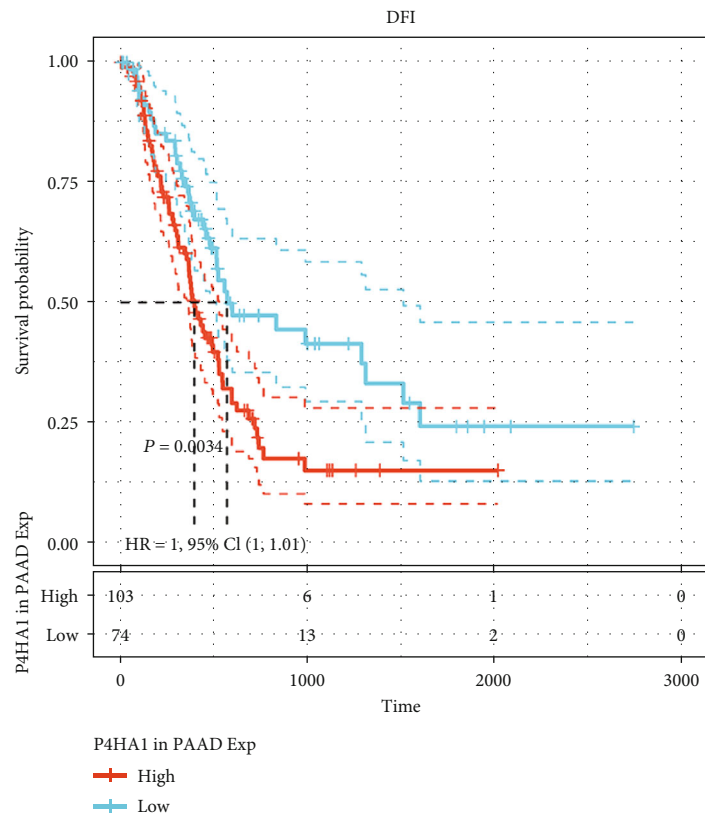


(b)

FIGURE 4: Continued.

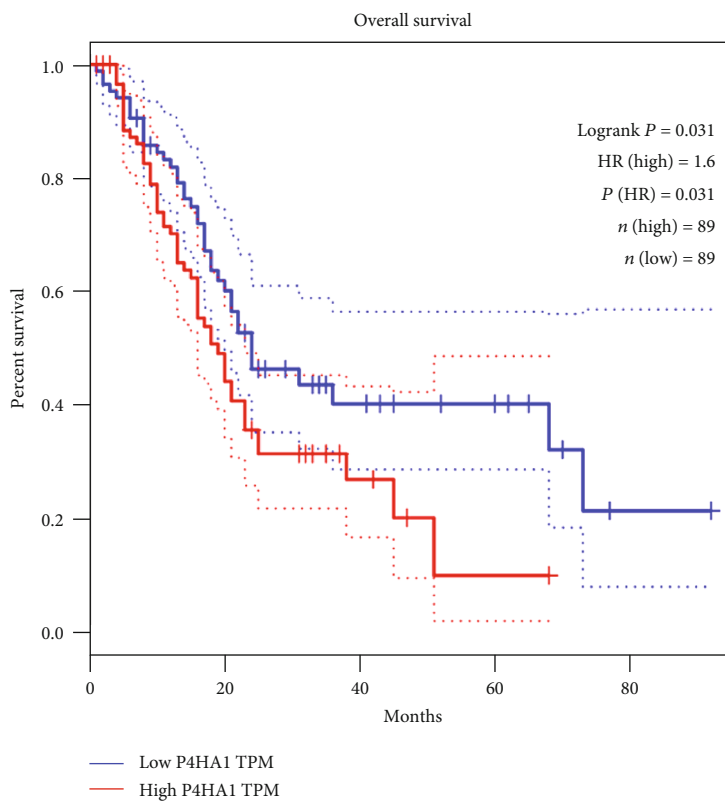


(c)

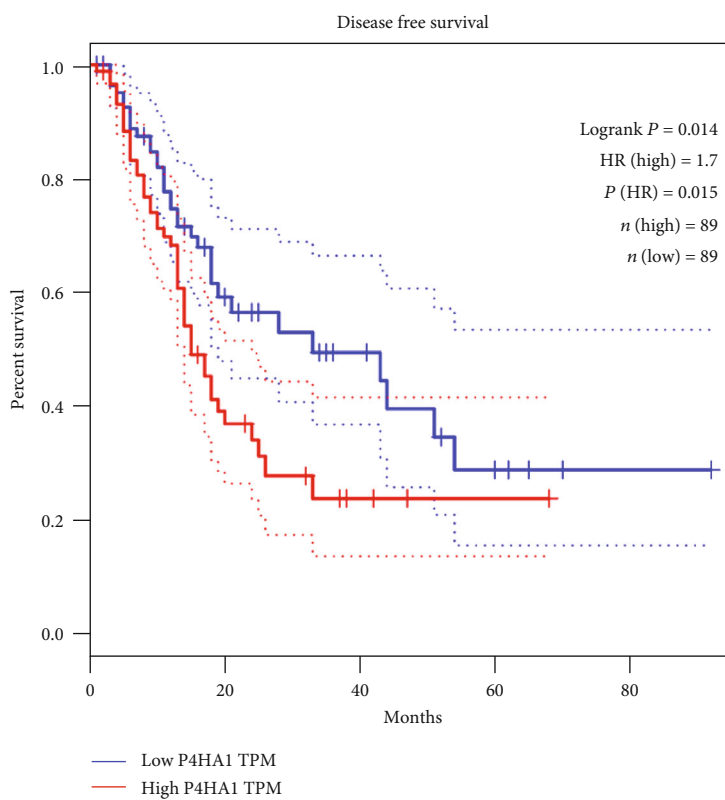


(d)

FIGURE 4: Continued.

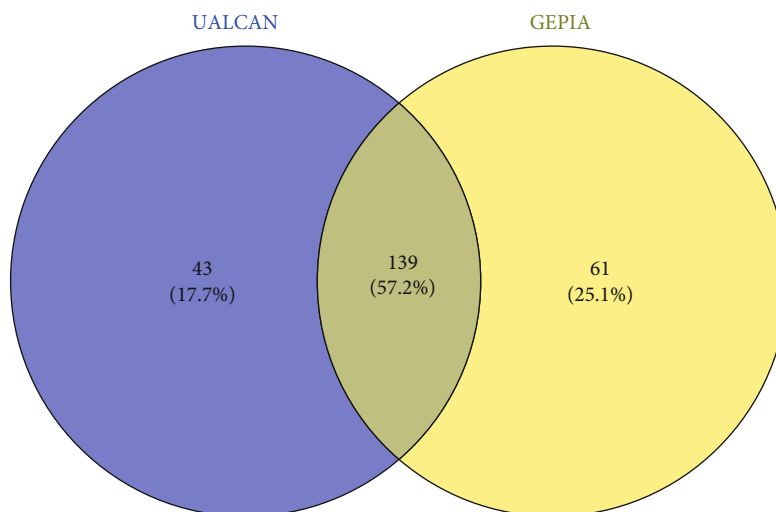


(e)



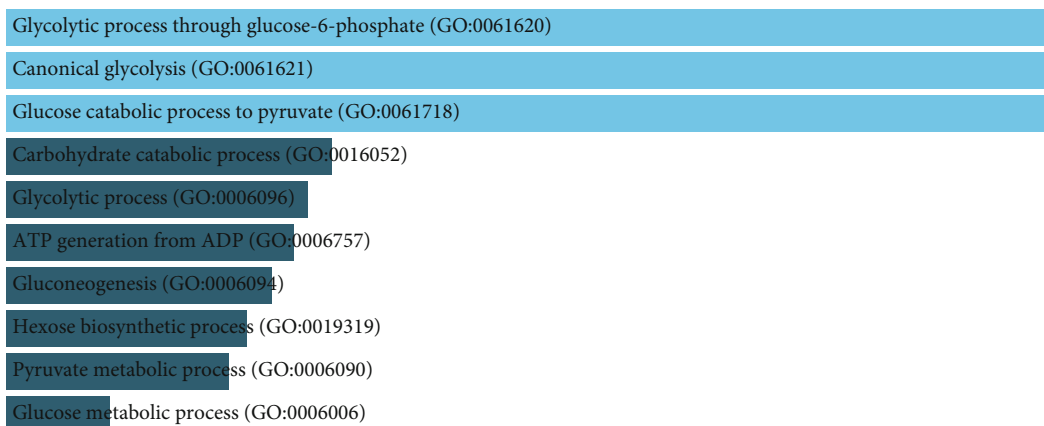
(f)

FIGURE 4: Prognostic value of P4HA1 in pancreatic cancer. (a–d) High expression of P4HA1 correlated with poor OS, DSS, PFI, and DFI, respectively, in PAAD (TCGA database). (e, f) High expression of P4HA1 correlated with poor OS and DFS of PAAD (GEPIA database).



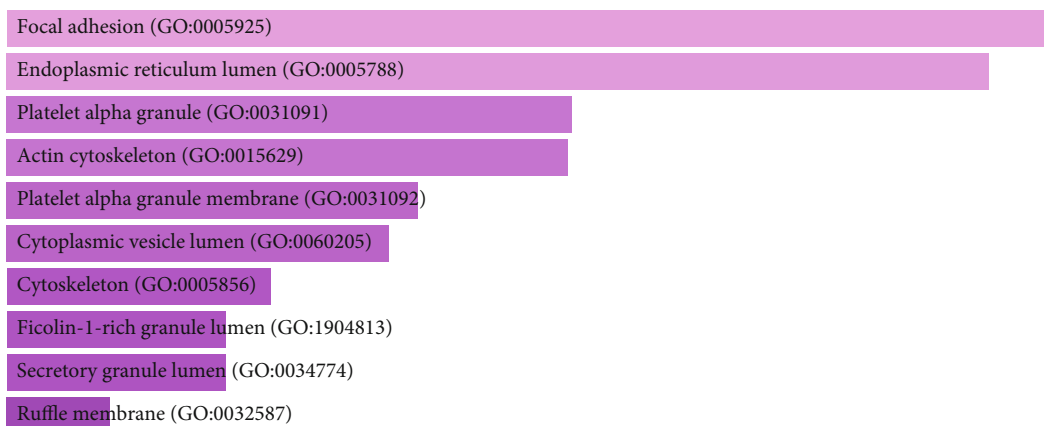
(a)

BP



(b)

CC



(c)

FIGURE 5: Continued.

