

RESEARCH

Open Access



Comprehensive analysis of the transcriptional expressions and prognostic value of S100A family in pancreatic ductal adenocarcinoma

Hong-Bin Li^{1†}, Jian-Li Wang^{2†}, Xiao-Dong Jin³, Lei Zhao⁴, Hui-Li Ye⁴, Yan-bei Kuang³, Yong Ma⁵, Xiang-Yan Jiang⁵ and Ze-Yuan Yu^{5*}

Abstract

Background: Pancreatic ductal adenocarcinoma (PDAC) remains a treatment-refractory malignancy with poor prognosis. It is urgent to identify novel and valid biomarkers to predict the progress and prognosis of PDAC. The S100A family have been identified as being involved in cell proliferation, migration and differentiation progression of various cancer types. However, the expression patterns and prognostic values of S100As in PDAC remain to be analyzed.

Methods: We investigated the transcriptional expressions, methylation level and prognostic value of S100As in PDAC patients from the Oncomine, GEPIA2, Linkedomics and cBioPortal databases. Real-time PCR was used to detect the expressions of S100A2/4/6/10/14/16 in four pancreatic cancer cell lines and pancreatic cancer tissues from PDAC patients undergoing surgery. To verify the results further, immunohistochemistry was used to measure the expression of S100A2/4/6/10/14/16 in 43 PDAC patients' tissue samples. The drug relations of S100As were analyzed by using the Drugbank database.

Results: The results suggested that, the expression levels of S100A2/4/6/10/14/16 were elevated to PDAC tissues than in normal pancreatic tissues, and the promoter methylation levels of S100A S100A2/4/6/10/14/16 in PDAC ($n = 10$) were lower compared with normal tissue ($n = 184$) ($P < 0.05$). In addition, their expressions were negatively correlated with PDAC patient survival.

Conclusions: Taken together, these results suggest that S100A2/4/6/10/14/16 might be served as prognostic biomarkers for survivals of PDAC patients.

Keywords: S100A, Pancreatic ductal adenocarcinoma, Biomarkers, Prognosis, Bioinformatics

* Correspondence: yuzy17@lzu.edu.cn

[†]Hong-Bin Li and Jian-Li Wang contributed equally to this work.

⁵Department of General Surgery, Gansu Provincial Cancer Hospital Lanzhou University Second Hospital, No.82 Cui-ying Men, Lanzhou 730030, Gansu Province, China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) remains one of the most malignant tumors with poor prognosis and high mortality. In recent years, its morbidity and mortality have increased significantly, and the 5-year survival rate is < 1% [1]. Due to the lack of early screening indicators, early PDAC is difficult to detect and therefore, spreads rapidly to surrounding organs, making it one of the deadliest cancers [2]. At present, the treatment of PDAC is very limited, and almost only surgical resection can provide a realistic opportunity of survival [3]. Thus, there is an urgent need to identify novel and valid markers to predict the progression and prognosis, and to understand the mechanism of PDAC. Calcium ions are the important factor that controls the balance between signal transduction, cell survival, and metabolism, whose dysregulation is thought to be closely related to cancer development [4]. The S100A family is a class of small-molecule calcium-binding proteins, which are identified only in vertebrates [5]. Many of their members are abnormally expressed in various cancer types, which are widely involved in cell proliferation, migration and differentiation during cancer development [6]. For instance, five calcium-binding proteins (S100A2/7/8/9/10) have been identified as overexpressed in gastric cancer, which might be important for gastric tumorigenesis [7]. Gupta S et al. reported that, the abnormal over-expression of S100A4 observed in cancer tissues are associated with an increase in tumor grade in prostate adenocarcinoma [8]. Analogously, Wang et al. found that, the expression level of S100A11 in colorectal cancer tissue was increased following stage progression of tumor progress [9].

Aberrant expressions and the prognostic values of some members of S100As family have been found in previous studies. However, to date, roles of S100As in the occurrence and development of PDAC remains largely elusive. Here, for the first time, we have used bioinformatics analysis to address this problem by analyzing the expression profiles, methylation level and mutations of different S100As family members and some relations with clinical parameters in PDAC patients. We investigated the expression of S100As in four PDAC cell lines. In addition, PDAC tissues and their matched adjacent normal pancreatic tissues from three PDAC patients undergoing surgery were collected, the expression levels of some S100As were measured using real-time PCR, and further verification was performed by immunohistochemistry in 43 cases of pancreatic cancer samples. Furthermore, we also analyzed genes, which have the similar expression pattern as S100As in PDAC tissues, and predicted the functions and pathways of related genes. This method based on the analysis of different omics levels, and the verification of clinical samples

make full use of bioinformatics data, which provides a complementary means for finding potential biomarkers of PDAC.

Materials and methods

OncoPrint database analysis

OncoPrint is a web-based cancer microarray database and integrated data-mining platform. The gene expression data referenced during the study are available in OncoPrint website (www.oncoPrint.com). In our study, transcriptional expressions of 16 different S100As members between different types of cancer and the corresponding normal controls were acquired from OncoPrint database. The differences in transcriptional expression were compared through students' t test. The relevant parameters are as follows: data type: mRNA, *p* value: 1E-4, fold change: 2, gene rank: 10%.

GEPIA 2 analysis

GEPIA2 (Gene Expression Profiling Interactive Analysis) was used to analyze transcriptional expressions of 16 different S100As members between different types of cancer and the corresponding adjacent normal control. The relevant parameters are as follows: |Log2FC| Cutoff:1, *q*-value Cut off:0.01, match the TCGA normal and Genotype Tissue Expression (GTEx) data, statistical methods: ANOVA. GEPIA 2 were also used to analyze Overall Survival of PDAC patients. Cutoff-High (%):50, Cutoff-Low (%):50.

UALCAN database analysis

UALCAN is an online tool for in-depth analyses of TCGA gene expression data [10]. The transcript levels of S100As in PDAC samples were analyzed using the UALCAN program. We also performed survival analysis (OS) based on the expression status of S100As signature and plot a Kaplan-Meier curve by using UALCAN.

LinkedOmics analysis

LinkedOmics website (<http://www.linkedomics.org>), which includes multi-omics data from all 32 TCGA cancer types, was used to screen the positively correlated significant genes of S100As, and the parameters are described as follows. Cancer cohort: TCGA_PAAD; search dataset: RNAseq; target dataset: RNAseq; statistical method: Pearson correlation test. Based on the score of person correlation, select the top 25 genes for further research.

cBioPortal

We retrieved the genomic profiles of S100As members, which contained mutations and clinical data from cBioportal (<http://www.cbioportal.org/>). The relevant parameters are as follows: Selected Studies: Pancreatic

Adenocarcinoma (TCGA, PanCancer Atlas)(184 total samples); Select Genomic Profiles: mutation, mRNA expression z-score relative to diploid (RNA Seq V2 RSEM), the z-score threshold ± 1.8 . Select Patient/Case Set: samples with mRNA data (RNA Seq V2)(177). Gene list: S100A2, S100A4, S100A6, S100A10, S100A14, S100A16.

Methylation analysis

MEXPRESS (<https://mexpress.be/index.html>) is a data visualization tool designed to easily visualize TCGA expression, DNA methylation and clinical data, and the relationship between them [11]. We analyzed the methylation levels of S100A2, S100A4, S100A6, S100A10, S100A14, S100A16 using the MEXPRESS database.

Drugbank analysis

The DrugBank database (<https://www.drugbank.ca/>) is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information [12]. The DrugBank (version 5.1.6) database was used to analyze drug information targeting S100As.

Protein–protein interaction network construction

Linkedomics database was used to search for significant genes positively correlated to S100As in pancreatic cancer samples. The significant genes were used for further PPI construction based on STRING online database (<http://string-db.org>). Furthermore, GO analysis and KEGG pathway enrichment of the screened S100As were carried out based on the Metascape [13].

Cell culture

Human pancreatic cell line hTERT-HPNE, human pancreatic cancer lines PANC-1, CFAPC-1, MIA PaCa-2 and ASPC-1 were obtained from the cell bank of Chinese Academy of Sciences (Shanghai, China). All of the cells were maintained in high glucose DMEM media supplemented with 10% fetal bovine serum, and were incubated at 37 °C in a humidified atmosphere containing 5% CO₂.

RNA preparation and real-time PCR

Total RNA was isolated by Trizol reagent (Invitrogen, USA), and quantified by 260/280 ratios determined using Nanodrop. cDNA was synthesized by PrimeScript RT Master Mix (Takara, China) from 500 ng of RNA. Real-time PCR was performed using TB Green Premix (Takara, China). Real-time PCR amplifications were performed on the QuantStudio 5 System. The expression levels of S100As were normalized to GAPDH in each individual sample. The relevant primers were listed in supplementary Table S1.

Immunohistochemistry

Immunohistochemistry (IHC) was performed using standard procedures described as previous study [14]. The intensity was scored as “0,” “1,” “2,” and “3”; that was negative, weak, moderate, and strong staining, respectively. The positively stained cells percentage was scored as “0 to 100” The final IHC score was obtained by product between intensity score and stained area. The primary antibodies were used for this study: goat anti-rat S100A2/4/6/10 dilution 1:200 (Abcam); goat anti-rat S100A14 and S100A16 dilution 1:200 (Proteintech). And a score of less than 150 is considered low expression, and a score of 150 or more is considered high expression.

Statistical analysis

SPSS version 26.0 (IBM, Armonk, NY, USA) was used to perform statistical analyses. Comparison between groups was made using the Student’s t-test, and the mean \pm standard derivation for normally distributed continuous data, as the median (interquartile range, Q25 - Q75) for abnormally distributed continuous data. Moreover, categorical data was analyzed using the Pearson chi-square test or Fisher’s exact test and presented as percentages or proportions. Kaplan-Meier analyses was used to analyze overall survival rate between different groups with a log-rank test was used to calculate the overall cumulative probability, and the difference was presented as hazard ratio (HR) value and *P* value. *P*-values < 0.05 were considered significant.

Results

Expression patterns of S100As in pancreatic cancer

Abnormal expression of genes in tumors is often the driving factor for tumor progression. In order to explore abnormal gene expression in PDAC patients, Oncomine database was used to analyze expression levels of S100As in cancer and normal samples. We found that multiple members of the S100A family (S100A2, S100A4, S100A6, S100A10, S100A11, S100A13, S100A14, and S100A16) are highly expressed in PDAC (Fig. 1).

