

A novel epigenetic signature for predicting the prognosis of pancreatic ductal adenocarcinoma

Zhiming Zhao

Chinese PLA General Hospital

Mengyang Li

Chinese PLA General Hospital <https://orcid.org/0000-0001-8973-0445>

Xianglong Tan

Chinese PLA General Hospital

Rong Liu (✉ liurong301@126.com)

<https://orcid.org/0000-0001-5170-6474>

Primary research

Keywords: Pancreatic ductal adenocarcinoma, Epigenetics, DNA methylation, qualitative signature, prognosis

Posted Date: May 28th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-27415/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Aberrant DNA methylation is often involved in carcinogenesis. This study is designed to establish an epigenetic signature to predict overall survival (OS) of pancreatic ductal adenocarcinoma (PDAC).

Methods

DNA methylation and RNA-seq data of PDAC patients were downloaded from the Cancer Genome Atlas database, Genotype-Tissue Expression (GTEx), and International Cancer Genome Consortium (ICGC) database. Methylation-related differentially expressed genes (DEGs) were identified using an R package MethylMix. Epigenetic signature and nomogram were established by the LASSO and multivariate Cox regression analysis, respectively. In addition, a joint survival analysis of the gene expression and methylation was performed to identify potential prognostic factors for patients with PDAC.

Results

There were a total of 56 methylation-related DEGs by MethylMix criteria. After LASSO Cox regression analysis, we developed an epigenetic signature composed of five genes according to their methylation level. The signature was able to divide patients into high-risk and low-risk groups, and the OS between the high-and low-risk groups was more significantly different in both training and validation cohort. The signature is independent of clinicopathological variables and indicated better predictive power. Moreover, we developed a novel prognostic nomogram that combines risk scores with three clinicopathological factors. The joint survival analysis of gene expression and methylation revealed that 24 genes could be independent prognostic factors for OS in PDAC.

Conclusions

The qualitative signature and nomogram that predict OS at the individualized level and guide therapy for patients with PDAC.

Background

Pancreatic ductal adenocarcinoma (PDAC) is the fourth leading cause of cancer-related deaths in the United States, with the 5-year survival rates of less than 9% [1]. PDAC is a fatal disease, and there is currently no reliable screening method to detect it in premalignant or at the early stage of malignancy. Moreover, due to the asymptomatic nature, it is usually detected at advanced stage. Although great progress has been made in treatment modalities, traditional treatments, such as surgery, radiotherapy and chemotherapy, are presently far from adequate in terms of improving the survival rate of patients

with PDAC. PDAC are heterogeneous and frequently are resistant to standard treatments [2]. At present, TNM staging system is widely used to predict the prognosis of these patients, but the prediction is not sufficiently accurate. In addition, no method exists to predict the prognosis of patients with individualized pancreatic cancer. Therefore, establishing a reliable prognostic model for screening high risk patients with poor overall survival is of great clinical significance for optimizing individualized treatments and improving the managements of cancer patients.

Epigenetic alterations are considered to be important factors for the initiation and progression of PDAC and potential therapeutic targets [3, 4]. Aberrant DNA methylation in oncogenes and tumor suppressor genes affect tumor progression and are related to the survival rate of PDAC patients [5, 6]. Several schematics exist to indicated the interrelationship between PDAC genome, epigenome and signaling pathway changes [7–9]. Tumor suppressor gene hypermethylation or oncogene hypomethylation is considered to be an important mechanism of tumorigenesis [10]. Therefore, the integrated analysis of the methylome and transcriptome is crucial to the understanding of the biological processes PDAC. The present study was to develop a novel epigenetic signature via the analysis integrating multi-omic data including DNA methylation, transcriptome, and clinical outcome of patients with PDAC.

Methods

Data and preprocessing

The DNA methylation (n = 189) and clinical data (n = 179) for PDAC used in this study was downloaded from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>). To compare the mRNA expression levels using the data from the TCGA (n = 182) and Genotype-Tissue Expression (GTEx) (n = 167) databases, we used the University of California Santa Cruz (UCSC) Xena Browser as described previously [11]. In addition, we downloaded methylation data (n = 228) from International Cancer Genome Consortium (ICGC) database (<https://icgc.org/>). The TCGA cohort was used for training because it contains methylation and transcriptome data and relatively high-quality clinical records. ICGC data was used as validation cohort to verify the performance of the model. Patients with uncertain TNM stage (n = 7) and tumor grade (n = 4) were excluded, 168 PDAC patients were finally reserved for further study. The clinicopathologic characteristics of the study cohort were shown in Table 1. The current study did not require ethics approval as all study datasets were downloaded and analyzed in accordance with the corresponding data policies of previous databases.