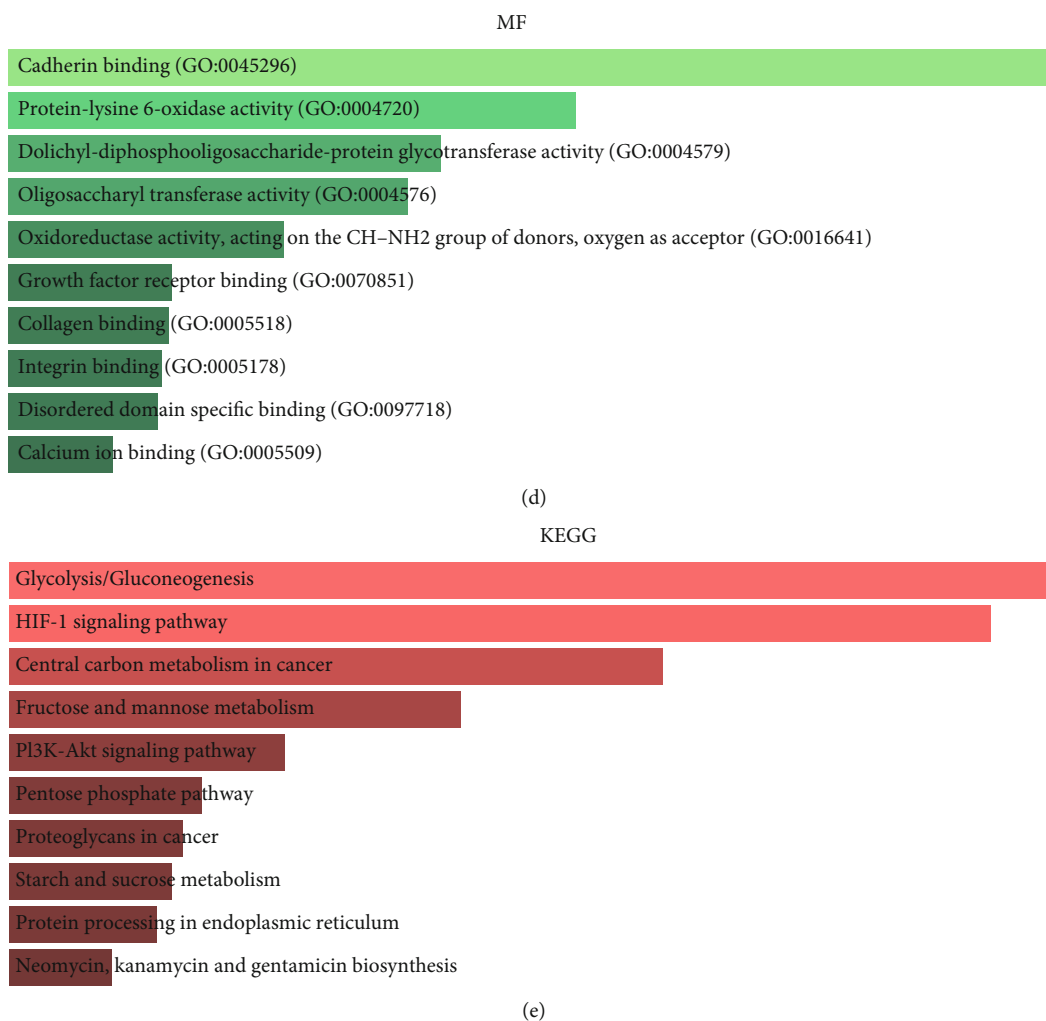


FIGURE 5: GO and KEGG pathway enrichment analyses of coexpressed genes. (a) A Venn diagram showing the intersection of coexpressed genes between UALCAN database and GEPIA database. A total of 139 coexpressed genes of P4HA1 were identified. (b–d) GO analysis revealed that the 139 coexpressed genes of P4HA1 were represented in biological process (BP), cellular component (CC), and molecular function (MF), respectively. (e) KEGG pathway enrichment analysis for the 139 coexpressed genes.

These results suggested that LINC01503 was the most promising lncRNA targeting to regulate miR-335-5p.

Finally, based on the coexpressed genes, the miRNA and lncRNA selected above, a LINC01503/miR-335-5p/P4HA1 ceRNA regulatory axis was constructed (Figure 9(g)). Additionally, a brief model diagram involving the potential mechanism of P4HA1 underlying the progression of pancreatic cancer was drawn. (Figure 9(h)).

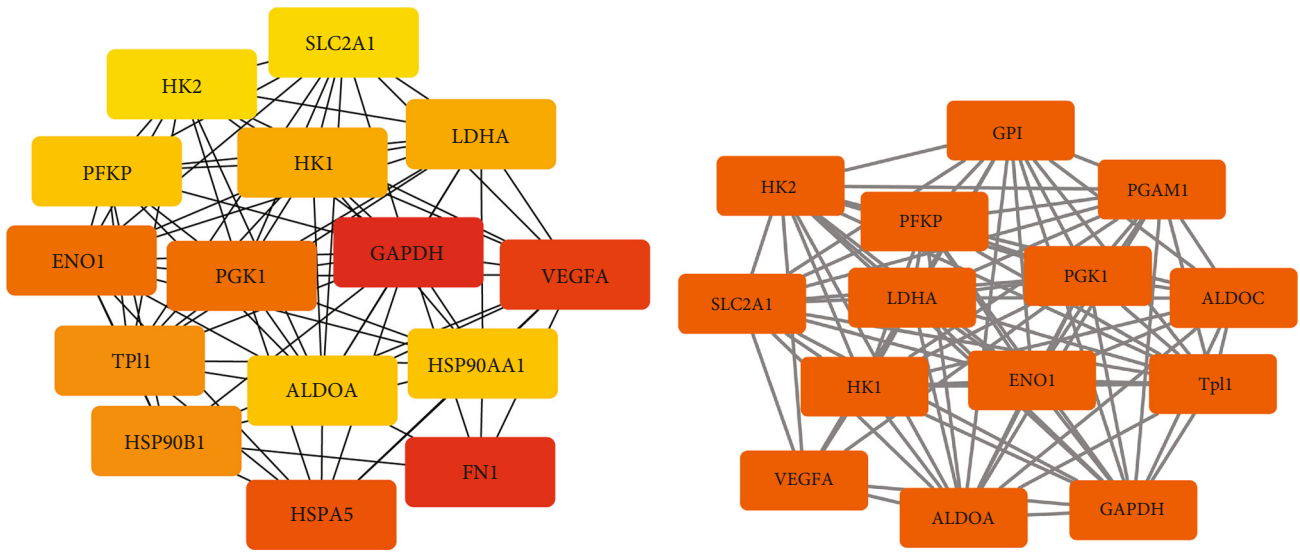
#### 4. Discussion

P4HA1 is a key intracellular enzyme that catalyzes the formation of 4-hydroxyproline, and it is necessary for the correct three-dimensional folding of newly synthesized procollagen chains, thereby maintaining the steady state of the extracellular matrix<sup>3, 29</sup>. P4HA1 is involved in regulating extracellular collagen synthesis and secretion and regulating extracellular matrix components. However, increasing evidence shows that P4HA1 is related to the occurrence, invasion, and metastasis of some human cancers including breast cancer

and prostate cancer, glioma, and hepatocellular carcinoma. Gilkes et al.<sup>4</sup> found that knocking down P4HA1 could reduce collagen deposition and inhibit the growth and metastasis of breast cancer cells. Additionally, the overexpression of P4HA1 was shown to promote the migration and metastasis of prostate cancer through the P4HA1-MMP1 signaling pathway<sup>5</sup>. Zhou et al.<sup>7</sup> found that the expression of P4HA1 may be regulated by HIF1 $\alpha$ , and its overexpression could promote invasive growth of gliomas by upregulating EMT induced by SNAI1 and SNAI2 and by causing angiogenesis of HBMECs. It was also shown that P4HA1 could be targeted by miR-30e, thereby inhibiting cell proliferation and colony formation in hepatocellular carcinoma<sup>30</sup>. Therefore, previous studies showed that P4HA1 played an important role in the process of malignant tumor progression. However, it was unclear whether P4HA1 plays an important role in the occurrence and malignant progression of pancreatic cancer, which was worthy of further exploration.

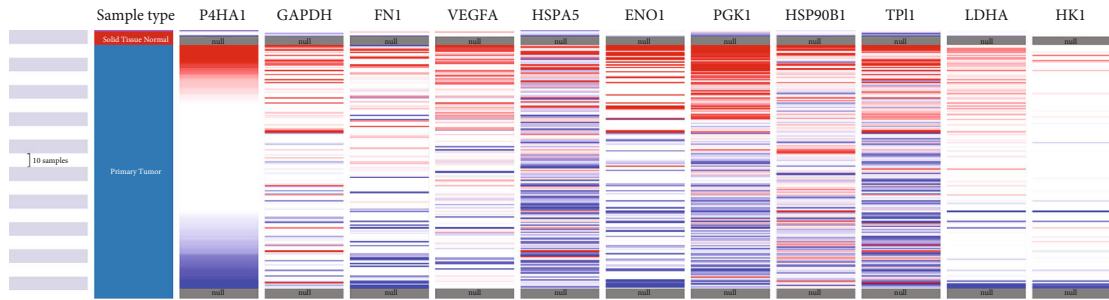
In this study, a brief pan-cancer analysis of P4HA1 in various tumors was first conducted. This facilitated an



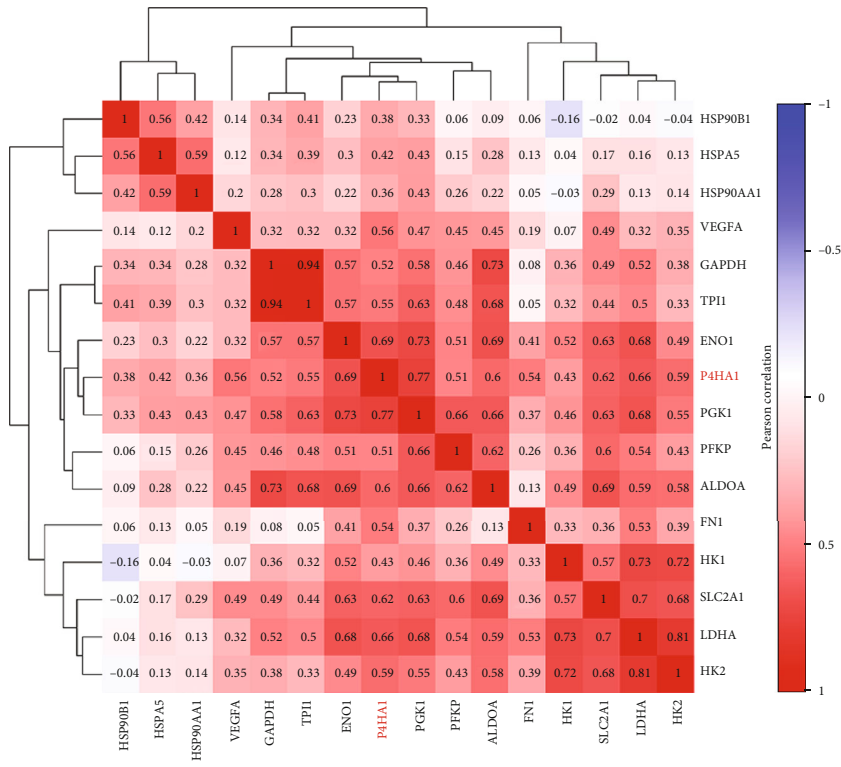


(a)

(b)



(c)



(d)

FIGURE 6: Continued.