Next, GEPIA 2 was used to obtain the expression matrix plots based on S100As family members. The color density indicates the median expression of one gene in specified tissue, and is further normalized by the maximum median expression value in all blocks. As shown in Figs. 2 and 3, significantly higher mRNA expressions of S100A2/3/4/6/8/9/10/11/13/14/16 were found in PDAC versus normal pancreatic tissue. Moreover, the transcription expression pattern of S100As family members were further analyzed by UALCAN whose resources were based on publicly available cancer OMICS data (TCGA and MET500), which was different

Analysis Type by Cancer	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal	Cancer vs. Normal																	
	S100A1	S100A2	S100A3	S100A4	S100A5	S100A6	S100A7	S100A8	S100A9	S100A10	S100A11	S100A12	S100A13	S100A14	S100A15	S100A16																	
Bladder Cancer		2				1			1		3				2																		
Brain and CNS Cancer	7			2		2		1		7	5			2		1																	
Breast Cancer	2						3	2		2						1																	
Cervical Cancer										1			1																				
Colorectal Cancer		16	3					1	4	4		2	15		7	1																	
Esophageal Cancer			2	1						1	2	4		3		2																	
Gastric Cancer		1		1							3		1																				
Head and Neck Cancer	1	3	2	1		1	4			1	2	2	1	1	3	3																	
Kidney Cancer		1				2				3	3	9	2	2	1	10																	
Leukemia Cancer						1	4	2		3	6	3	2	1																			
Liver Cancer						3				3		3	2			1																	
Lung Cancer		1	7		8		1			8		3	6		2	1																	
Lymphoma																																	
Melanoma	1															1																	
Myeloma			1																														
Other Cancer																																	
Ovarian Cancer			1			1				1	1	1		2	1	1																	
Pancreatic Cancer	2		1			2					4		4		2																		
Prostate Cancer																																	
Sarcoma		5				1	1			8		2	1	2	1	1																	
Significant Unique Analyses	4	19	33	10	6	9	17	22		1	18	12	9	10	29	12	21	29	21	54	15	4	20	17	7	13	22	6	1	7	6		
Total Unique Analyses	443	429	436	403	394	337	414	438	420	446	408	438	437	342	247	265																	



Fig. 1 The transcription levels of S100As in various tumors (OncoPrint). Difference of transcriptional expression was compared by students' t-test. The color indicates the best gene rank percentile for the analyses. Cut-off of p value and fold change were as following: Data type: mRNA, p value: 1E-4, fold change: 2, gene rank: 10%

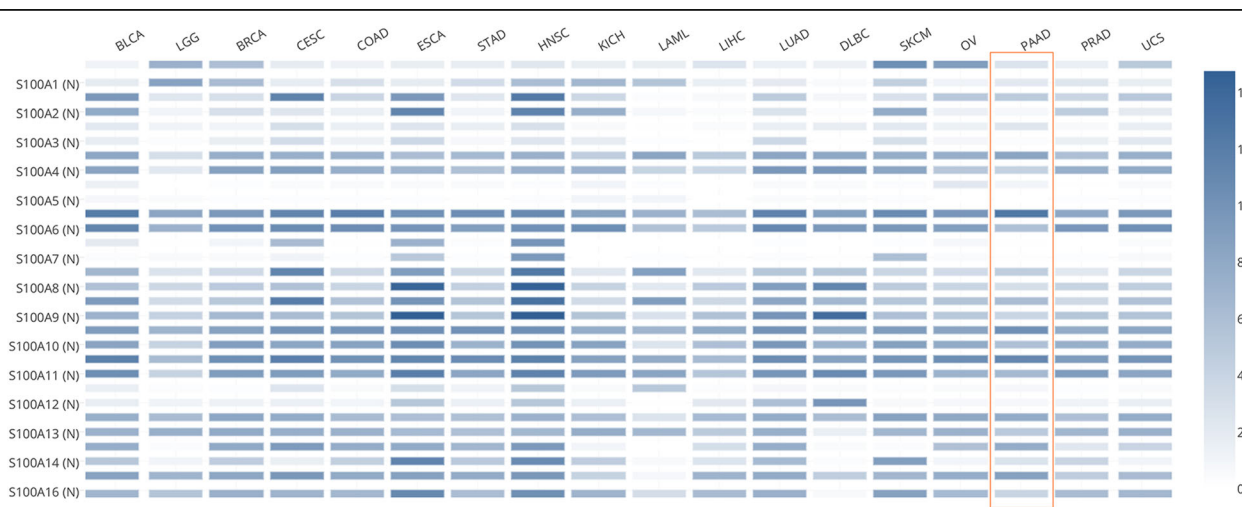


Fig. 2 The transcription levels of S100As in various tumors (GEPIA 2). The density of color in each block represents the median expression value of a gene in a given tissue, normalized by the maximum median expression value across all blocks. Different genes in same tumors or normal tissues can be compared in one plot. PDAC, Pancreatic ductal adenocarcinoma; BLCA, Bladder Urothelial Carcinoma; LGG, Brain Lower Grade Glioma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, Colon adenocarcinoma; ESCA, Esophageal carcinoma; STAD, Stomach adenocarcinoma; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; LAML, Acute Myeloid Leukemia; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; SKCM, Skin Cutaneous Melanoma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PRAD, Prostate adenocarcinoma; UCS, Uterine Carcinosarcoma

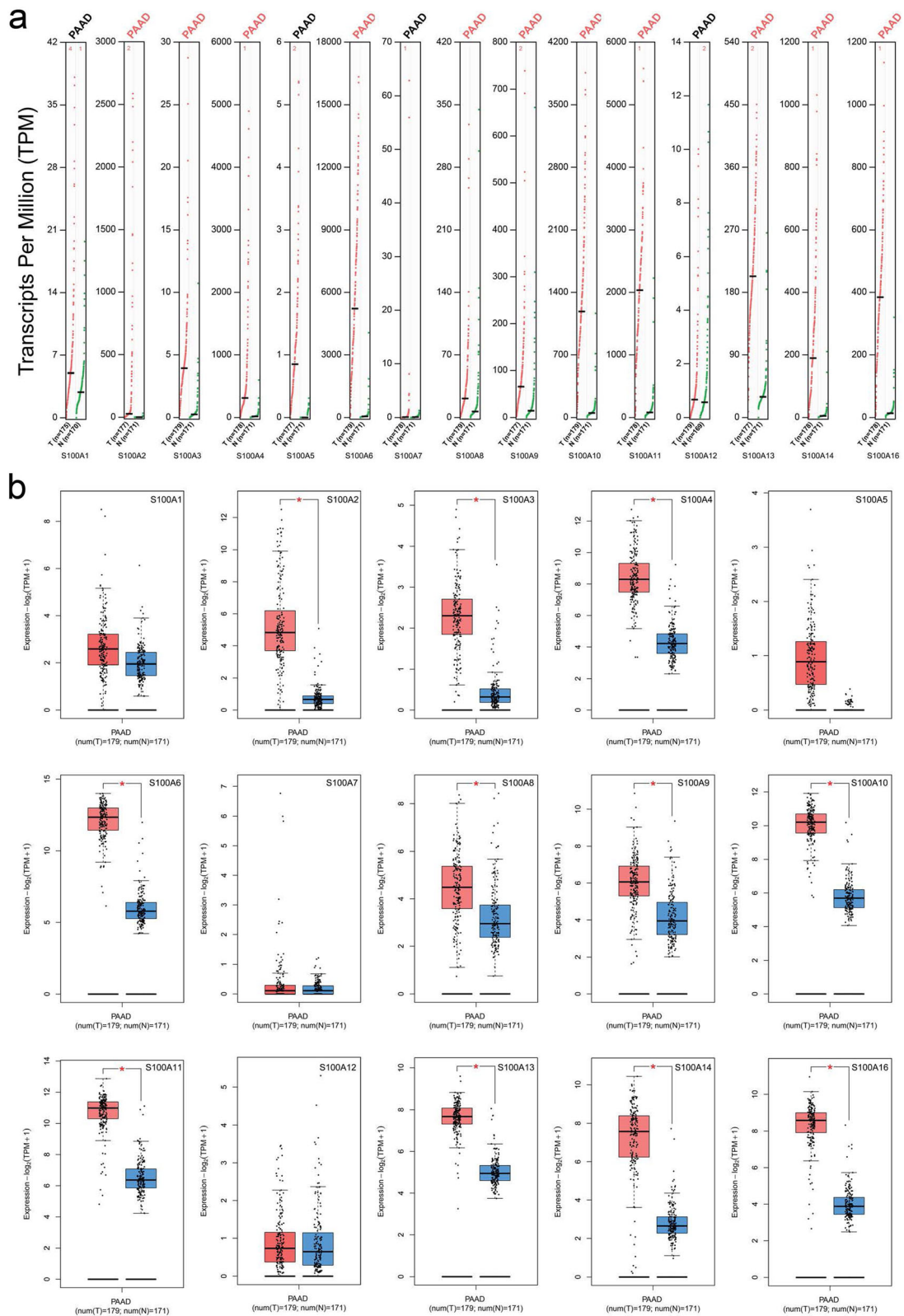
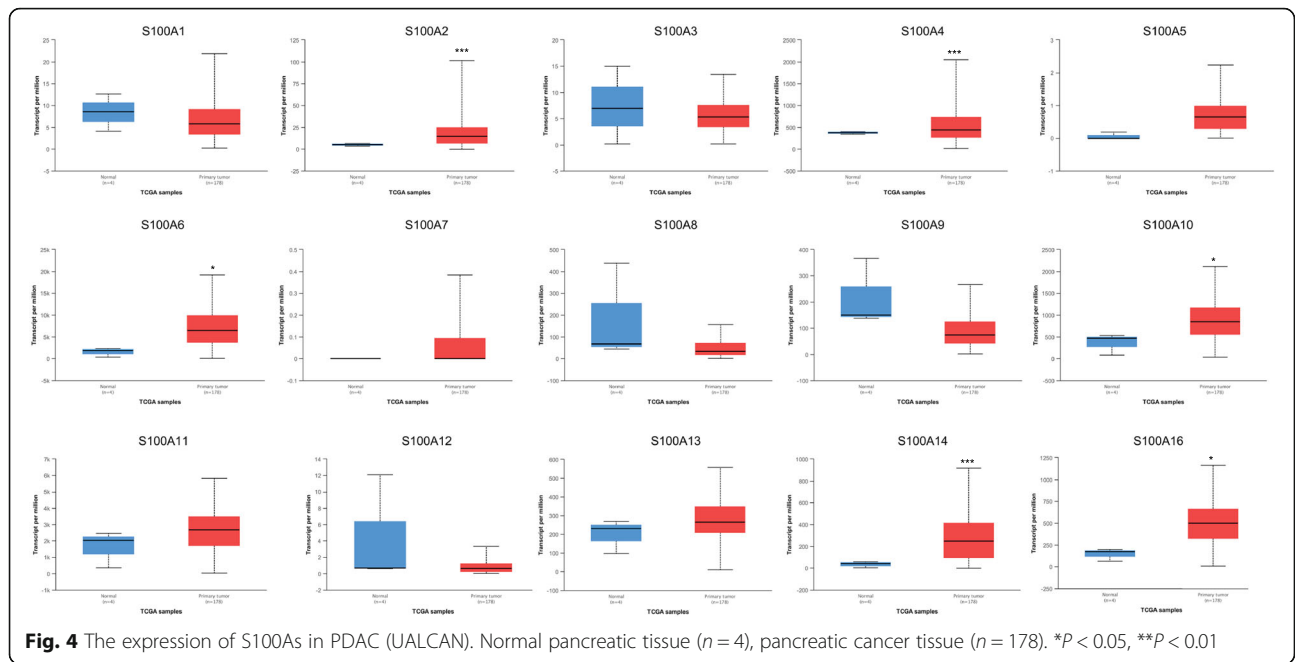