Table 1
Clinicopathologic characteristics of pancreatic ductal adenocarcinoma patients.

Variables	TCGA cohort	ICGC cohort
	(n = 168)	(n = 228)
	N (%)	N (%)
Age (M ± SD, years)	65.0 ± 10.9	68.0 ± 10.0
Gender		
Female	75 (44.6)	102 (44.7)
Male	93 (55.4)	126 (55.3)
Tumor location		
Body/Tail	24 (44.6)	/
Head	135 (80.4)	/
Other	9 (5.4)	/
Surgery type		
Whipple	133 (79.2)	/
Distal Pancreatectomy	20 (17.9)	/
Unknown	15 (8.9)	/
Radiation therapy		
Yes	30 (17.9)	/
No	87 (51.8)	/
Unknown	51 (20.3)	/
Grade		
I	25 (14.9)	/
II-III	143 (85.1)	/
Stage		
I-II	159 (94.6)	/
III-IV	9 (5.4)	/
T stage		
T1-2	27 (16.1)	/

Variables	TCGA cohort	ICGC cohort
	(n = 168)	(n = 228)
	N (%)	N (%)
T3-4	141 (83.9)	/
N Stage		
N0	46 (27.4)	/
N1	122 (72.6)	/

Identification of methylation-related differentially expressed genes

The limma R package was used to identify differentially methylated genes (DMGs) and differentially expressed genes (DEGs). Based on cut-off value, adjusted P values < 0.05 and |fold change (FC)| > 1.5, and adjusted P values < 0.05 and |FC| > 2 were used for obtaining DMGs and DEGs, respectively. We explored the correlation between methylation data and DEG expression using Pearson's correlation analysis. Pearson coefficient between methylation level and gene expression level was calculated to find significantly negatively related genes using the MethylMix package in R [12]. Methylation-regulated DEGs were identified using the criterion of Pearson coefficient < - 0.3 with P < 0.05.

Development and validation of signature for prognosis prediction

To identify potential OS-related genes in patients with PDAC, univariate Cox regression analysis was conducted based on these gene methylation level. Genes with P < 0.05 were defined as candidate genes and were then subjected to LASSO-penalized Cox regression analysis for identifying optimal prognostic genes affecting the OS of PDAC patients. An epigenetic signature was established for predicting OS of PDAC patients by combining the prognostic coefficients of the DMGs and their methylation level.

The formula for calculating the risk scores was as follows:

$$\text{Risk score} = \sum_{i=1}^N \text{Methylation} \times \text{Coef.}$$

Coef and Methylation represented regression coefficient and methylation level of corresponding gene, respectively. Based on the median risk score value, all patients were classified into high risk score group and low risk score group. The Kaplan-Meier survival curves were generated to graphically compare the difference of OS between two groups. The predictive ability of the model was evaluated by time-dependent receiver operating characteristic (ROC) curve analysis. In order to evaluate the feasibility of the

prognostic model, we built a validation prognostic model in the ICGC cohort. The risk score for each patient was calculated with the same formula as TCGA cohort.

Construction of prognostic nomogram

To provide clinicians with a quantitative method to predict OS of PDAC patients, we established a nomogram integrating risk score and clinicopathological factors. Discrimination between the outcomes of observations and predictions using the area under the curve (AUC) of ROC curve. A calibration map was generated by comparing the nomogram prediction probability and the observation for the 1-year and 3-year OS rates.

Joint survival analysis of DMGs and DEGs

To investigate the prognostic value of methylation level and expression level of the same methylation-regulated gene, we performed a joint survival analysis to further explore prognosis-related genes in patients with PDAC.

Statistical analysis

Continuous variables were presented as mean \pm standard deviation (SD) or median as appropriate. The tdROC curves were used to assess the predictive performance of prognostic models [13]. Prognostic nomograms were constructed based on the factors of significant multivariate prognostic using “rms” package. All statistical analyses were carried out through R software (version 3.6.1, <https://www.r-project.org/>). P values lower than 0.05 were considered significant.