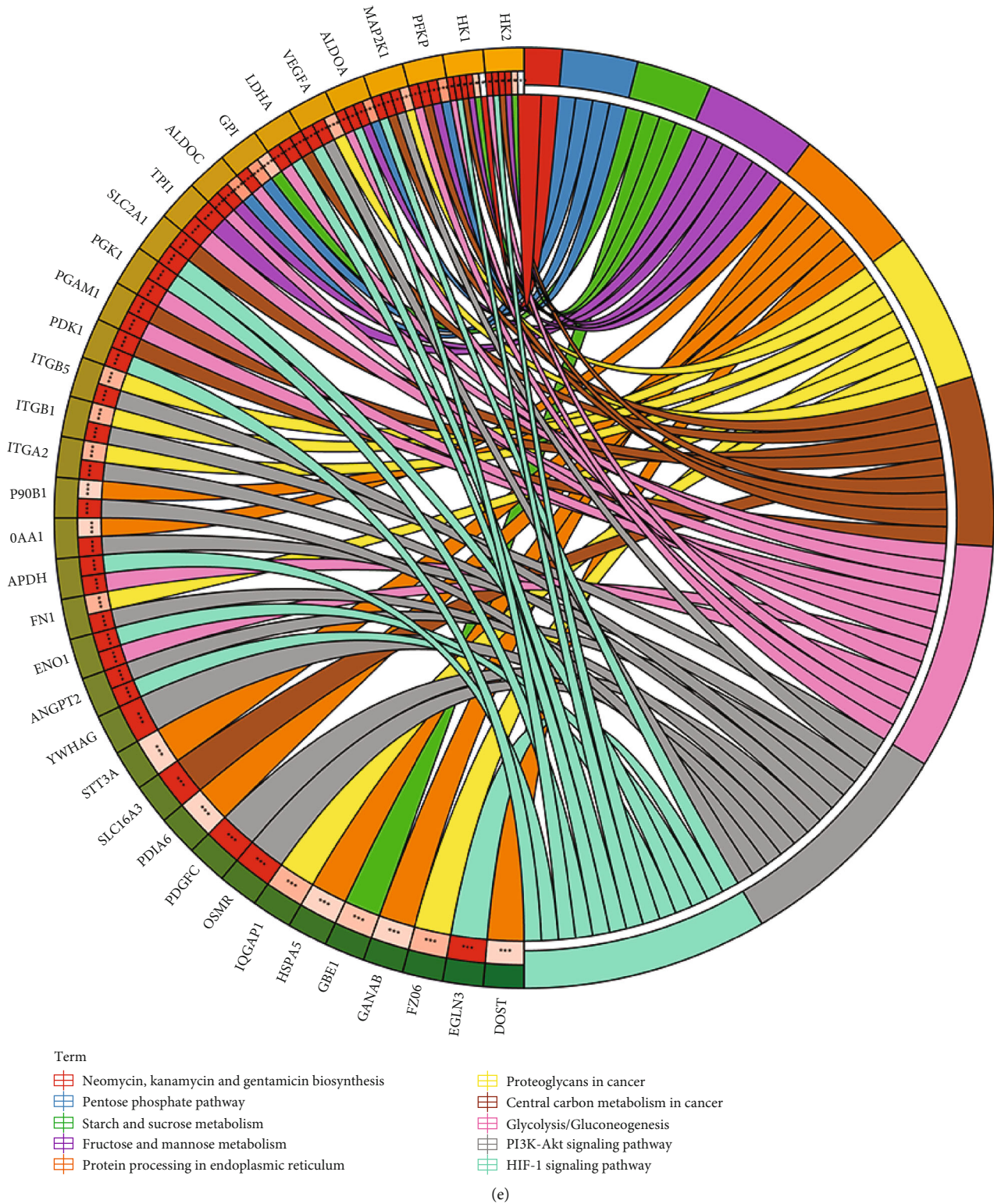
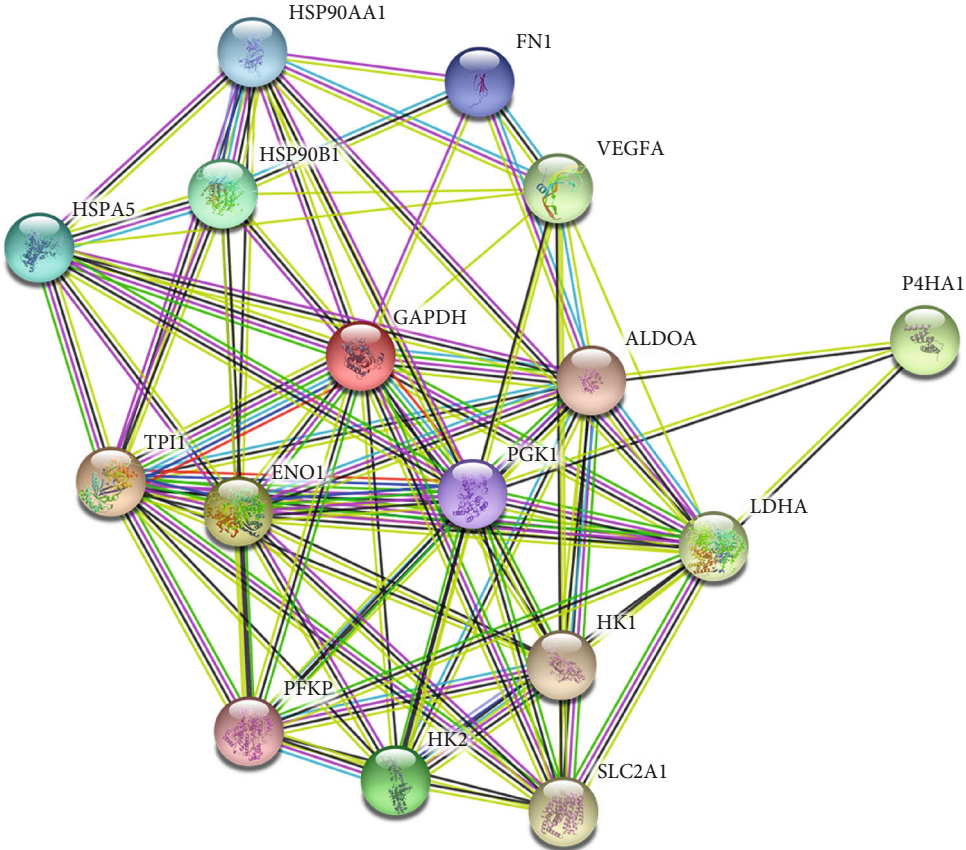
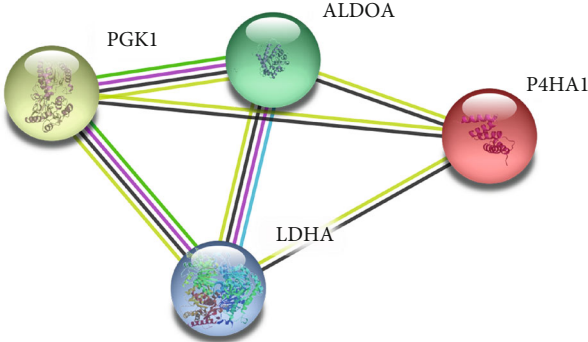


FIGURE 6: Construction of a protein-protein interaction (PPI) network of coexpressed genes of P4HA1 and analysis of Hub genes. (a) PPI network of coexpressed genes of P4HA1 was constructed and visualized using the STRING and Cytoscape tools, and the top 15 Hub genes were identified using CytoHubba tool kits in Cytoscape. (b) The most significant module in PPI networks was identified using MCODE tool kits in Cytoscape. (c) A heat map showing P4HA1 and top10 Hub genes expression in PAAD was constructed using UCSC. (d) Correlation clustering heat map of P4HA1 and top 15 Hub genes. (e) A circle diagram showing results of KEGG enrichment analysis of top 15 coexpressed genes.



(a)



(b)

FIGURE 7: Continued.