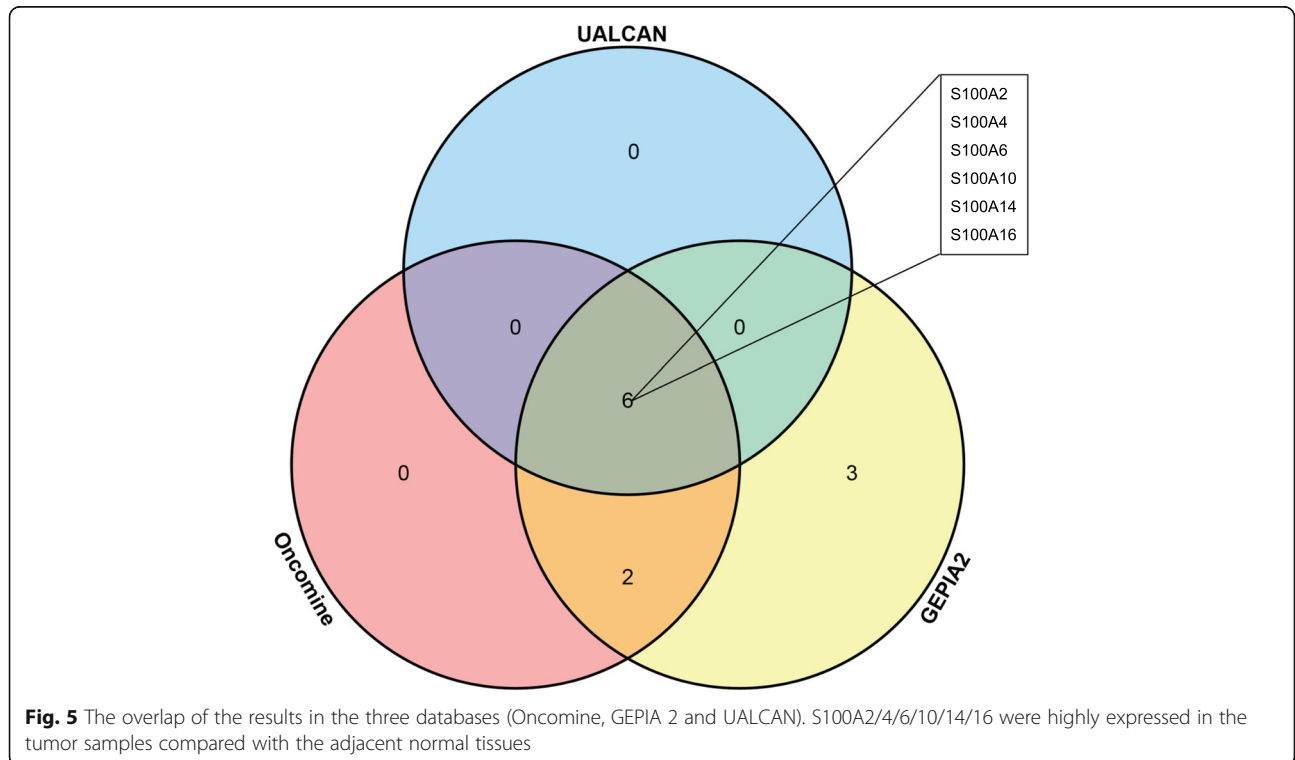
Fig. 3 The expression of S100As in PDAC (GEPiA2). The expression of S100As in PDAC (a, scatter diagram; b, box plot). * $P < 0.05$. T: pancreatic tumor tissue; N: Normal pancreatic tissue

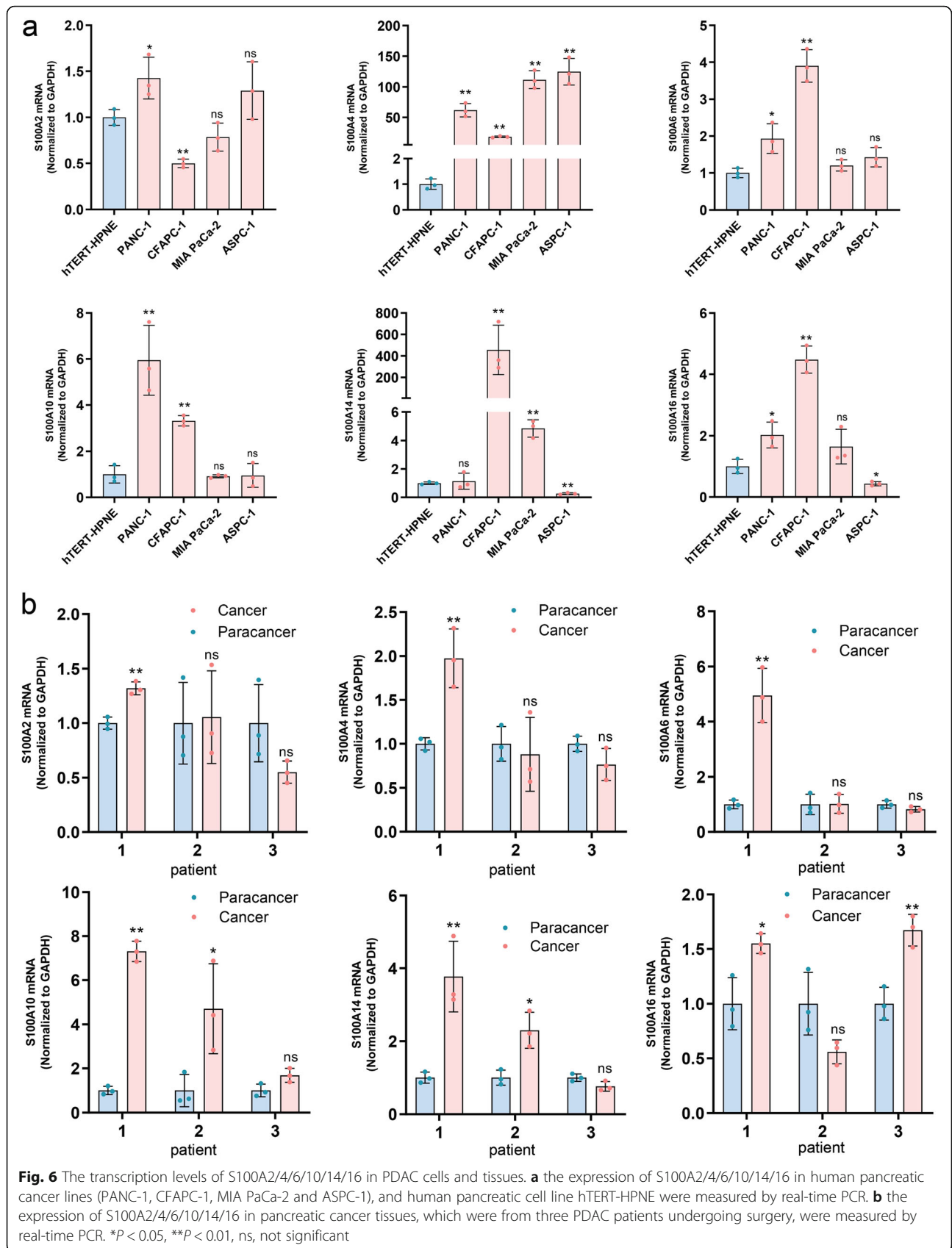


from Oncomine database, and the results indicated that, S100A2/4/6/10/14/16 were over-expressed in PDAC samples compared with normal pancreas samples (Fig. 4). P values less than 0.05 between the indicated groups considered statistically.

The transcription levels of S100A2/4/6/10/14/16 in PDAC cells and tissues

Based on the results above, we chose the overlap of the results in the three databases (Oncomine, GEPIA 2 and UALCAN), as shown in Fig. 5, S100A2, S100A4,





S100A6, S100A10, S100A14 and S100A16 were over-expressed in PDAC versus normal samples. After completing the analysis based on bioinformatics, we further tried to explore the expression patterns of S100As in pancreatic cancer cell lines and pancreatic cancer tissue.

We first analyzed the expression of S100A2/4/6/10/14/16 in human pancreatic cancer lines (PANC-1, CFAPC-1, MIA PaCa-2 and ASPC-1), and human pancreatic cell line hTERT-HPNE served as control by real-time PCR. As shown in Fig. 6a (* $P < 0.05$, ** $P < 0.01$, ns $P > 0.05$), the expression level of S100A2 was higher in PANC-1 than that in hTERT-HPNE. S100A4 was up-regulated in all four pancreatic cancer cells relative to hTERT-HPNE cells. S100A6 and S100A10 were up-regulated in PANC-1, CFAPC-1 cells compared with normal control. S100A14 mRNA expression was up-regulated in CFAPC-1, MIA PaCa-2 cells and down-regulated in ASPC-1 compared to normal cells. Compared with hTERT-HPNE cells, the mRNA expression of S100A16 was significantly up-regulated in PANC-1, CFAPC-1 cells, and down-regulated in ASP-1 cells. We then tested the expression of S100A2/4/6/10/14/16 in pancreatic cancer tissues, which were from three PDAC patients undergoing surgery. From the results, we noted high expression of S100A2/4/6/10/14/16 in the pancreatic cancer tissue of patient one. However, in pancreatic cancer tissue of patient 2, only S100A10 and S100A14 are highly expressed compared with adjacent cancer (Fig. 6b). Next, IHC was used to measure S100As protein level in pancreatic cancer tissues ($n = 43$) and their counterparts ($n = 28$). The data showed that S100A4, S100A14 and S100A16 were significantly higher in PDAC tissues compared with their counterparts, but the protein level of S100A2, S100A6 and S100A10 were similar between cancer tissues and their counterparts (Fig. 7).