Results

Identification of methylation related EDGs

The flowchart of this study is demonstrated in Fig. 1. We identified 12,102 DEGs (7,447 up-regulated genes and 4,655 down-regulated genes) between 178 PDAC samples and 171 normal samples from the TCGA and GTEx databases. Then, methylation and DEGs data were integrated into the same sample by MethylMix algorithm for correlation analysis. Totally, 56 DEGs were identified as methylation-related DEGs ($|FC| > 1.5$, $P < 0.05$, $Cor < -0.3$, Fig. 2).

Construction and validation of epigenetic signature

We obtained 25 prognosis-associated genes by using univariate Cox regression analysis, which were significantly correlated to OS in PDAC. Then, we performed LASSO-penalized Cox regression, and the results demonstrated that five genes were still significant prognostic factor for PDAC. An epigenetic signature was established based on the methylation level of the five genes. The formula was as follows: $(-0.81179 \times \text{methylation level of AHNAK2}) + (-0.80006 \times \text{methylation level of ARNTL2}) + (-0.38797 \times \text{methylation level of OAS2}) + (-0.43012 \times \text{methylation level of CLEC2B}) + (0.54849 \times \text{methylation level of HIST1H2BH})$. The distribution map of methylated status of five genes is demonstrated in Fig. 3a, and the

correlation between DNA methylation and gene expression of the five genes is summarized in Fig. 3b. The distribution of risk score, survival status, and expression pattern of the 5-gene signature are shown in Fig. 4a, b, and c, respectively. We classified patients into two groups (high-risk and low-risk groups) according to the median risk score as the cut-off. Patients with high-risk group had worse OS compared to the low-risk group in PDAC ($P < 0.05$) (Fig. 4d). The AUC value of risk score was 0.776 which showed predictive ability of the signature in predicting survival risk of PDAC patients (Fig. 4e).

Validation of the signature in ICGC cohort

To validate the predictive power of the signature, risk scores were calculated with the same formula for patient in ICGC cohort. The distribution of risk score, survival status, and expression pattern of the 5-gene signature in ICGC cohort are shown in Fig. 5a, b, and c, respectively. Similarly, there were significant differences in OS among patients between low- and high-risk groups in PDAC ($P < 0.05$) (Fig. 5d). The AUC value of the signature was 0.705 which showed predictive ability of the signature in predicting PDAC patients' survival risk (Fig. 5e).

Construction of prognostic nomogram

Table 2 displays all risk factors between clinicopathological characteristics and prognosis. According to the results based on the multivariate analysis, age, N stage, surgery type, and risk score were confirmed as independent predictors for OS ($P < 0.05$). The four variables were used to create the nomograms for OS. The prognostic nomogram for 1- and 3-year OS was shown in Fig. 6a. We further constructed ROC curve to verify the accuracy of the nomogram. The AUCs of the nomogram predicting the 1- and 3-year OS were 0.748 and 0.771, respectively (Fig. 6b). The calibration curves, shown no obviously deviations from the reference line, which illustrated optimal agreement between model prediction and actual observations for 1-and 3-year OS (Fig. 6c).

Table 2
The results for screening independent prognostic factors.

Clinical characteristics	Univariate Analysis		Multivariate analysis	
	HR (95%CI)	P Value	HR (95%CI)	P Value
Age (M ± SD, years)	1.024 (1.002–1.047)	0.035	1.027 (1.014 – 1.041)	< 0.001
Gender				
Female	1			
Male	0.761 (0.493–1.174)	0.217		
Surgery type				
Whipple	1		1	
Distal Pancreatectomy	0.413 (0.246–0.696)	0.001	0.431 (0.258–0.718)	0.001
Grade				
I	1			
II-III	1.327 (0.951–1.852)	0.096		
T stage				
T1-2	1		1	
T3-4	1.902 (1.036–3.491)	0.038	1.216 (0.648–2.283)	0.543
N Stage				
N0	1		1	
N1	2.175 (1.258–3.764)	0.005	1.773 (1.009–3.115)	0.047
Risk score	3.922 (1.769–8.696)	0.001	3.878 (1.668–9.018)	0.002

Joint survival analysis

The joint survival analysis combining of methylation and expression of 24 genes was significantly associated with OS in patients with PDAC ($P < 0.05$; Fig. 6). The high-methylation and low-expression survival rate of 16 genes were poor, while the low-methylation and high-expression survival rate of the eight genes were also higher (Fig. 7).