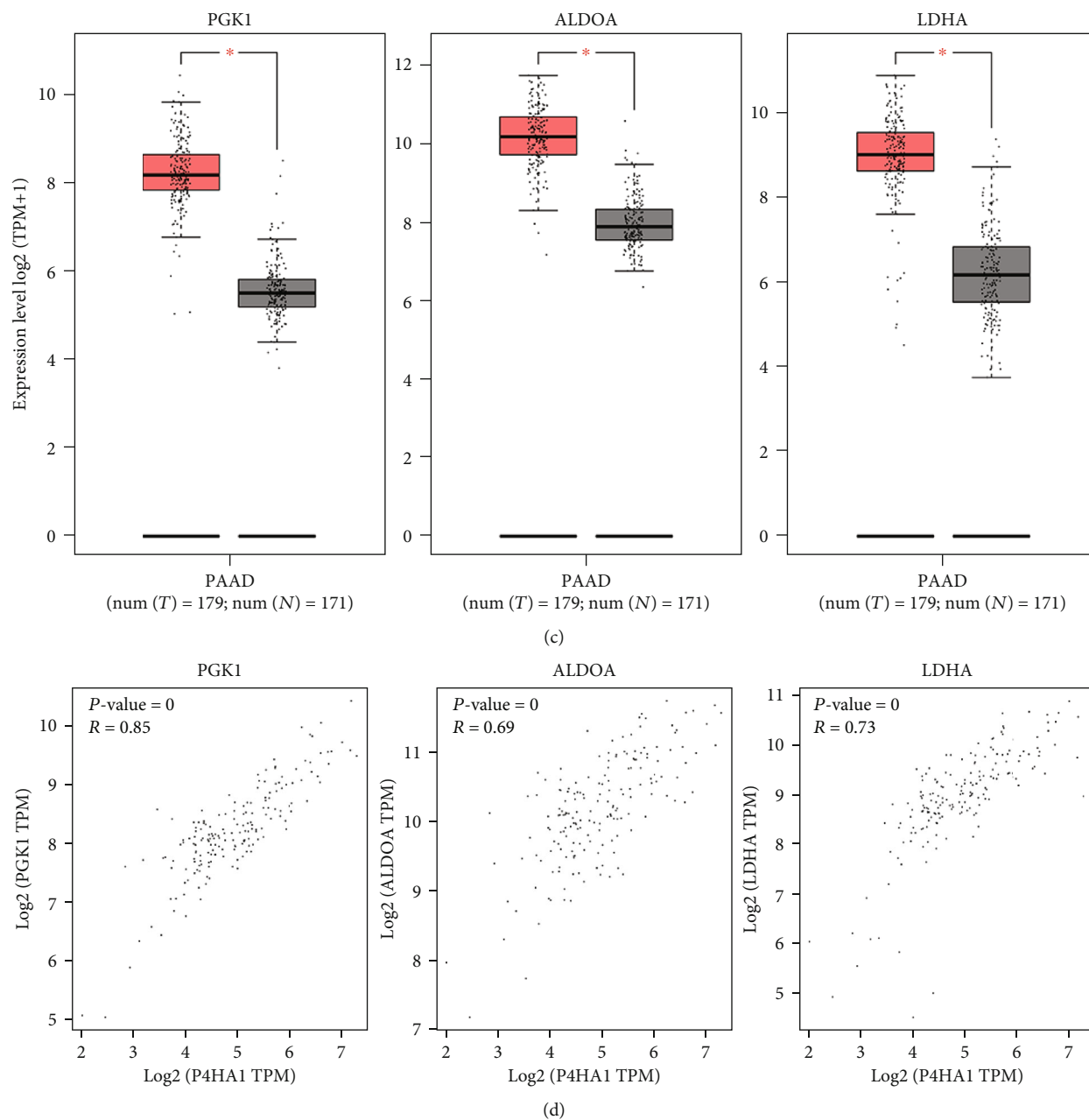


FIGURE 7: Continued.

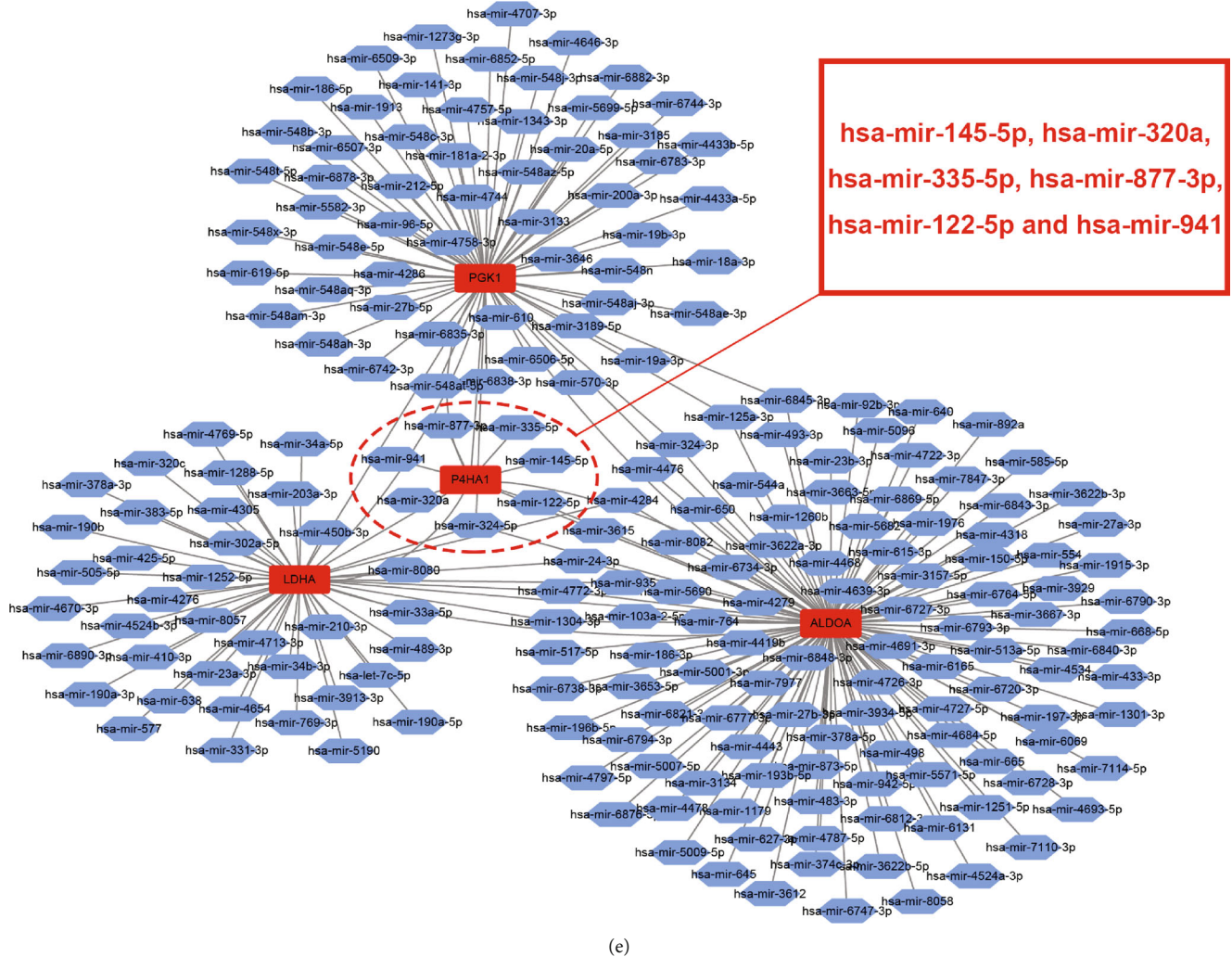
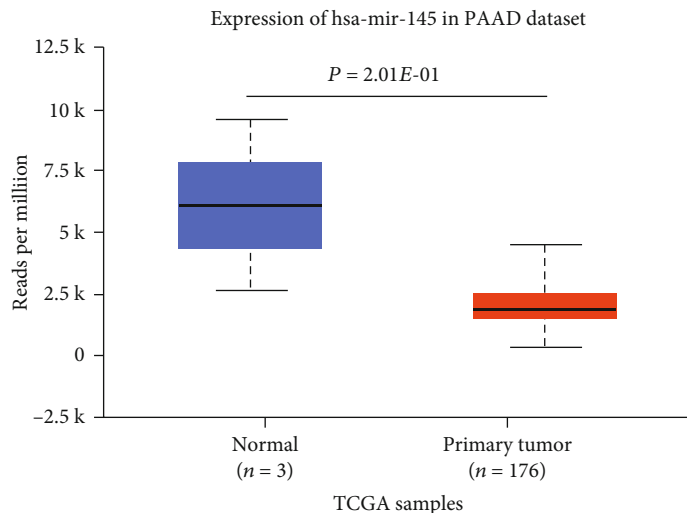


FIGURE 7: A PPI network showing Hub genes interacting with P4HA1 and the construction of a miRNA-mRNA network. (a) A PPI network showing the interaction among P4HA1 and its top 15 coexpressed genes; the interconnecting lines between those proteins indicate that there were interaction relationships between one and another. (b) A diagram of the PPI network for P4HA1 and three coexpressed genes directly interacting with P4HA1. (c) The expression of three coexpressed genes interacting with P4HA1 in PAAD using GEPIA. (d) Correlation diagram of the three coexpressed genes interacting with P4HA1 in PAAD using GEPIA. (e) The miRNA-mRNA network of the three coexpressed genes and P4HA1. Notably, hsa-miR-877-3p, hsa-miR-941, hsa-miR-320a, hsa-miR-324-5p, hsa-miR-122-5p, hsa-miR-145-5p, and hsa-miR-335-5p were predicted to target P4HA1.

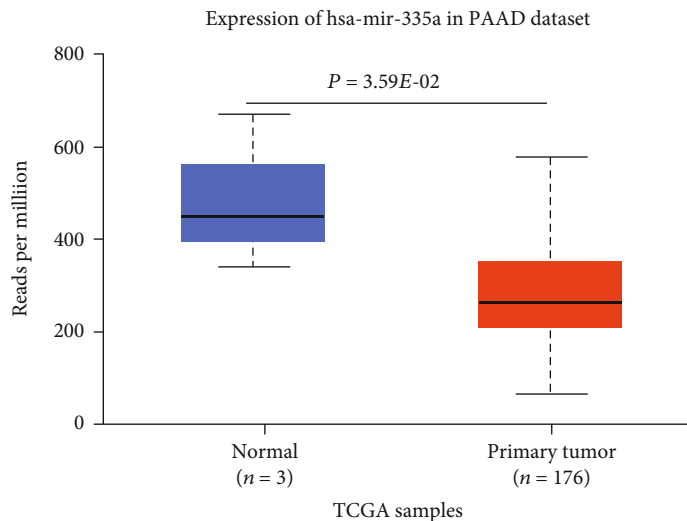
understanding of the expression pattern and prognosis of P4HA1 in various tumors. In pancreatic cancer, P4HA1 not only was highly expressed but also was related to prognosis in patients. This provided a basis for further analysis of the relationship between P4HA1 and pancreatic cancer. The study then sought to further verify the correlation between expression of P4HA1 and prognosis of pancreatic cancer. To achieve this, GEO data from the Oncomine database was used to observe the expression level of P4HA1 in pancreatic cancer tissues compared to normal ones. It was shown that expression in pancreatic cancer was significantly higher than that in normal tissues, consistent with the results obtained from analysis using the GEPIA database. In addition, the HPA database confirmed that the level of the P4HA1 protein in pancreatic cancer was higher than that in normal pancreatic tissues. Moreover, this study found that

high expression of P4HA1 was associated with poor clinicopathological features in pancreatic cancer. Our results showed that P4HA1 expression was not related to the M stage but considered the low number of M-positive patients included in the TCGA database; it may also have a certain deviation to our results. Therefore, this still needs to be confirmed by large-scale studies in the future. Subsequently, the prognostic analysis revealed that high expression of P4HA1 was related to poor prognosis in patients with pancreatic cancer. These results indicated that P4HA1 played an important and unfavorable role in the malignant transformation of pancreatic cancer and could be a promising biomarker of pancreatic cancer in the future.