Analysis of S100As methylation levels

The level of DNA methylation is inversely proportional to transcriptional activity and affects genome stability. Additionally, the methylation levels of the promoter region have a greater effect on gene expression activity. We first measured the promoter methylation level of the S100A2/4/6/10/14/16 in PDAC by searching the UALCAN database. The results showed that (Fig. 8a), compared with normal tissues, the promoter methylation level of S100A2, S100A4, S100A6, S100A10, S100A14, S100A16 in PDAC ($n = 10$) were lower compared with normal tissue ($n = 184$), ** $P < 0.01$, *** $P < 0.001$. We further used the MEXPRESS database to analyze specific methylation sites. According to MEXPRESS's algorithm, PDAC tissue samples were divided into two groups according to their S100As expression level (blue indicates a low level of expression, while violet indicates a high level of expression). The horizontal line at each probe position represents the median percentage of methylation (β value), while the vertical line

represents the range between 25 and 75%. As shown in Fig. 8b, in PDAC tissues, the expression of S100A2/4/6/10/14/16 were negative correlation (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$) to the level of promoter methylation.

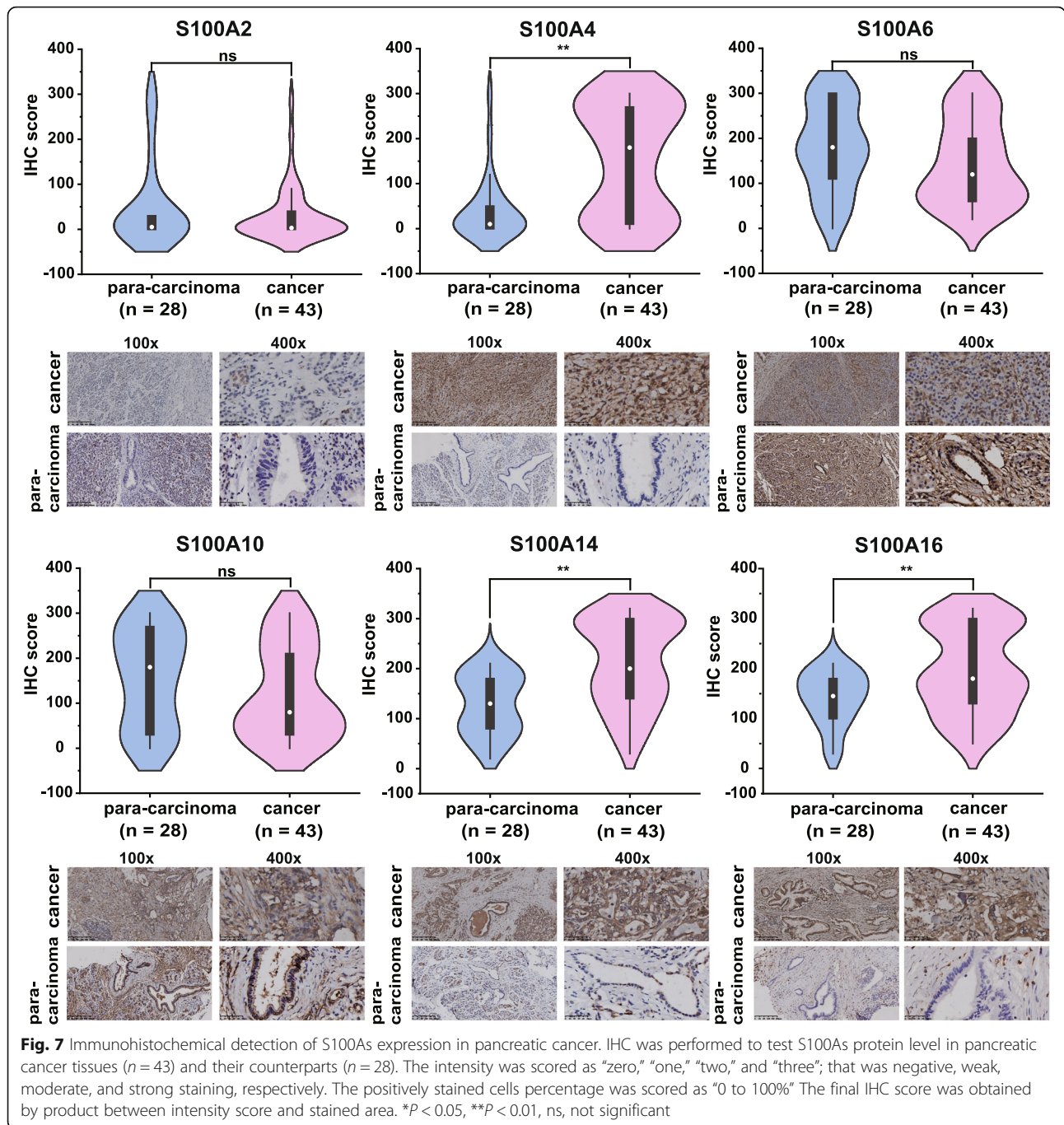
Correlation analysis of mRNA expression of S100As and clinicopathological parameters in pancreatic cancer

After finding that the transcriptional level of S100A2/4/6/10/14/16 was highly expressed in pancreatic cancer, we further analyzed the expression of S100A2/4/6/10/14/16 with tumor stage for pancreatic cancer using GEPIA2 and Linkedomics database. As shown in Fig. 9a, S100A4/6/10/14/16 groups significantly varied, whereas S100A2 did not significantly (P values less than 0.05 was considered to be statistically significant). The results of Linkedomics database analysis from different sample sources are consistent with those of GEPIA2 (Fig. 9b). In short, we could found from the results that, mRNA expressions of S100A4/6/10/14/16 were significantly associated with tumor grades, which may have potential value of clinical transformation.

Prognostic value analysis of S100As mRNA expression in pancreatic cancer

In order to explore the critical efficiency of S100A2/4/6/10/14/16 in the survival of patients with pancreatic cancer, GEPIA2, UALCAN and Kaplan-Meier Plotter tools, whose data sources include GEO, EGA, and TCGA, were used to analyze the prognostic values (OS) of the mRNA expression of S100As in pancreatic cancer patients. Based on the high expression of S100As members in pancreatic cancer, the association between mRNA expression of S100A2/4/6/10/14/16 and prognosis of pancreatic cancer patients were further analyzed. As shown in Fig. 10a, the results of the GEPIA2 database indicate that, higher mRNA expression of S100A2/10/14/16 were significantly associated with shorter OS in PDAC patients ($P < 0.05$). The results analyzed by the UALCAN database dedicated that (Fig. 10b), expression levels of S100A14 and S100A16 were significantly correlated with the OS of PDAC patients (* $P < 0.05$). In addition, Kaplan-Meier plotter were used to analyze the prognostic values of S100A2/4/6/10/14/16 in PDAC. As shown in Fig. 10c, mRNA expression of S100A2/4/6/10/14/16 were inversely correlated with patient survival ($P < 0.01$).

Next, we further evaluated the prognostic value of S100A2/4/6/10/14/16 in 43 cases of PDAC patients. It was revealed from Kaplan-Meier analyses that the S100A4, S100A14 and S100A16-high patients have worse overall survival rate than S100A4, S100A14 and S100A16-low patients. However, there are no differences of overall survival rate between patients who have different protein level of S100A2, S100A6 and S100A10 (Fig. 10d). The correlation between S100A2, S100A4, S100A6, S100A10, S100A14 and



S100A16 level and clinicopathologic characteristics were shown in supplementary Tables S3–8, respectively.

Correlation analysis of drugs targeting S100As

To explore the potential use of S100As as host targets, we queried DrugBank (version 5.1.6) to identify drugs targeting them. As shown in Table 1, there are seven kinds of drugs targeting S100A2, which can be further divided into antagonist, inactivator, and regulator groups. There are two drugs (Trifluoperazine and

Copper) targeting S100A4. In addition, three drugs that target S100A6 have been approved. Three drugs with actions of ligand were identified as drugs targeting S100A16. While, drugs targeting S100A10 and S100A14 have not yet been found now.

Genetic mutations in S100As and their associations with OS and (progression-free survival) (PFS) of PDAC patients
 Since gene mutation is an important factor affecting its expression, we further analyzed the genetic alteration in