Discussion

PDAC is a highly destructive malignant tumor of the digestive system. It is insidious, difficult to diagnose early, and prone to distant metastasis, which is an important cause of poor prognosis. Pancreatic cancer is associated with late diagnosis, early invasion, and chemoradiotherapy resistance. Therefore, a more

accurate quantitative prediction method is urgently needed to assist clinical procedures. Here, we integrate DNA methylation and RNA-sequencing data from three public databases (TCGA, GTEx, and ICGC) to develop a robust qualitative signature for OS of patients with PDAC. This study elucidated five methylated genes involved in the prognostic character of PDAC, which were used to develop an epigenetic signature. We validated its predictive power in ICGC cohort and proved that the signature is indeed significant irrespective of dataset in the independent validation cohort. The proposed risk score is independent of clinicopathological variables and showed a favorable prognostic ability. Moreover, a nomogram combining the epigenetic signature and clinicopathological factors was constructed to visually predict OS of PDAC patients. In addition, the methylation and gene expression combined survival analysis indicated that 24 independent genes can also be used as prognostic factors for stratifying OS in PDAC. Taken together, these results indicate that the signature not only serves as a biomarker of PDAC independent of clinicopathological features, but also to predicts clinical outcomes in PDAC patients.

Many genes show irregular expression due to abnormal CpG island methylation in the regulatory region of DNA, rather than sequence changes [14]. Due to the stable nature of DNA and its amplification, DNA methylation can be easily converted from laboratory settings to routine hospital operations. The methylation profile of gene promoters is different for each type of cancer, suggesting that detection of abnormal methylation can be used as a potential molecular biomarker for cancers. Moreover, epigenetic changes expected to be therapeutic targets as epigenetic changes are reversible [15]. Therefore, detecting DNA methylation can provide new insights for further assessing cancer risk and treatment. Many reports suggested that aberrant DNA methylation is involved in the development and progression of pancreatic cancer. Tumor suppressor genes often show hypermethylation of the promoter, thereby suppressing its expression and contributing to the pathogenesis of PDAC [16, 17]. Zhou et al. [17] found that methylation frequencies of CpG sites within promoter of LITAF were significantly higher in PDAC tissues, and LITAF promoter hypermethylation was associated with low LITAF expression. Moreover, promoter demethylation dose-dependently increased the LITAF transcription. LITAF demethylation inhibited proliferation and cell cycle progression, and promoted apoptosis of PDAC cells. Similarly, PCDH10 expression was significantly down-regulated in pancreatic cancer cells [18]. Methylation in PCDH10 promoter was observed in pancreatic cancer cells, and the expression of PCDH10 treated with Potent DNA methyltransferase 1 inhibitor was significantly up-regulated. Overexpression of PCDH10 inhibited the proliferation, migration, invasion ability of PDAC cells and induce apoptosis [18].

Several reports show that hypermethylation of the promoter can also serve as a diagnostic and prognostic biomarker in PDAC [19–21]. Nishizawa et al. [19] revealed that methylation level CDO1 promoter was significantly higher in PDAC tissues, and hypermethylation was significantly associated with worse disease-specific survival in PDAC. Interestingly, it was significantly higher in prospectively collected PDAC cytology samples, including both pancreatic juice and needle aspiration cytology. Promoter methylation of ADAMTS1 and BNC1 in cell-free tumor DNA were detected in early stages of pancreatic cancer, so they may be used as indicators for early diagnosis of pancreatic cancer [21]. If a person is suspected of having pancreatic cancer, his plasma sample may be collected and analyzed for his methylation status. Promoter methylation of other genes in blood can also be used as a potential

indicator for early detection of PDAC, such as MUC2 and SPARC [20, 22]. In addition, Curia et al. [6] indicated that CpG methylation was observed in PDAC tissue, and hypermethylation level of PCDH10 was significantly associated with poor progression-free survival (PFS) rates.