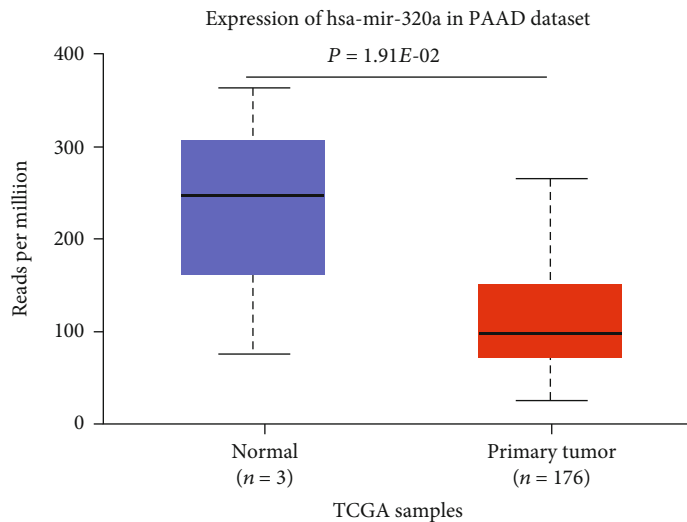
In order to better understand the potential function of P4HA1 in pancreatic cancer, genes coexpressed with the P4HA1 were predicted. GO enrichment analysis



(a)

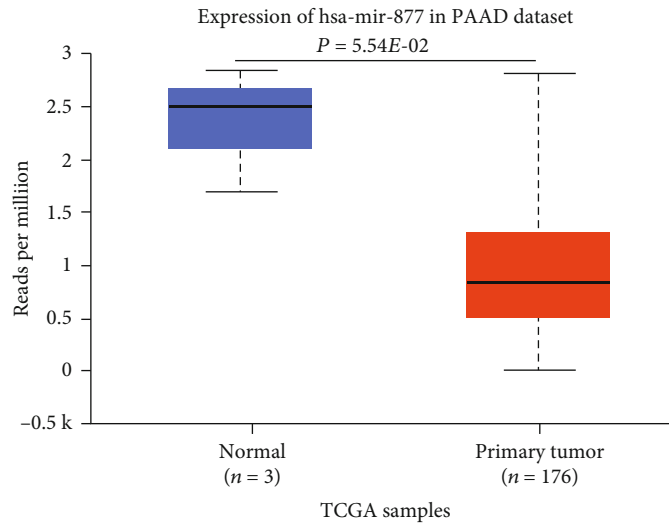


(b)

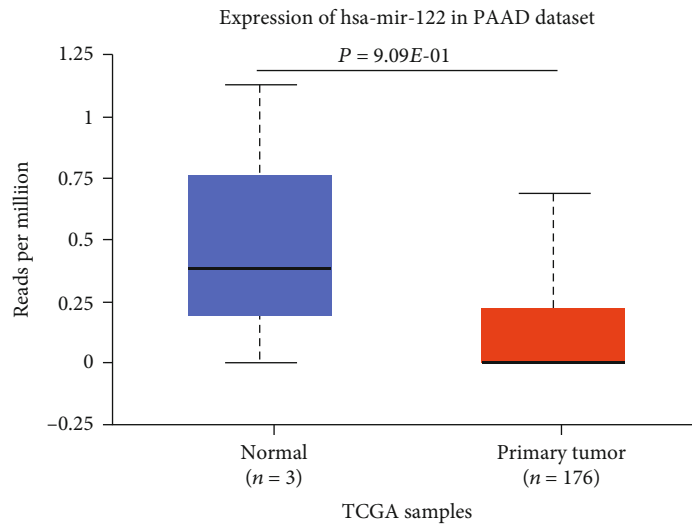


(c)

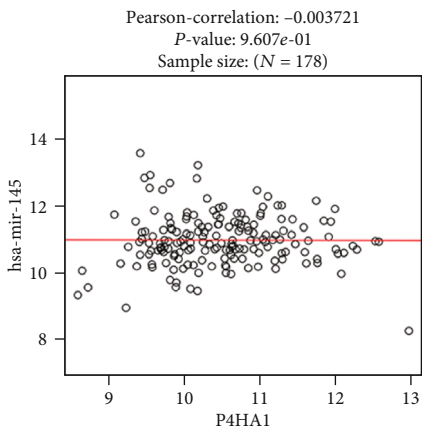
FIGURE 8: Continued.



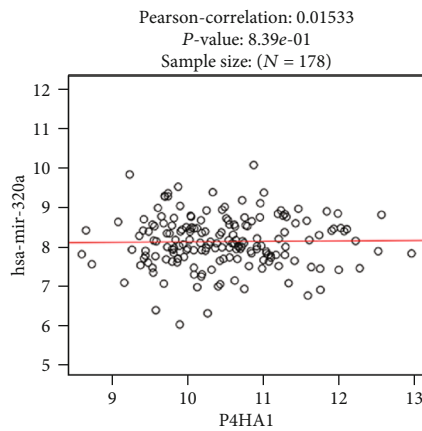
(d)



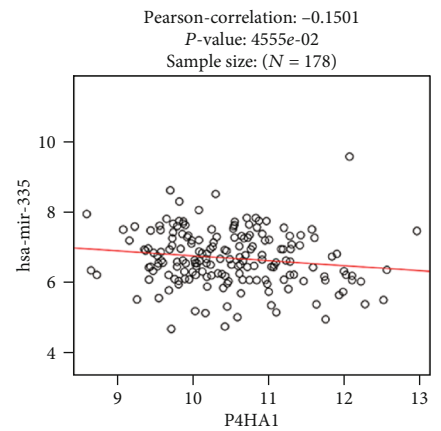
(e)



(f)



(g)



(h)

FIGURE 8: Continued.

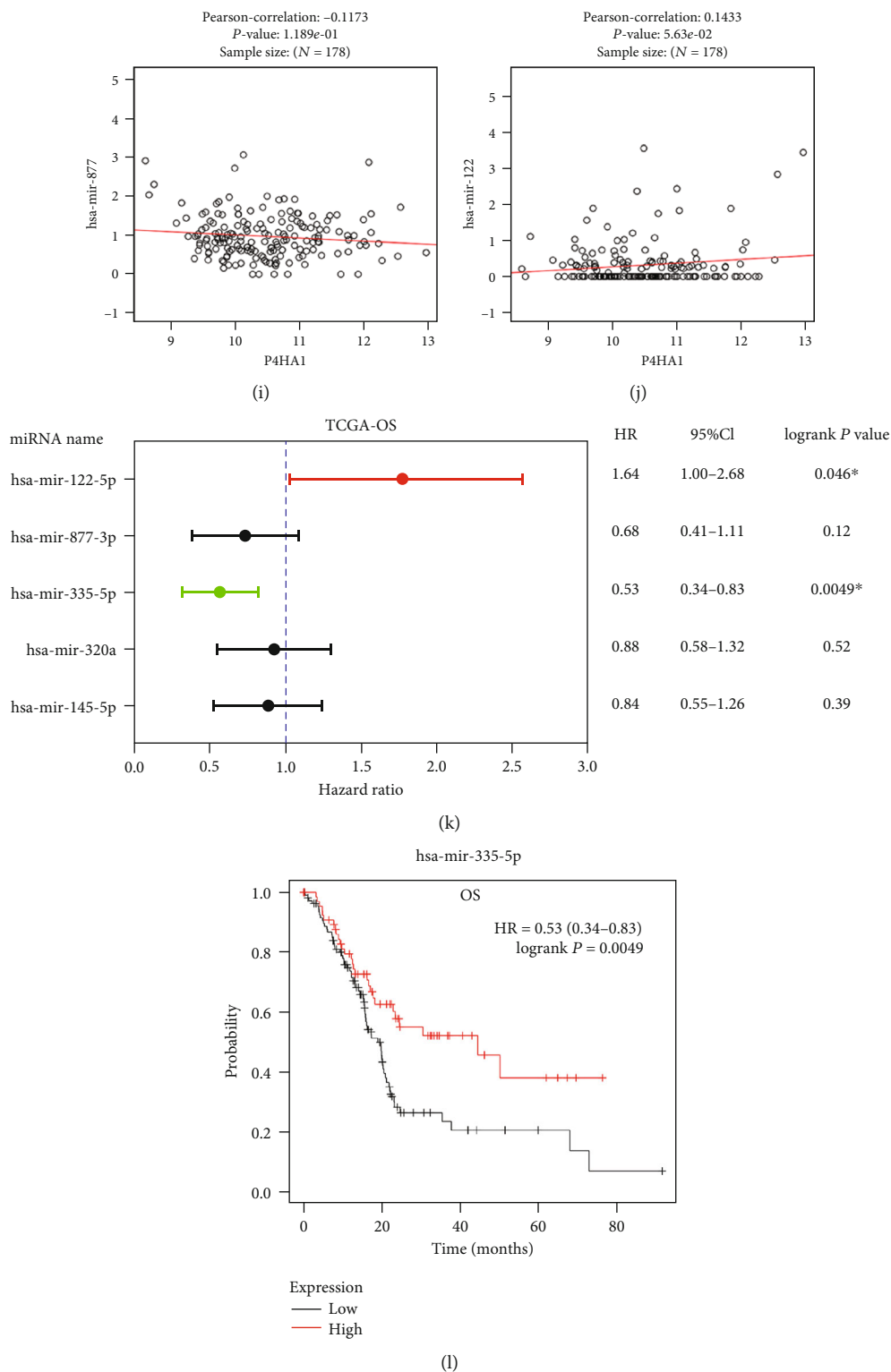


FIGURE 8: The expression, correlation, and prognostic value of 5 predicted upstream miRNAs of P4HA1. (a-e) The expression of 5 miRNAs in pancreatic cancer based on UALCAN. (f-j) Correlation between the expression of 5 miRNAs and P4HA1 expression in PAAD based on the UALCAN. (k) A forest map showing the prognostic value of 5 miRNAs in PAAD (TCGA). Green bars indicate a favorable prognosis; red bars indicate an unfavorable prognosis; black bars represent no significant statistical correlation. (l) The impact of miR-335-5p on overall survival of PAAD using Kaplan Meier-Plotter database.