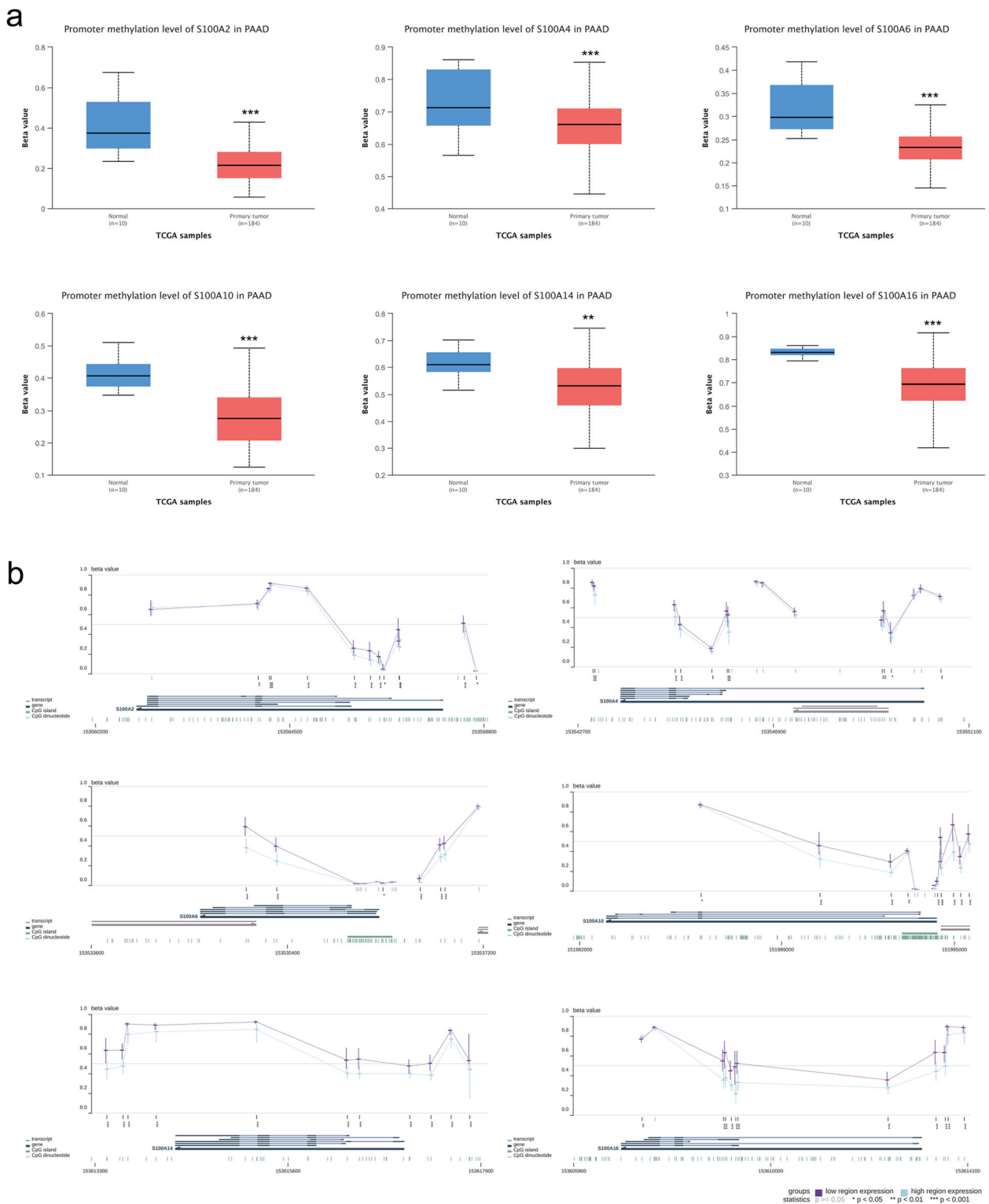
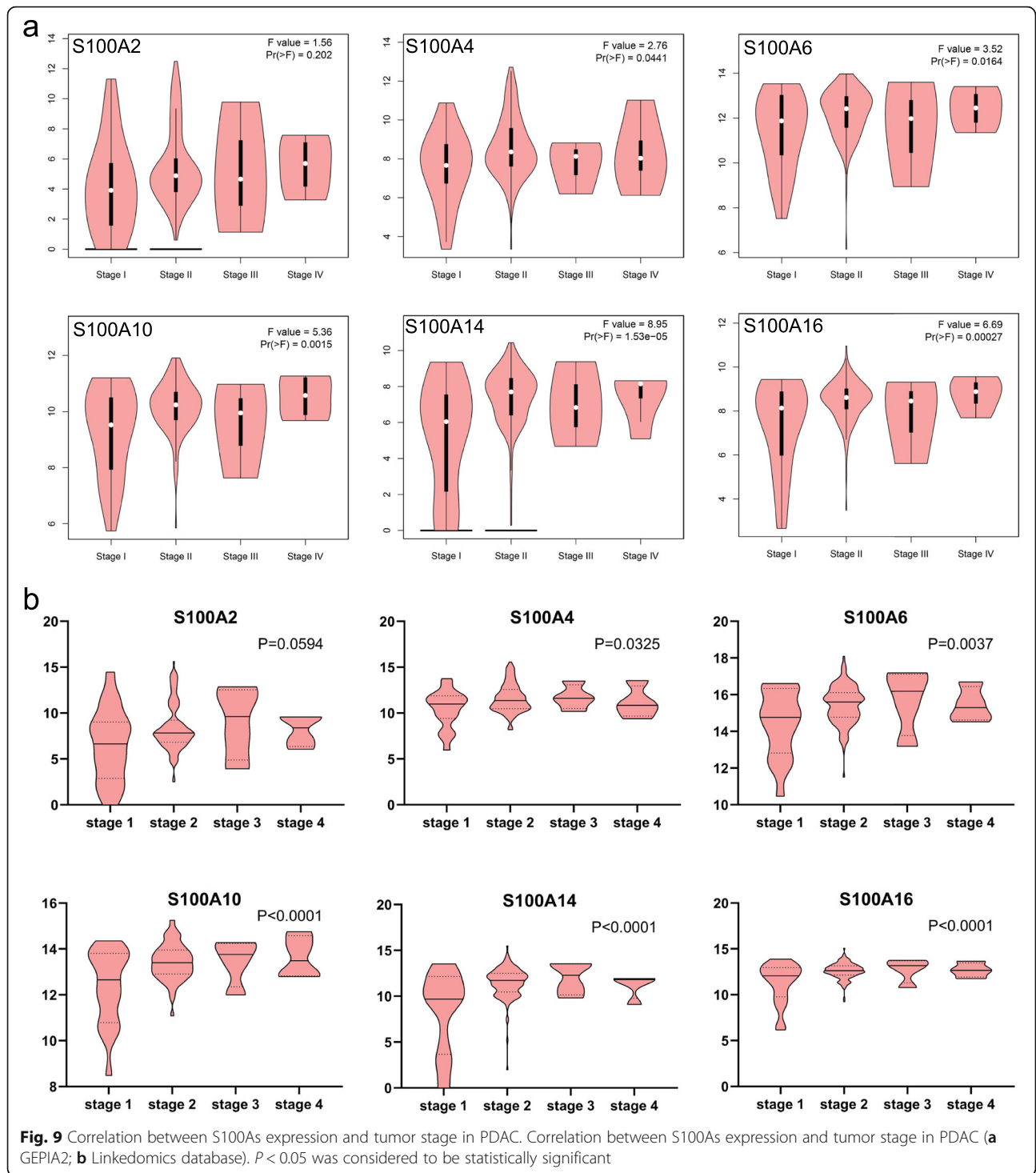


Fig. 8 Analysis of S100As methylation levels in PDAC. **a** promoter methylation level of S100As were detected by UALCAN database. ****** $P < 0.01$, ******* $P < 0.001$. **b** specific methylation sites were analyzed by MEXPRESS database (blue indicates a low level of expression, while violet indicates a high level of expression)



S100A2/4/6/10/14/16 and their correlation with OS and PFS in patients with pancreatic cancer by using cBioPortal online tool. As shown in Fig. 11a-b, S100A2/4/6/10/14/16 were altered in 55 samples of 177 patients with pancreatic cancer (31%). The mutation rates of S100A2, S100A4, S100A6, S100A10, S100A14 and S100A16 were 9, 12, 11, 14, 8, 16 and 13%, respectively. In additional,

results from Kaplan-Meier plot and log-rank test indicated that genetic alteration in S100A2/4/6/10/14/16 were associated with shorter OS (Fig. 11c, $p = 0.0343$) and PFS (Fig. 11d, $p = 0.0356$) of PDAC patients. Survival analysis showed that gene mutation of S100As might also significantly affect the prognosis of PDAC patients.

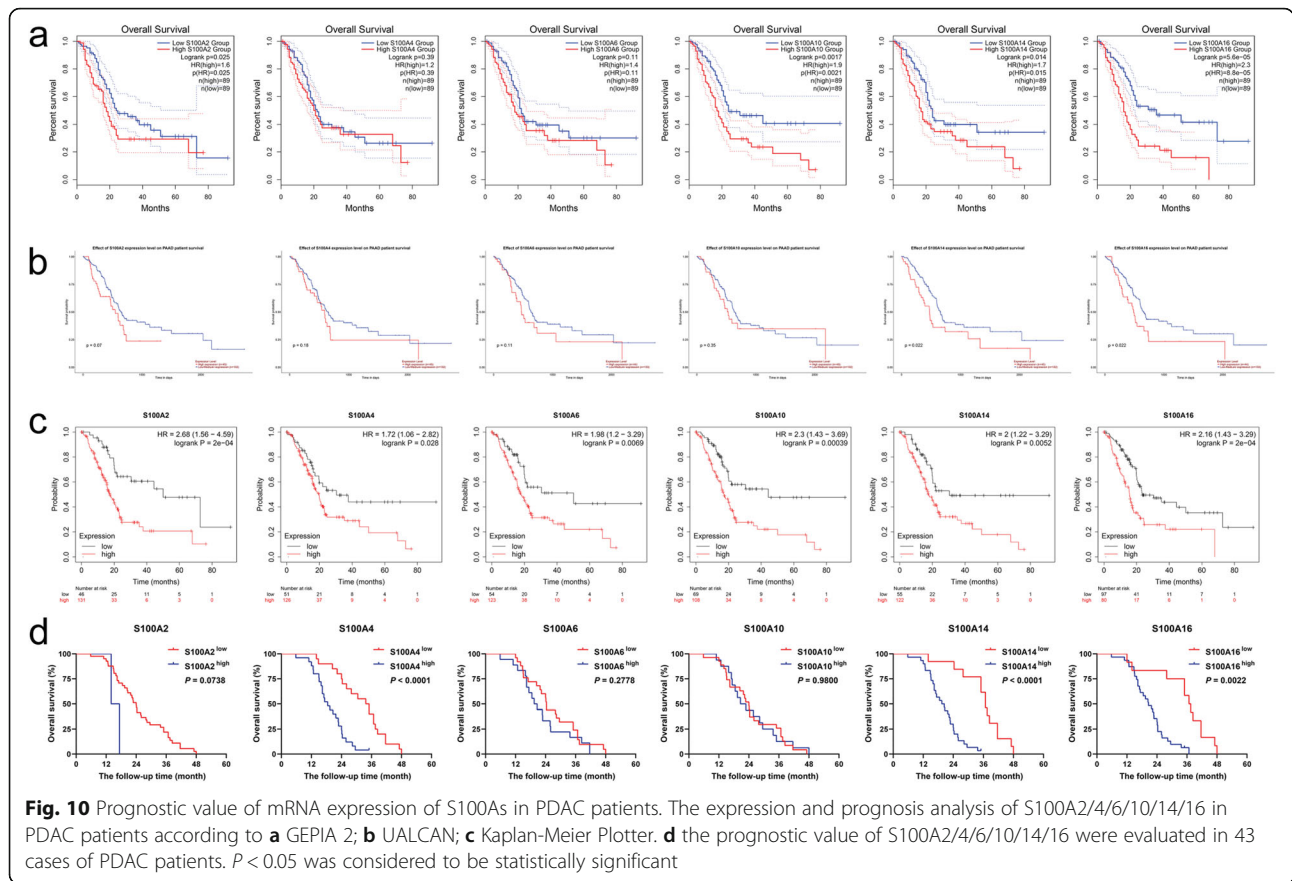
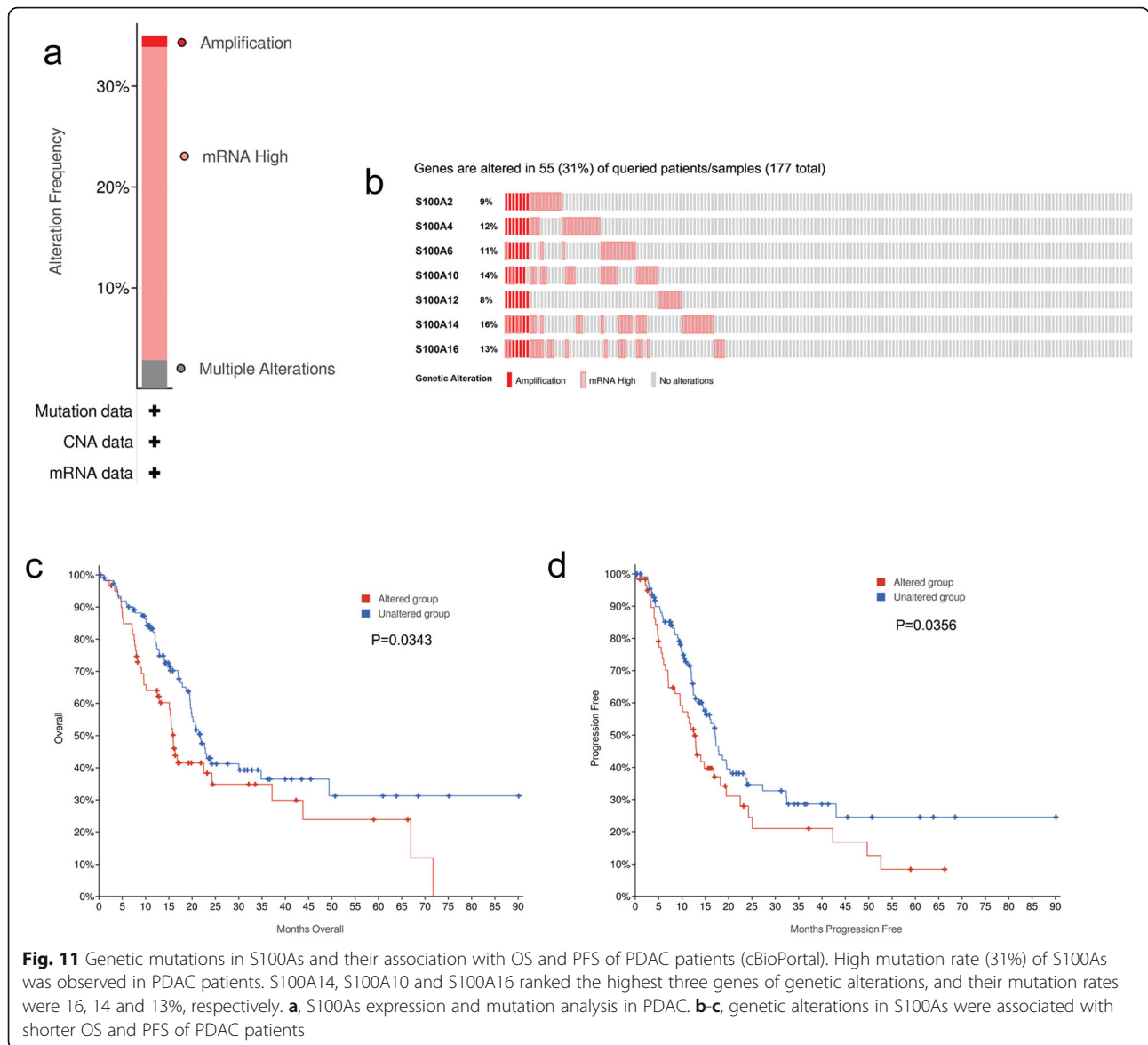