The proposed signature included five genes (AHNAK2, ARNTL2, OAS2, CLEC2B, and HIST1H2BH). Only AHNAK2 involved in this signature have been previously investigated in PDAC. Several studies in PDAC have confirmed AHNAK2 as an oncogene [23, 24]. Lu et al. [24] indicated that AHNAK2 is upregulated in 79 PDAC tissues compared with adjacent normal tissue, and patients with high AHNAK2 expression had worse OS. AHNAK2 protein also be detected in PDAC samples by the immunohistochemical staining [25]. Other genes in the signature have not been reported to related to PDAC biology, the functions and mechanisms of these genes in PDAC need to be further investigated.

Though the epigenetic risk score model is promising, this study inevitably has some limitations that can be explored in the future. First, the prognostic signature or nomogram were established based on the limited number of patients. A larger sample validation cohort is needed to further validate the predictive accuracy of our model. Second, although the signature of the two genes showed favorable predictive ability in PDACC, the mechanism behind it was not yet clear and further researches were needed. Further functional experiments were needed to validate the functions of the five gene signature.

Conclusion

we have identified five prognostic genes and developed an independently validated epigenetic signature to predict the OS of patients with PDAC via the analysis integrating data of DNA methylation and gene expression. The qualitative signature could robustly predict OS of PDAC patients at the individualized levels.

Availability of data and materials

The datasets used in this study are available from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>), and Genotype-Tissue Expression (GTEx) databases, and International Cancer Genome Consortium (ICGC) database (<https://icgc.org/>).

List Of Abbreviations

PDAC Pancreatic ductal adenocarcinoma

OS Overall survival

GTEx Genotype-Tissue Expression

ICGC International Cancer Genome Consortium database

DEGs Methylation-related differentially expressed genes

AUC area under the curve

ROC receiver operating characteristic

DMGs differentially methylated genes

DEGs differentially expressed genes

Declarations

Availability of data and materials

The datasets used in this study are available from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>), and Genotype-Tissue Expression (GTEx) databases, and International Cancer Genome Consortium (ICGC) database (<https://icgc.org/>).

Competing interests

The authors declare that they have no competing interests.

Declaration

Ethics approval and consent to participate

Not applicable

Consent for publication

All the authors agreed with the publication

Funding

None.

Authors' contributions

Conceptualization: LR, ZZM.

Data curation: zzm, LMY.

Formal analysis: LR, TXL, LMY.

Investigation: TXL, LMY.

Methodology: LR, ZZM.

Project administration: ZZM.

Resources: ZZM.

Software: TXL, LMY.

Supervision: ZZM, LR.

Validation: LR.

Visualization: LMY.

Writing – original draft: ZZM, LMY, TXL.

Writing – review & editing: ZZM, LMY, TXL, LR.

Acknowledgements

None.

References

1. Siegel RL, Miller KD, Jemal A. **Cancer statistics, 2019**. *CA Cancer J Clin* 2019, 69(1).
2. Matera R, Saif MW. New therapeutic directions for advanced pancreatic cancer: cell cycle inhibitors, stromal modifiers and conjugated therapies. *Expert Opin Emerg Drugs*. 2017;22(3):223–33.
3. Neureiter D, Jäger T, Ocker M, Kiesslich T. Epigenetics and pancreatic cancer: pathophysiology and novel treatment aspects. *World journal of gastroenterology*. 2014;20(24):7830–48.
4. Li Y, Go VLW, Sarkar FH. **The Role of Nutraceuticals in Pancreatic Cancer Prevention and Therapy: Targeting Cellular Signaling, MicroRNAs, and Epigenome**. *Pancreas* 2015, 44(1).
5. Thompson MJ, Rubbi L, Dawson DW, Donahue TR, Pellegrini M. Pancreatic cancer patient survival correlates with DNA methylation of pancreas development genes. *PLoS one*. 2015;10(6):e0128814.
6. Curia MC, Fantini F, Lattanzio R, Tavano F, Di Mola F, Piantelli M, Battista P, Di Sebastiano P, Cama A. High methylation levels of PCDH10 predict poor prognosis in patients with pancreatic ductal adenocarcinoma. *BMC Cancer*. 2019;19(1):452.
7. van Kampen JGM, Marijnissen-van Zanten MAJ, Simmer F, van der Graaf WTA, Ligtenberg MJL, Nagtegaal ID. Epigenetic targeting in pancreatic cancer. *Cancer treatment reviews*. 2014;40(5):656–64.
8. Paradise BD, Barham W, Fernandez-Zapico ME: **Targeting Epigenetic Aberrations in Pancreatic Cancer, a New Path to Improve Patient Outcomes?** *Cancers* 2018, 10(5).
9. Lomberk GA, Urrutia R. The Triple-Code Model for Pancreatic Cancer: Cross Talk Among Genetics, Epigenetics, and Nuclear Structure. *Surg Clin North Am*. 2015;95(5):935–52.
10. Pan Y, Liu G, Zhou F, Su B, Li Y. **DNA methylation profiles in cancer diagnosis and therapeutics**. *Clin Exp Med* 2018, 18(1).