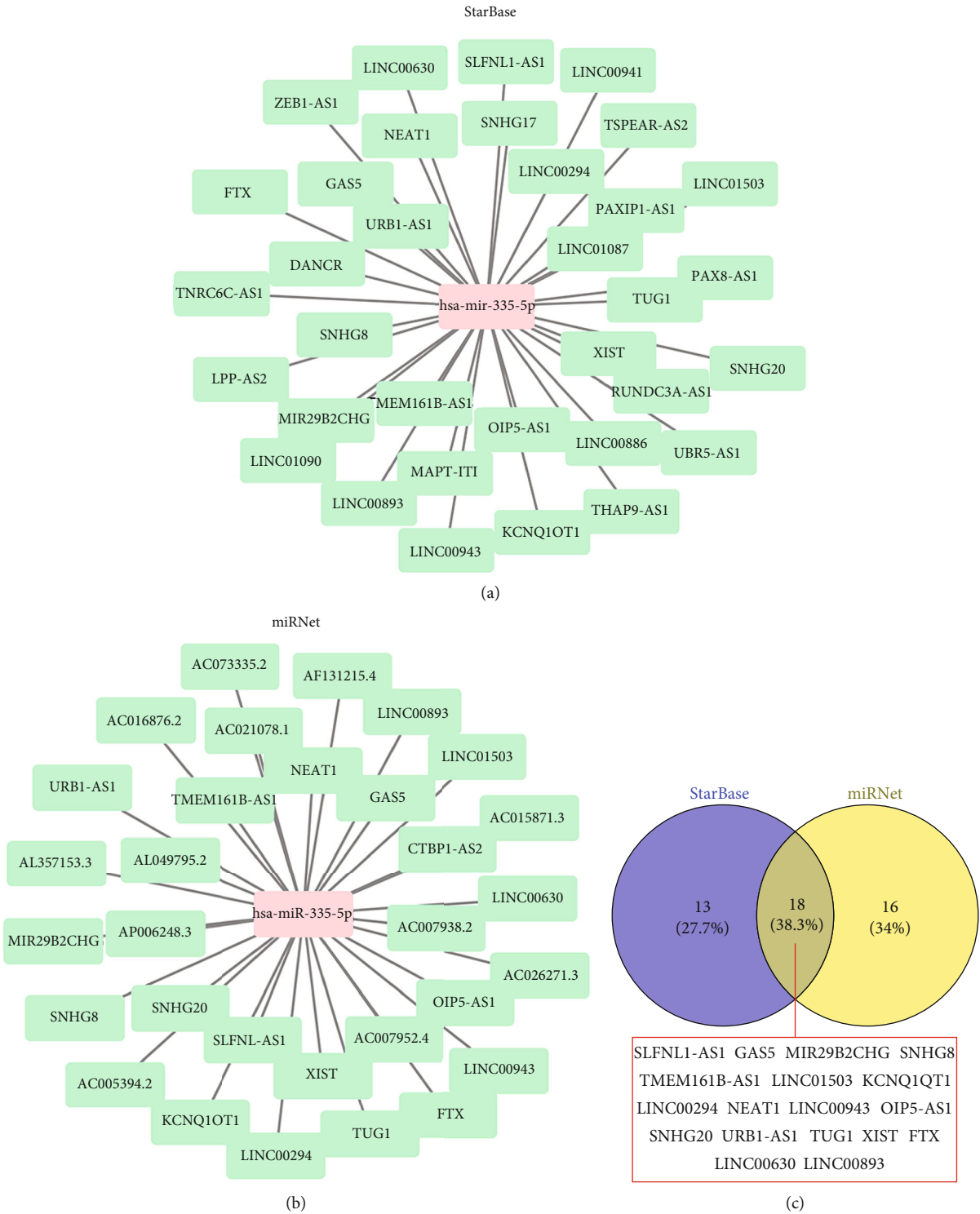
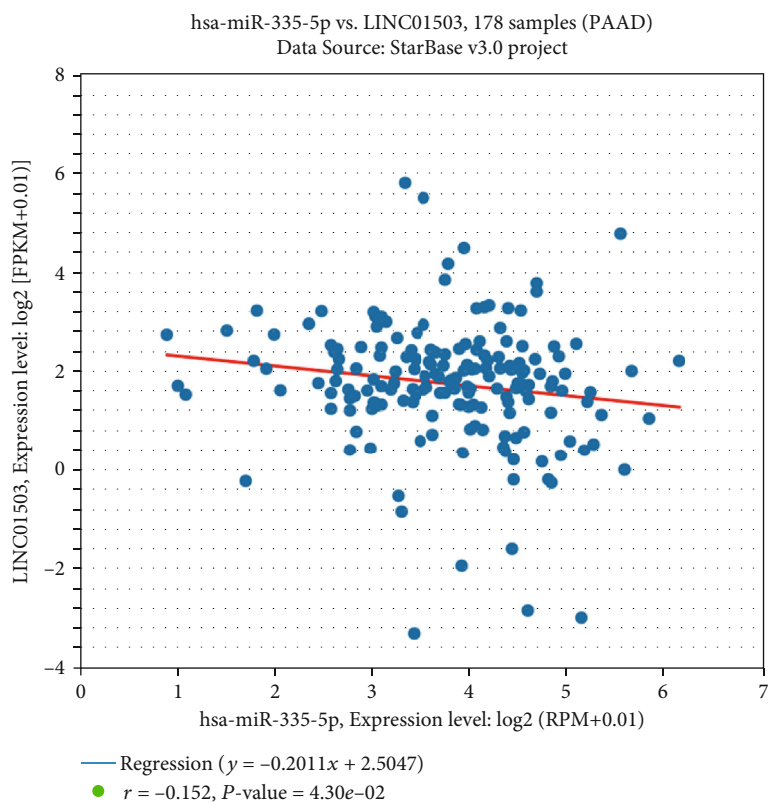
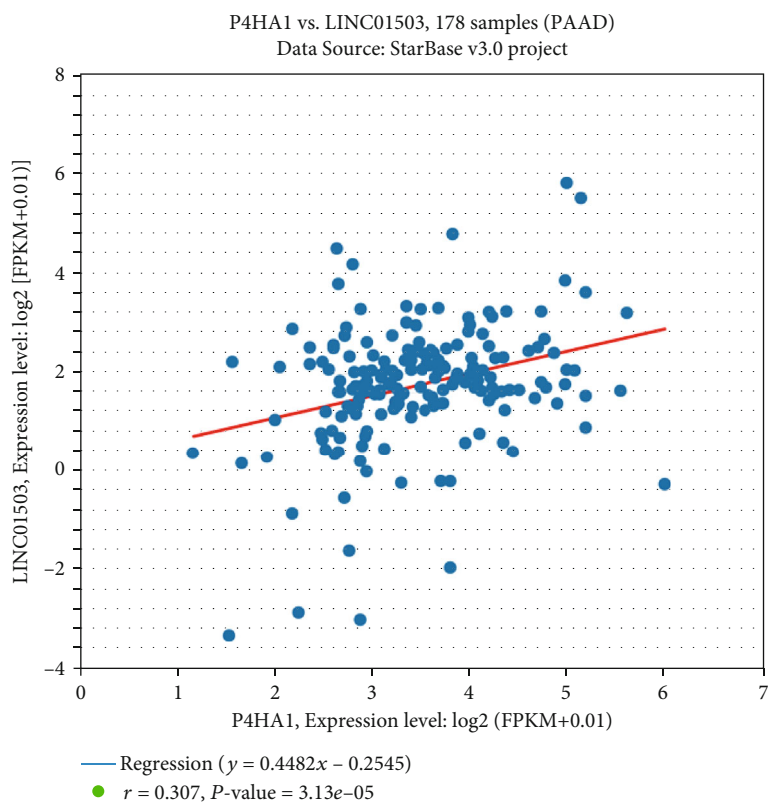


FIGURE 9: Continued.



(d)



(e)

FIGURE 9: Continued.

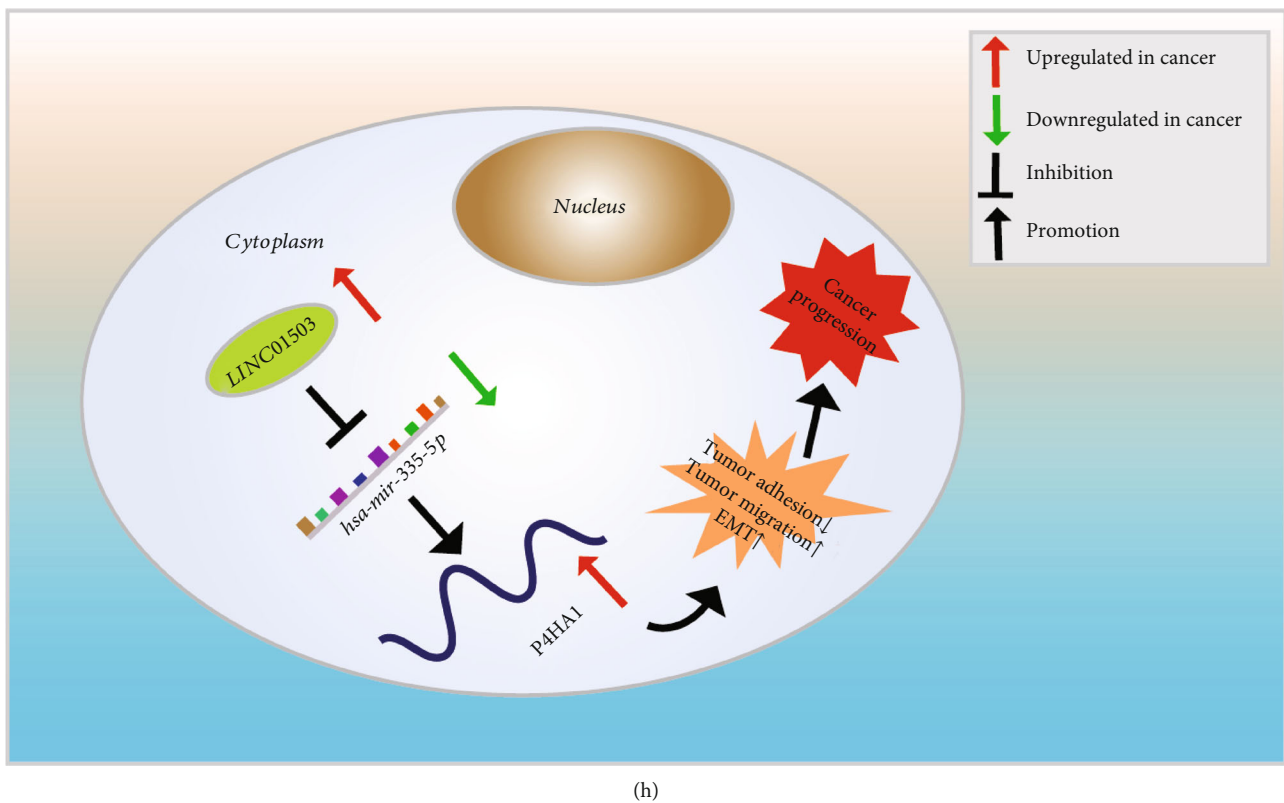
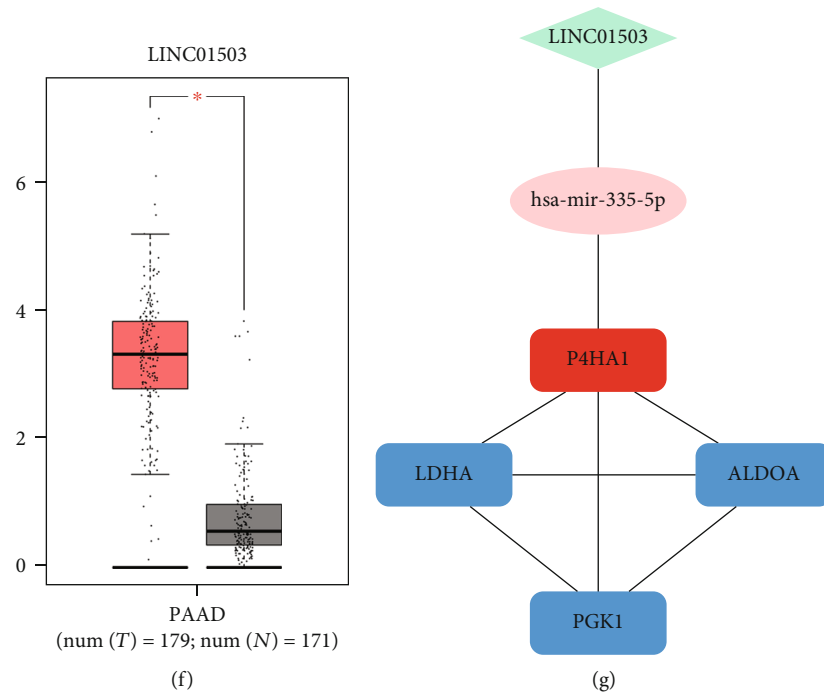