Table 1 Drug relations of S100As

Gene	Drug relations				
S100A2	Drugbank ID	Drug Name	Drug Group	Actions	Reference
	DB00768	Olopatadine	approved	Antagonist	[15]
	DB01373	Calcium	nutraceutical		[16]
	DB01593	Zinc	approved, investigational		[16]
	DB09130	Copper	approved, investigational		[17]
	DB14487	Zinc acetate	approved, investigational		[16]
	DB14533	Zinc chloride	approved, investigational	Inactivator, regulator	[16]
	DB14548	Zinc sulfate, unspecified form	approved, experimental	Inactivator, regulator	[16]
S100A4	DB00831	Trifluoperazine	approved, investigational	Inhibitor	[18]
	DB09130	Copper	approved, investigational		[17]
S100A6	DB11093	Calcium citrate	approved, investigational	Ligand	UniProt: Protein S100-A6
	DB11348	Calcium Phosphate	approved	Ligand	UniProt: Protein S100-A6
	DB14481	Calcium phosphate dihydrate	approved		UniProt: Protein S100-A6
S100A10	NA				
S100A14	NA				
S100A16	DB11093	Calcium citrate	approved, investigational	Ligand	UniProt: Protein S100-A16
	DB11348	Calcium Phosphate	approved	Ligand	UniProt: Protein S100-A16
	DB14481	Calcium phosphate dihydrate	approved		UniProt: Protein S100-A16



Predicted functions and pathways of the genes that positively correlated to S100As in PDAC patients

Linkedomics database was used to search for significant genes positively correlated to S100As in pancreatic cancer samples. As shown in Table S2, we selected top25 genes, which were positively related to expression pattern of S100As, according to the person correlation. Furthermore, the positively correlated significant genes were loaded into Metascape to acquire gene ontology (GO) enrichment. As shown in Fig. 12a-b, the enriched ontology clusters colored by cluster ID, which predicted the functional roles of S100As based on molecular functions, cellular components and biological processes. We could find that GO:0050839 (cell adhesion molecule binding), GO:0005509 (calcium ion binding), GO:0009611 (response to wounding), GO:0008544 (epidermis

development) were significantly regulated (top 4) by the positively related to expression pattern of S100As in pancreatic cancer (Fig. 12c). In addition, KEGG pathways enrichment analysis was performed by the online free tool provided by omicshare. The heatmap are colored by their P-values (Fig. 12d). The result suggested that, focal adhesion; arrhythmogenic right ventricular cardiomyopathy and hypertrophic cardiomyopathy were significantly associated with the genes which were positively correlated significant Genes of S100As. STRING, a database that can analyze the interaction between known proteins and predicted proteins, was used to predict protein-protein interactions of the genes positively related to S100As expression described above. We can see from Fig. 12e, S100As family members have close interactions themselves. In addition, they also interact with

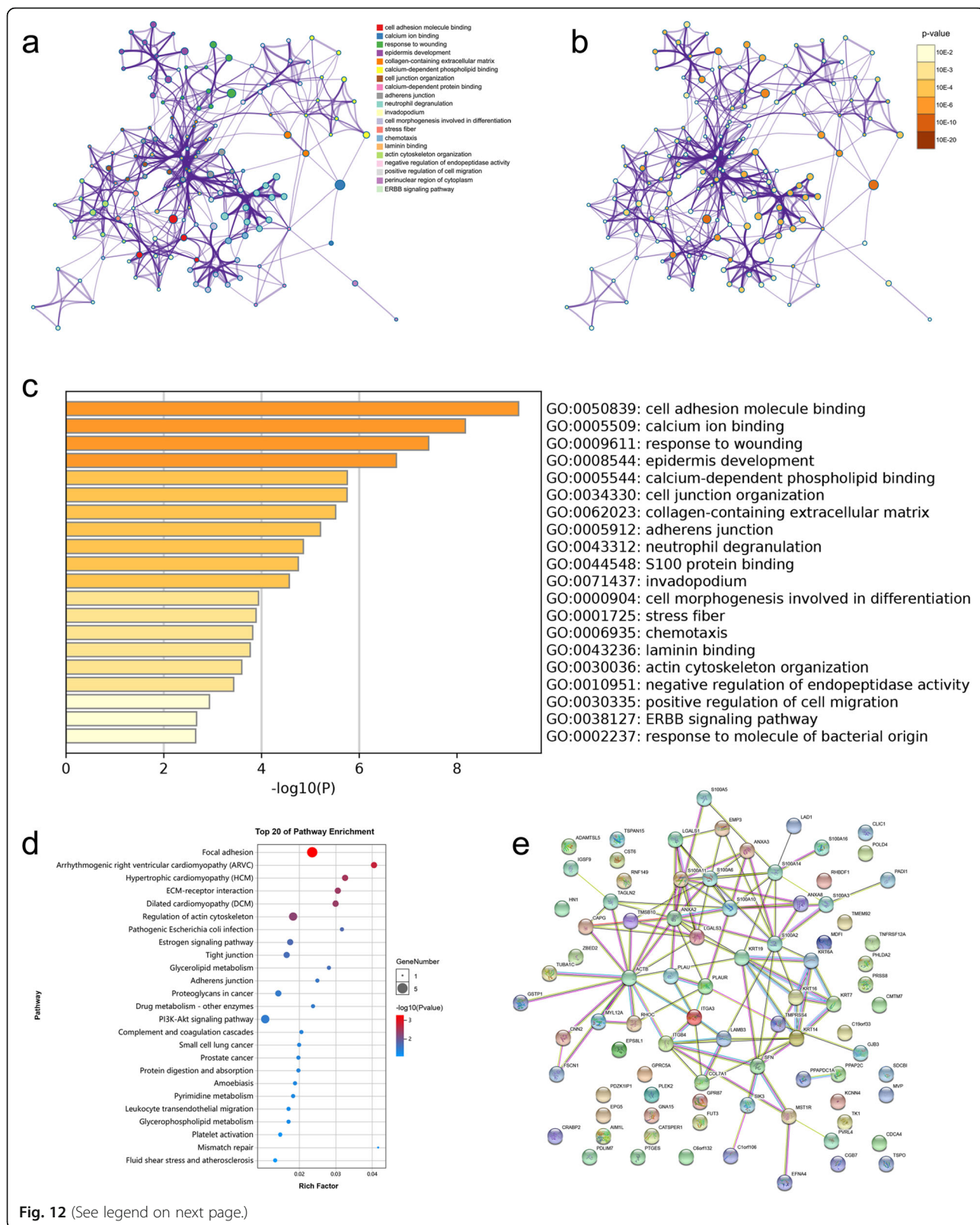


Fig. 12 (See legend on next page.)

(See figure on previous page.)

Fig. 12 Protein-protein interaction enrichment analysis. Predicted functions and pathways of significant genes positively correlated to S100As in pancreatic cancer. Network of enriched terms: **a** enriched ontology clusters colored by cluster ID (legend on right), where nodes that share the same cluster ID are typically close to each other; **b** colored by *p*-value, where terms containing more genes tend to have a more significant *p*-value. **c** Bar graph of enriched terms across significant genes positively correlated to S100As, colored by *p*-values. **d** top 20 of pathway enrichment of genes positively correlated to S100As. **e** information of interactions of genes positively correlated to S100As were extracted from online database STRING

tumor cell adhesion-related proteins such as ANXA2, LGALS3, ANXA8, and ACTB.

Discussion

PDAC remains a treatment-refractory malignancy with poor prognosis. The discovery of some promising molecular targets for the treatment of PDAC provided a new direction for the diagnosis and treatment of PDAC. For example, Jin et al. reported that, PES1 might be a promoting factor of tumor growth and a PDAC prognosis associated protein [19]. Goehrig et al. found that, β ig-h3 stromal-derived protein is a main factor driving the immune paracrine interaction mechanism of PDAC, which might be novel target in PDAC treatment [20]. However, compared with other malignant tumors, effective diagnosis and treatment targets of pancreatic cancer are still quite lacking. It is urgent to develop screening procedures for early detection and more efficacious treatment strategies for PDAC [21]. Bioinformatics provides an opportunity to quest for the genetic alterations in pancreatic cancer, and proved a useful way to identify new biomarkers in pancreatic cancer. To verify the results from bioinformatics further, we tested the expression of S100As in pancreatic cancer at the cellular and tissue levels by real-time PCR and immunohistochemistry. Unfortunately, pancreatic cancer tissues from PDAC patients undergoing surgery are scarce due to the few cases of pancreatic cancer surgery. Because the sample size is too small, the heterogeneity between patients obscures some of the original expression levels. Here, we investigated the mRNA expression and prognostic values of S100A family in PDAC for the first time. We expect that our findings could help us to understand the mechanism of PDCA, improve the clinical treatment design, and improve the reliability of prognostic assessment.

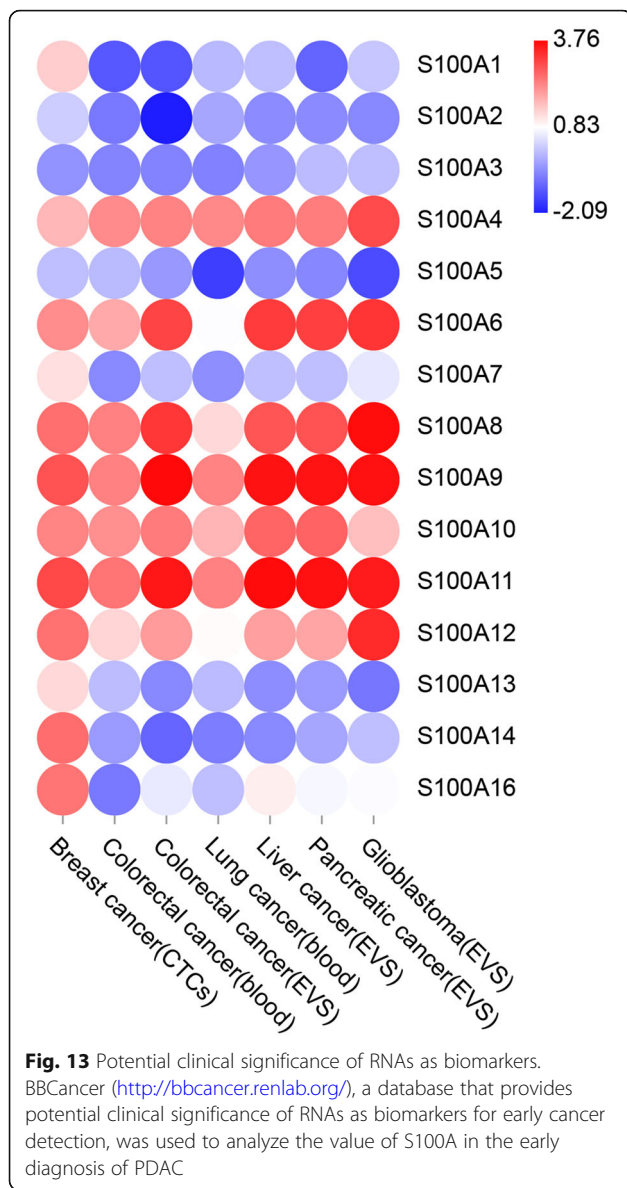
The S100 protein family belong to a family of calcium binding proteins (CaBPs) and contain 25 known members, which have a high degree of sequence and structural similarity. The S100A subfamily is among the most distinctive of EF-hand CaBPs and only found in vertebrates. Studies have shown that multiple members of the S100As are abnormally expressed in tumors and are widely involved in processes related to tumorigenesis and progression [22, 23]. However, the function of most S100A has not been fully characterized, especially in

pancreatic cancer, the research on S100A is still very limited.

Among the S100As, S100A4 is the most intensively studied. S100A4 is a member of the S100As, which contains two EF-hand calcium-binding motifs. Among its related pathways are Ca^{2+} , cAMP and Lipid Signaling. Fei et al. reported that, the expression of S100A4 was positively correlated with the malignancy degree and the level in lymph node metastasis of human colorectal cancer [24]. Recently, a study about hepatocellular carcinogenesis indicated that, high expression of S100A4 in the fibrotic region of was liver tumor tissue observed, and the expression of S100A4 was associated with advanced disease severity [25]. It was also reported that, S100A4 was one of the key factors in EMT mediated by Shh-Gli1 signaling pathway in PDAC [26]. In present study, Oncomine, GEPIA 2 and UALCAN datasets revealed that the expression of S100A4 was higher in PDAC than in normal pancreatic tissues. We also noticed that, S100A4 was up-regulated in all four pancreatic cancer cells relative to hTERT-HPNE cells. In addition, BBCancer (<http://bbcancer.renlab.org/>), a database that provides potential clinical significance of RNAs as biomarkers for early cancer detection, was used to analyze the value of S100A in the early diagnosis of pancreatic cancer. We can see from Fig. 13, S100A4/6/8/9/10/11/12/16 were up regulated in extracellular vesicles of pancreatic cancer, which might have diagnostic value.

As an important paralog of S100A4, S100A2 is reported to be abnormally highly expressed in various tumors, such as colorectal cancer [27] and lung cancer [28]. In our study, we indicated that the expression of S100A2 in pancreatic cancer tissues was higher than that in normal pancreatic tissues. It has also been suggested as a negative prognostic biomarker in pancreatic cancer [29]. However, the prognostic role of S100A2 in pancreatic cancer remains to be studied. GO annotations of S100A2 include calcium ion binding and identical protein binding. Among its related pathways are p53 pathway.

S100A6 is also an up-regulated S100As in pancreatic cancer confirmed by bioinformatics and in vitro experiments. Results of BBCancer also showed that, S100A6 was up-regulated in extracellular vesicles of PDAC. Diseases associated with S100A6 include Pilomatrixoma and Retinitis Pigmentosa. The latest research also



showed that S100A6 is involved in the process of neurodegenerative diseases [30]. GO annotations of S100A6 were calcium ion binding and calcium-dependent protein binding. Among its related pathways are prostaglandin synthesis and regulation and DNA damage. Further research is needed to explain the role of S100A6 in the development of pancreatic cancer.

Diseases associated with S100A10 include Trachea Leiomyoma and Barrett’s Adenocarcinoma. There is evidence that S100A10 has potential value as a biomarker that represents the high-grade cell state of breast cancer [31]. Among its related pathways are Prostaglandin Synthesis and Regulation and Response to elevated platelet cytosolic Ca²⁺. GO annotations related to this gene include calcium ion binding and lipid binding. An important paralog of this gene are S100A1 and S100A14. In our

study, both bioinformatics analysis and experimental verification indicated that S100A10 was up-regulated in pancreatic cancer (vs normal controls/paraneoplastic tissue). Low expression of S100A10 was associated significantly with better survival in pancreatic cancer by univariate analysis (*P* < 0.05). The above results reveal the potential value of S100A10 as a diagnosis and treatment of pancreatic cancer.

Both S100A14 and S100A16 were up-regulated in pancreatic cancer, which was also verified in our experiments. S100A14 has a potential to be clinically useful as prognostic biomarker in several cancer types. For example, Meta-analysis from Hu et al. suggested that S100A14 overexpression might be a predictive biomarker for poor prognosis in patients with breast cancer and ovarian cancer [32]. However, the role of S100A14 in pancreatic cancer is rarely reported. Survival analysis showed that, low expression of S100A14 was associated with significantly better survival. S100A16 was discovered late and its research is not deep enough. In this study, we found that S100A16 is upregulated in PDAC, and its low expression was closely related to a better survival.

Given that the signaling pathways involved in calcium ions have been shown to change the sensitivity of chemotherapy drugs [33]. We therefore also analyzed the drug relations of S100As by Drugbank. As listed in Table 1, Drugs targeted at S100As include Olopatadine, Trifluoperazine and calcium citrate. These drugs are mainly used as inactivators, regulators, corrosion inhibitors and ligands to play the role of drugs. Therefore, S100As have certain potential value in the development of pancreatic cancer chemotherapy resistance and new drug development.

DNA methylation generally occurs at the CpG site, and the conversion of cytosine to 5-methylcytosine is catalyzed by DNA methyltransferase in human genes. Approximately 80% of CpG sites are methylated, and the level of DNA methylation is inversely proportional to transcriptional activity and affects genome stability. Compared with the methylation of the gene-coding region, the methylation of the promoter region has a greater effect on gene expression activity. In our study, we found that the methylation level of the promoter region was negatively correlated with its expression, which may be a reason for the abnormal expression of S100As in pancreatic cancer.

Naturally, there are also many limitations in our research. First, although high mRNA expressions of S100A2/4/6/10/14/16 were independent prognostic factors for shorter OS of pancreatic cancer patients, the online databases used in this study lacked proteome-level data. Further studies consist of larger sample sizes (more clinical samples and different pancreatic cancer cell

lines) are required to validate the findings above and to explore the clinical application of S100A2/4/6/10/14/16 in the treatment of PDAC. In addition, we did not explore the potential molecular mechanism of S100As in the occurrence and development of pancreatic cancer. Future research worth to investigate the detailed mechanism between S100As and PDAC.

Conclusions

In summary, we comprehensively explored the expression pattern and prognostic value of S100As in PDAC, and provided an understanding of the molecular biological properties of PDAC. The current research results indicate that the increased expression of S100A2/4/6/10/14/16 in PDAC tissue may play an important role in the occurrence of pancreatic cancer. The high expression of S100A2/10/14/16 could also be used as molecular markers to identify high-risk subgroups of PDAC patients. S100A2/10/14/16 were potential therapeutic targets for PDAC, and regulating the expression of S100A2/4/6/10/14/16 might be potential prognostic markers to improve PDAC survival rate and prognostic accuracy. It is undeniable that cancer biomarkers prediction and analysis based on bioinformatics can only provide preliminary data, whose clinical applications are yet to be further developed. Further experimental verification and molecular mechanism research are also the focus of our future work.