FIGURE 9: Identification of upstream lncRNAs of miR-335-5p, a ceRNA network, and a diagram showing a brief mechanism involving P4HA1 promote pancreatic cancer progress. (a) The predicted lncRNAs targeting miR-335-5p based on the StarBase database. (b) The potential lncRNAs targeting miR-335-5p as determined on the miRNet database. (c) 18 intersected lncRNAs obtained from the StarBase and miRNet databases. (d, e) Analysis of the correlation between LINC01503 and miR-335-5p and the correlation between LINC01503 and P4HA1. (f) LINC01503 expression in PAAD based on the analysis of GEPIA. (g) A ceRNA network was constructed. (h) A brief mechanism diagram was constructed.

TABLE 1: Correlation analysis between 18 lncRNAs and miR-335-5p and between 18 lncRNAs and P4HA1.

	miR-335-5p		P4HA1	
	Cor	<i>P</i> value	Cor	<i>P</i> value
SLFN1-AS1	0.111	1.41E – 01	-0.140	6.28E – 02
GAS5	0.223	<b>2.73E – 03</b>	0.084	2.63E – 01
MIR29B2CHG	0.257	<b>5.35E – 04</b>	-0.135	7.14E – 02
SNHG8	0.451	<b>2.56E – 10</b>	-0.266	<b>3.31E – 04</b>
TMEM161B-AS1	0.277	<b>1.85E – 04</b>	-0.203	<b>6.47E – 03</b>
LINC01503	-0.152	<b>4.30E – 02</b>	0.307	<b>3.13E – 05</b>
KCNQ1OT1	-0.044	5.60E – 01	0.057	4.50E – 01
LINC00294	-0.146	5.23E – 02	-0.137	6.74E – 02
NEAT1	-0.033	6.65E – 01	0.190	<b>1.09E – 02</b>
LINC00943	-0.128	8.90E – 02	0.029	6.99E – 01
OIP5-AS1	-0.217	<b>3.57E – 03</b>	0.253	<b>6.41E – 04</b>
SNHG20	0.261	<b>4.30E – 04</b>	-0.294	<b>6.73E – 05</b>
URB1-AS1	0.200	<b>7.33E – 03</b>	-0.102	1.75E – 01
TUG1	-0.097	1.98E – 01	0.278	<b>1.69E – 04</b>
XIST	-0.031	6.84E – 01	0.057	4.51E – 01
FTX	0.056	4.56E – 01	0.127	9.24E – 02
LINC00630	-0.098	1.95E – 01	0.427	<b>2.68E – 09</b>
LINC00893	0.178	<b>1.72E – 02</b>	-0.066	3.84E – 01

Cor: *R* value of Pearson's correlation; bold values indicate  $P < 0.05$ .

(Figures 6(b)–6(d)) and KEGG pathway enrichment analysis (Figure 6(e)) of these genes showed that they were mainly involved in the biological processes of glycolysis, glucose metabolism, and production of ATP from ADP. The genes were mainly associated with glycolysis/gluconeogenesis, the HIF-1 signaling pathway, central carbon metabolism of cancer, fructose and mannose metabolism, the PI3K-Akt signaling pathway, and the pentose phosphate pathway. These results were consistent with previous reports on P4HA1 and collagen synthesis as well as tissue cell hypoxia<sup>29, 31, 32</sup>. Therefore, we speculated that P4HA1 plays an important role in tumors probably through these biological processes. Moreover, three genes (PGK1, ALDOA, and LDHA) directly interacting with P4HA1 were selected from the top 15 Hub genes. Analysis using the GEPIA database further confirmed that the genes were overexpressed in pancreatic cancer. Additionally, there was a significant positive correlation between the three genes and P4HA1. Therefore, the results suggested that these coexpressed genes might play an important synergistic role with P4HA1 in the progression of pancreatic cancer through glycolysis, HIF-1 $\alpha$ , and metabolism.

Previous studies showed that the interaction of lncRNAs and miRNAs played an important role in the process of cancer and miRNAs could mediate tumor progression by regulating their downstream mRNA<sup>12</sup>. Therefore, in order to explore the potential regulatory mechanism of P4HA1 in the progression of pancreatic cancer, a lncRNA-miRNA-mRNA interaction network was constructed. First, six potential miRNAs were identified as candidates that targeted P4HA1 (miR-145-5p, miR-320a, miR-335-5p, miR-877-3p,

miR-122-5p, and miR-941), and predicted targeting relationship was verified in TarBase and miRTarBase. Further, expression levels of the six potential miRNAs in pancreatic cancer were analyzed. The results showed that two of them (miR-320a and miR-335-5p) were significantly downregulated in pancreatic cancer. According to the ceRNA hypothesis, miRNA has a negative regulatory relationship with the target gene, so miR-335-5p was further screened with negative correlation between miRNA and P4HA1 by correlation analysis.

The effect of these miRNAs on the OS of patients with pancreatic cancer was then analyzed using the Kaplan-Meier Plotter database. The results further proved that miR-335-5p was a protective factor in pancreatic cancer [HR = 0.53, 95% CI (0.34-0.83),  $P = 0.0049$ ]. Moreover, previous studies reported that miR-335-5p had tumor suppressive effects in lung cancer<sup>33</sup>, gastric cancer<sup>34</sup>, colorectal cancer<sup>35</sup>, and breast cancer<sup>36</sup>. Therefore, miR-335-5p was considered to be the most promising miRNA that could target and regulate P4HA1.

Similarly, StarBase and miRNet were used to predict upstream lncRNA targeting miR-335-5p. 18 lncRNAs predicted by the two both databases were obtained. According to the ceRNA theory, lncRNAs are negatively correlated with miRNAs but positively correlated with mRNA; therefore, the only one lncRNA (LINC01503) was eligible. The expression level of LINC01503 in pancreatic cancer was then examined and found that it was significantly upregulated in PAAD. In addition, previous studies identified LINC01503 played an oncogene role in tumors such as colorectal cancer<sup>37</sup>, gastric

cancer<sup>38</sup>, glioma<sup>39</sup>, cholangiocarcinoma<sup>40</sup>, and cervical cancer<sup>41</sup> and promotes tumor proliferation and invasion. Therefore, LINC01503 was considered to be the most promising lncRNA that could target and regulate miR-335-5p. Finally, the LINC01503/miR-335-5p/P4HA1 ceRNA regulatory axis was constructed by combining the miRNA and lncRNA selected in the above analytical process.

In summary, the study utilized bioinformatics tools to provide a reliable, comprehensive analysis of the expression pattern, prognostic value, and potential regulatory mechanism of P4HA1 in pancreatic cancer. It was confirmed that the overexpression of P4HA1 not only was associated with poor prognosis but was also related to poor clinicopathology in the disease. In addition, the potential regulatory mechanism of P4HA1 in promoting the progression of pancreatic cancer was explored by constructing a ceRNA regulatory axis.

## 5. Conclusion

P4HA1 can be used as a novel and promising prognostic biomarker for patients with pancreatic cancer, and it may provide a new therapeutic direction for the treatment of pancreatic cancer. However, considering that the findings were based on bioinformatics analyses, more laboratory experiments and clinical trials are needed to further validate in the future.

## Abbreviations

P4HA1:	Prolyl 4-hydroxylase subunit alpha 1
ceRNA:	Competing endogenous RNA
lncRNA:	Long noncoding RNA
MRE:	miRNA recognition elements
ncRNA:	Nonprotein coding RNA
KICH:	Kidney Chromophobe
TCGA:	The Cancer Genome Atlas
GTEX:	Genotype-tissue expression
HPA:	Human Protein Atlas
GO:	Gene Ontology
KEGG:	Kyoto Encyclopedia of Genes and Genomes
MOCDE:	Molecular Complex Detection
PPI:	Protein-protein interaction
CC:	Cellular components
BP:	Biological processes
MF:	Molecular functions
FC:	Fold change
OS:	Overall survival
DSS:	Disease-specific survival
PFI:	Progress-free interval
DFI:	Disease-free interval
DFS:	Disease-free survival
HR:	Hazard ratio
CI:	Confidence intervals.

## Data Availability

Source data of this study were derived from the public repositories, as indicated in Materials and Methods of the manuscript. And the datasets used and/or analyzed during

the current study are available from the corresponding author on reasonable request.

## Conflicts of Interest

The authors declare that they have no competing interest.

## Authors' Contributions

Tianyou Liao and Yangzhi Hu are responsible for the study concept and design. Tianyou Liao and Zhili Hu developed the analytical methodologies. Zhili Hu, Yangzhi Hu, and Fang Song performed the analysis and interpreted the data. Zhili Hu, Fang Song, and Yangzhi Hu discussed the results and wrote the paper. Yangzhi Hu, Tianyou Liao, and Fang Song performed the statistical analysis. Tianyou Liao and Zhili Hu corrected and revised manuscript. Zhili Hu, Fang Song, and Yangzhi Hu contributed equally to this work.

## Acknowledgments

This study was financially supported by the Science and Technology Plan Project of Chenzhou Municipal Science and Technology Bureau (No. ZDYF2020179 and No. ZDYF201974). We thank Freescience Editorial Team for its linguistic assistance during the revision of this manuscript.

## References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] E. R. Witkowski, J. K. Smith, and J. F. Tseng, "Outcomes following resection of pancreatic cancer," *Journal of surgical oncology*, vol. 107, no. 1, pp. 97–103, 2013.
- [3] L. Chen, Y. H. Shen, X. Wang et al., "Human prolyl-4-hydroxylase  $\alpha$ (I) transcription is mediated by upstream stimulatory factors," *The Journal of biological chemistry*, vol. 281, no. 16, pp. 10849–10855, 2006.
- [4] D. M. Gilkes, P. Chaturvedi, S. Bajpai et al., "Collagen prolyl hydroxylases are essential for breast cancer metastasis," *Cancer research*, vol. 73, no. 11, pp. 3285–3296, 2013.
- [5] B. V. S. K. Chakravarthi, S. S. Pathi, M. T. Goswami et al., "The miR-124-prolyl hydroxylase P4HA1-MMP1 axis plays a critical role in prostate cancer progression," *Oncotarget*, vol. 5, no. 16, pp. 6654–6669, 2014.
- [6] W. M. Hu, J. Zhang, S. X. Sun et al., "Identification of P4HA1 as a prognostic biomarker for high-grade gliomas," *Pathology, research and practice*, vol. 213, no. 11, pp. 1365–1369, 2017.
- [7] Y. Zhou, G. Jin, R. Mi et al., "Knockdown of P4HA1 inhibits neovascularization via targeting glioma stem cell-endothelial cell transdifferentiation and disrupting vascular basement membrane," *Oncotarget*, vol. 8, no. 22, pp. 35877–35889, 2017.
- [8] F. S. Collins, M. Morgan, and A. Patrinos, "The Human Genome Project: lessons from large-scale biology," *Science*, vol. 300, no. 5617, pp. 286–290, 2003.
- [9] K. R. Chi, "The dark side of the human genome," *Nature*, vol. 538, no. 7624, pp. 275–277, 2016.

- [10] C. A. Davis, B. C. Hitz, C. A. Sloan et al., "The Encyclopedia of DNA elements (ENCODE): data portal update," *Nucleic acids research*, vol. 46, no. D1, pp. D794–d801, 2018.
- [11] L. Salmena, L. Poliseno, Y. Tay, L. Kats, and P. P. Pandolfi, "A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?," *Cell*, vol. 146, no. 3, pp. 353–358, 2011.
- [12] C. M. Klinge, "Non-coding RNAs: long non-coding RNAs and microRNAs in endocrine-related cancers," *Endocrine-related cancer*, vol. 25, no. 4, pp. R259–r282, 2018.
- [13] D. R. Rhodes, J. Yu, K. Shanker et al., "ONCOMINE: a cancer microarray database and integrated data-mining platform," *Neoplasia*, vol. 6, no. 1, pp. 1–6, 2004.
- [14] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic acids research*, vol. 45, no. W1, pp. W98–w102, 2017.
- [15] F. Pontén, J. M. Schwenk, A. Asplund, and P.-H. D. Edqvist, "The Human Protein Atlas as a proteomic resource for biomarker discovery," *Journal of Internal Medicine*, vol. 270, no. 5, pp. 428–446, 2011.
- [16] A. S. Zweig, D. Karolchik, R. M. Kuhn, D. Haussler, and W. J. Kent, "UCSC genome browser tutorial," *Genomics*, vol. 92, no. 2, pp. 75–84, 2008.
- [17] M. Goldman, B. Craft, T. Swatloski et al., "The UCSC Cancer Genomics Browser: update 2013," *Nucleic acids research*, vol. 41, no. D1, pp. D949–D954, 2013.
- [18] D. S. Chandrashekar, B. Bashel, S. A. H. Balasubramanya et al., "UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses," *Neoplasia*, vol. 19, no. 8, pp. 649–658, 2017.
- [19] E. Y. Chen, C. M. Tan, Y. Kou et al., "Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool," *BMC Bioinformatics*, vol. 14, no. 1, p. 128, 2013.
- [20] M. V. Kuleshov, M. R. Jones, A. D. Rouillard et al., "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic acids research*, vol. 44, no. W1, pp. W90–W97, 2016.
- [21] A. Franceschini, D. Szklarczyk, S. Frankild et al., "STRING v9.1: protein-protein interaction networks, with increased coverage and integration," *Nucleic acids research*, vol. 41, no. Database issue, pp. D808–D815, 2012.
- [22] M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, and T. Ideker, "Cytoscape 2.8: new features for data integration and network visualization," *Bioinformatics*, vol. 27, no. 3, pp. 431–432, 2011.
- [23] J. Xia, M. J. Benner, and R. E. W. Hancock, "NetworkAnalyst-integrative approaches for protein-protein interaction network analysis and visual exploration," *Nucleic Acids Research*, vol. 42, no. W1, pp. W167–W174, 2014.
- [24] G. Zhou, O. Soufan, J. Ewald, R. E. W. Hancock, N. Basu, and J. Xia, "NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis," *Nucleic acids research*, vol. 47, no. W1, pp. W234–w241, 2019.
- [25] S. V. Vasaikar, P. Straub, J. Wang, and B. Zhang, "LinkedOmics: analyzing multi-omics data within and across 32 cancer types," *Nucleic acids research*, vol. 46, no. D1, pp. D956–d963, 2018.
- [26] B. Györfy, P. Surowiak, J. Budczies, and A. Lánckzy, "Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer," *PLoS One*, vol. 8, no. 12, article e82241, 2013.
- [27] J. H. Yang, J. H. Li, P. Shao, H. Zhou, Y. Q. Chen, and L. H. Qu, "starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data," *Nucleic acids research*, vol. 39, suppl\_1, pp. D202–D209, 2011.
- [28] J.-H. Li, S. Liu, H. Zhou, L.-H. Qu, and J.-H. Yang, "starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data," *Nucleic acids research*, vol. 42, no. D1, pp. D92–D97, 2013.
- [29] J. Myllyharju, "Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis," *Matrix biology: journal of the International Society for Matrix Biology*, vol. 22, no. 1, pp. 15–24, 2003.
- [30] G. Feng, H. Shi, J. Li et al., "MiR-30e suppresses proliferation of hepatoma cells via targeting prolyl 4-hydroxylase subunit alpha-1 (P4HA1) mRNA," *Biochemical and biophysical research communications*, vol. 472, no. 3, pp. 516–522, 2016.
- [31] D. M. Gilkes, S. Bajpai, P. Chaturvedi, D. Wirtz, and G. L. Semenza, "Hypoxia-inducible factor 1 (HIF-1) promotes extracellular matrix remodeling under hypoxic conditions by inducing P4HA1, P4HA2, and PLOD2 expression in fibroblasts," *The Journal of biological chemistry*, vol. 288, no. 15, pp. 10819–10829, 2013.
- [32] R. Xu, "P4HA1 is a new regulator of the HIF-1 pathway in breast cancer," *Cell stress*, vol. 3, no. 1, pp. 27–28, 2019.
- [33] W. Du, H. Tang, Z. Lei et al., "miR-335-5p inhibits TGF- $\beta$ -1 induced epithelial-mesenchymal transition in non-small cell lung cancer via ROCK1," *Respiratory Research*, vol. 20, no. 1, p. 225, 2019.
- [34] L. L. Zhang, L. F. Zhang, X. H. Guo, D. Z. Zhang, F. Yang, and Y. Y. Fan, "Downregulation of miR-335-5p by long noncoding RNA ZEB1-AS1 in gastric cancer promotes tumor proliferation and invasion," *DNA and cell biology*, vol. 37, no. 1, pp. 46–52, 2018.
- [35] D. Zhang and N. Yang, "MiR-335-5p inhibits cell proliferation, migration and invasion in colorectal cancer through downregulating LDHB," *Journal of BUON: official journal of the Balkan Union of Oncology*, vol. 24, no. 3, pp. 1128–1136, 2019.
- [36] Q. Jia, L. Ye, S. Xu et al., "Circular RNA 0007255 regulates the progression of breast cancer through miR-335-5p/SIX2 axis," *Thoracic cancer*, vol. 11, no. 3, pp. 619–630, 2020.
- [37] S. R. Lu, Q. Li, J. L. Lu, C. Liu, X. Xu, and J. Z. Li, "Long non-coding RNA LINC01503 promotes colorectal cancer cell proliferation and invasion by regulating miR-4492/FOXK1 signaling," *Experimental and therapeutic medicine*, vol. 16, no. 6, pp. 4879–4885, 2018.
- [38] J. Ding, F. Shi, G. Xie, and Y. Zhu, "Long non-coding RNA LINC01503 promotes gastric cancer cell proliferation and invasion by regulating Wnt signaling," *Digestive diseases and sciences*, 2020.
- [39] H. Wang, Z. G. Sheng, and L. Z. Dai, "Long non-coding RNA LINC01503 predicts worse prognosis in glioma and promotes tumorigenesis and progression through activation of Wnt/ $\beta$ -catenin signaling," *European review for medical and pharmacological sciences*, vol. 23, no. 4, pp. 1600–1609, 2019.
- [40] Y. K. Qu, X. S. Qu, G. Chen et al., "LINC01503 promotes cell proliferation, invasion and EMT process in cholangio-carcinoma," *European review for medical and pharmacological sciences*, vol. 23, no. 15, pp. 6445–6452, 2019.
- [41] X. Peng, J. Gao, C. Cai, and Y. Zhang, "LncRNA LINC01503 aggravates the progression of cervical cancer through sponging miR-342-3p to mediate FXRD3 expression," *Bioscience Reports*, vol. 40, no. 6, 2020.