Abbreviations

PDAC: Pancreatic ductal adenocarcinoma; CaBPs: Calcium binding proteins; GO: Gene ontology; IHC: Immunohistochemistry; OS: Overall survival

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-021-08769-6>.

Additional file 1: Supplementary Table S1. Primers used in this study.

Additional file 2: Supplementary Table S2. Positively Correlated Significant Genes.

Additional file 3: Supplementary Table S3–8. Basic characteristics of 43 PDAC patients.

Acknowledgements

None.

Authors' contributions

Conception and design: All authors; Administrative support: H Li, J Wang, Z Yu; Provision of study materials or patients: All authors; Collection and assembly of data: All authors; Data analysis and interpretation: H Li, J Wang, X Jin; Manuscript writing: H Li, J Wang, X Jiang; Final approval of manuscript: All authors.

Funding

This work is jointly supported by the National Natural Science Foundation of China (Grant No. 82003216) and the Fundamental Research Funds for the Central Universities (Grant Nos. lzujbky-2017-78).

Availability of data and materials

The datasets presented in this study are available in online repositories.

Declarations

Ethics approval and consent to participate

All procedures involving human samples were approved by the Ethics Committee of Lanzhou University Second Hospital.

Consent for publication

Not applicable.

Competing interests

All of the authors declare that they have no competing interests.

Author details

¹School of Life Science and Engineering, Lanzhou University of Technology, Lanzhou 730050, China. ²Gansu Provincial Cancer Hospital, Gansu Provincial Academic Institute for Medical Research, Lanzhou 730050, China. ³Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730030, China. ⁴Cui-ying Biomedical Research Center, Lanzhou University Second Hospital, Lanzhou 730030, China. ⁵Department of General Surgery, Gansu Provincial Cancer Hospital Lanzhou University Second Hospital, No.82 Cui-ying Men, Lanzhou 730030, Gansu Province, China.

Received: 17 January 2021 Accepted: 5 September 2021

Published online: 16 September 2021

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424. <https://doi.org/10.3322/caac.21492>.
- Wang Z, Hausmann S, Lyu R, Li TM, Lofgren SM, Flores NM, et al. SETD5-coordinated chromatin reprogramming regulates adaptive resistance to targeted pancreatic cancer therapy. *Cancer Cell*. 2020;37(6):834–49. <https://doi.org/10.1016/j.ccell.2020.04.014>.
- Liermann J, Shinoto M, Syed M, Debus J, Herfarth K, Naumann P. Carbon ion radiotherapy in pancreatic cancer: a review of clinical data. *Radiother Oncol*. 2020;147:145–50. <https://doi.org/10.1016/j.radonc.2020.05.012>.
- Marchi S, Vitto VAM, Danese A, Wieckowski MR, Giorgi C, Pinton P. Mitochondrial calcium uniporter complex modulation in cancerogenesis. *Cell Cycle*. 2019;18(10):1068–83. <https://doi.org/10.1080/15384101.2019.1612698>.
- Cmoch A, Groves P, Palczewska M, Pikula S. S100A proteins in propagation of a calcium signal in norm and pathology. *Postepy Biochem*. 2012;58(4):429–36.
- Tong LLW, Lim RR, Chaurasia SS. S100A proteins as molecular targets in the ocular surface inflammatory diseases. *Ocul Surf*. 2014;12(1):23–31. <https://doi.org/10.1016/j.jtos.2013.10.001>.
- El-Rifai W, Moskaluk CA, Abdrabbo MK, et al. Gastric cancers overexpress S100A calcium-binding proteins. *Cancer Res*. 2002;62(23):6823–6.
- Gupta S, Hussain T, MacLennan GT, et al. Differential expression of S100A2 and S100A4 during progression of human prostate adenocarcinoma. *J Clin Oncol*. 2003;21(1):106–12. <https://doi.org/10.1200/JCO.2003.03.024>.
- Wang G, Wang X, Wang S, Song H, Sun H, Yuan W, et al. Colorectal cancer progression correlates with upregulation of S100A11 expression in tumor tissues. *Int J Color Dis*. 2008;23(7):675–82. <https://doi.org/10.1007/s00384-008-0464-6>.
- Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia*. 2017;19(8):649–58. <https://doi.org/10.1016/j.neo.2017.05.002>.
- Koch A, Jeschke J, Van Criekeing W, et al. MEXPRESS update 2019. *Nucleic Acids Res*. 2019;47(W1):W561–5. <https://doi.org/10.1093/nar/gkx445>.
- Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074–82. <https://doi.org/10.1093/nar/gkx1037>.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of

- systems-level datasets. *Nat Commun.* 2019;10(1):1523. <https://doi.org/10.1038/s41467-019-09234-6>.
14. Yu ZY, Jiang XY, Zhao RR, Qin JJ, Luo CJ, Ren YX, et al. Effect of KIF22 on promoting proliferation and migration of gastric cancer cells via MAPK-ERK pathways. *Chin Med J.* 2020;133(8):919–28. <https://doi.org/10.1097/CM9.0000000000000742>.
 15. Okada M, Tokumitsu H, Kubota Y, Kobayashi R. Interaction of S100 proteins with the antiallergic drugs, olopatadine, amlexanox, and cromolyn: identification of putative drug binding sites on S100A1 protein. *Biochem Biophys Res Commun.* 2002;292(4):1023–30. <https://doi.org/10.1006/bbrc.2002.6761>.
 16. Koch M, Bhattacharya S, Kehl T, Gimona M, Vašák M, Chazin W, et al. Implications on zinc binding to S100A2. *Biochim Biophys Acta.* 1773;2007(3): 457–70. <https://doi.org/10.1016/j.bbamcr.2006.12.006>.
 17. Haase-Kohn C, Wolf S, Lenk J, Pietzsch J. Copper-mediated cross-linking of S100A4, but not of S100A2, results in proinflammatory effects in melanoma cells. *Biochem Biophys Res Commun.* 2011;413(3):494–8. <https://doi.org/10.1016/j.bbrc.2011.08.132>.
 18. Malashkevich VN, Dulyaninova NG, Ramagopal UA, Liriano MA, Varney KM, Knight D, et al. Phenothiazines inhibit S100A4 function by inducing protein oligomerization. *Proc Natl Acad Sci U S A.* 2010;107(19):8605–10. <https://doi.org/10.1073/pnas.0913660107>.
 19. Jin X, Fang R, Fan P, Zeng L, Zhang B, Lu X, et al. PES1 promotes BET inhibitors resistance and cells proliferation through increasing c-Myc expression in pancreatic cancer. *J Exp Clin Cancer Res.* 2019;38(1):463. <https://doi.org/10.1186/s13046-019-1466-7>.
 20. Goehrig D, Nigri J, Samain R, Wu Z, Cappello P, Gabiane G, et al. Stromal protein β ig-h3 reprogrammes tumour microenvironment in pancreatic cancer. *Gut.* 2019;68(4):693–707. <https://doi.org/10.1136/gutjnl-2018-317570>.
 21. Collaborators GPC. The global, regional, and national burden of pancreatic cancer and its attributable risk factors in 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet Gastroenterol Hepatol.* 2019;4:934–47.
 22. Hua X, Zhang H, Jia J, Chen S, Sun Y, Zhu X. Roles of S100 family members in drug resistance in tumors: status and prospects. *Biomed Pharmacother.* 2020;127:110156. <https://doi.org/10.1016/j.biopha.2020.110156>.
 23. Gonzalez LL, Garrie K, Turner MD. Role of S100 proteins in health and disease. *Biochim Biophys Acta, Mol Cell Res.* 1867;2020(6):118677. <https://doi.org/10.1016/j.bbamcr.2020.118677>.
 24. Fei F, Liu K, Li C, du J, Wei Z, Li B, et al. Molecular mechanisms by which S100A4 regulates the migration and invasion of PGCCs with their daughter cells in human colorectal Cancer. *Front Oncol.* 2020;10:182. <https://doi.org/10.3389/fonc.2020.00182>.
 25. Li Y, Wang J, Song K, Liu S, Zhang H, Wang F, et al. S100A4 promotes hepatocellular carcinogenesis by intensifying fibrosis-associated cancer cell stemness. *Oncoimmunology.* 2020;9(1):1725355. <https://doi.org/10.1080/2162402X.2020.1725355>.
 26. Xu X, Su B, Xie C, Wei S, Zhou Y, Liu H, et al. Sonic hedgehog-Gli1 signaling pathway regulates the epithelial mesenchymal transition (EMT) by mediating a new target gene, S100A4, in pancreatic cancer cells. *PLoS One.* 2014;9(7):e96441. <https://doi.org/10.1371/journal.pone.0096441>.
 27. Sabates-Bellver J, Van der Flier LG, de Palo M, et al. Transcriptome profile of human colorectal adenomas. *Mol Cancer Res.* 2007;5(12):1263–75. <https://doi.org/10.1158/1541-7786.MCR-07-0267>.
 28. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A.* 2001; 98(24):13790–5. <https://doi.org/10.1073/pnas.191502998>.
 29. Bachet JB, Maréchal R, Demetter P, Bonnetain F, Cros J, Svrcek M, et al. S100A2 is a predictive biomarker of adjuvant therapy benefit in pancreatic adenocarcinoma. *Eur J Cancer.* 2013;49(12):2643–53. <https://doi.org/10.1016/j.ejca.2013.04.017>.
 30. Filipek A, Leśniak W. S100A6 and its brain ligands in neurodegenerative disorders. *Int J Mol Sci.* 2020;21(11). <https://doi.org/10.3390/ijms21113979>.
 31. Arai K, Iwasaki T, Sonoda A, Endo A. Membranous overexpression of S100A10 is associated with a high-grade cellular status of breast carcinoma. *Med Mol Morphol.* 2020;53(2):104–14. <https://doi.org/10.1007/s00795-019-00236-3>.
 32. Hu L, Kong F, Pan Y. Prognostic and clinicopathological significance of S100A14 expression in cancer patients: a meta-analysis. *Med Sci Sports Exerc.* 2019;98(28):e16356. <https://doi.org/10.1097/MD.00000000000016356>.
 33. Hatoum D, Yagoub D, Ahadi A, Nassif NT, McGowan EM. Annexin/S100A protein family regulation through p14ARF-p53 activation: a role in cell survival and predicting treatment outcomes in breast cancer. *PLoS One.* 2017;12(1):e0169925. <https://doi.org/10.1371/journal.pone.0169925>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