11. Chang Y, Wang X, Xu Y, Yang L, Qian Q, Ju S, Chen Y, Chen S, Qin N, Ma Z, et al. Comprehensive characterization of cancer-testis genes in testicular germ cell tumor. *Cancer medicine*. 2019;8(7):3511–9.
12. Pan Y, Song Y, Cheng L, Xu H, Liu J. Analysis of methylation-driven genes for predicting the prognosis of patients with head and neck squamous cell carcinoma. *Journal of cellular biochemistry*. 2019;120(12):19482–95.
13. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *J Biometrics*. 2000;56(2):337–44.
14. Hervouet E, Peixoto P, Delage-Mourroux R, Boyer-Guittaut M, Cartron PF. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clin Epigenetics*. 2018;10:17.
15. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nature reviews Cancer*. 2011;11(10):726–34.
16. Hong L, Sun G, Peng L, Tu Y, Wan Z, Xiong H, Li Y, Xiao W. The interaction between miR-148a and DNMT1 suppresses cell migration and invasion by reactivating tumor suppressor genes in pancreatic cancer. *Oncol Rep*. 2018;40(5):2916–25.
17. Zhou Y, Huang J, Yu X, Jiang X, Shi Y, Weng Y, Kuai Y, Lei L, Ren G, Feng X, et al. : **is a potential tumor suppressor in pancreatic cancer**. *Oncotarget*. 2018;9(3):3131–42.
18. Qiu C, Bu X, Jiang Z. Protocadherin-10 acts as a tumor suppressor gene, and is frequently downregulated by promoter methylation in pancreatic cancer cells. *Oncol Rep*. 2016;36(1):383–9.
19. Nishizawa N, Harada H, Kumamoto Y, Kaizu T, Katoh H, Tajima H, Ushiku H, Yokoi K, Igarashi K, Fujiyama Y, et al. Diagnostic potential of hypermethylation of the cysteine dioxygenase 1 gene (CDO1) promoter DNA in pancreatic cancer. *Cancer Sci*. 2019;110(9):2846–55.
20. Zhou H, Zhu Y, Wei F, Shao Y, Pan J, Wang G, Xu K, Cheng Y. Significance of MUC2 gene methylation detection in pancreatic cancer diagnosis. *Pancreatology: official journal of the International Association of Pancreatology (IAP) [et al]*. 2019;19(8):1049–53.
21. Eissa MAL, Lerner L, Abdelfatah E, Shankar N, Canner JK, Hasan NM, Yaghoobi V, Huang B, Kerner Z, Takaesu F, et al. Promoter methylation of ADAMTS1 and BNC1 as potential biomarkers for early detection of pancreatic cancer in blood. *Clin Epigenetics*. 2019;11(1):59.
22. Singh N, Rashid S, Rashid S, Dash NR, Gupta S, Saraya A. Clinical significance of promoter methylation status of tumor suppressor genes in circulating DNA of pancreatic cancer patients. *J Cancer Res Clin Oncol*. 2020;146(4):897–907.
23. Almeida PP, Cardoso CP, de Freitas LM. PDAC-ANN: an artificial neural network to predict pancreatic ductal adenocarcinoma based on gene expression. *BMC Cancer*. 2020;20(1):82.
24. Lu D, Wang J, Shi X, Yue B, Hao J. AHNK2 is a potential prognostic biomarker in patients with PDAC. *Oncotarget*. 2017;8(19):31775–84.
25. Klett H, Fuellgraf H, Levit-Zerdoun E, Hussung S, Kowar S, Küsters S, Bronsert P, Werner M, Wittel U, Fritsch R, et al. Identification and Validation of a Diagnostic and Prognostic Multi-Gene Biomarker Panel for Pancreatic Ductal Adenocarcinoma. *Frontiers in genetics*. 2018;9:108.

Figures

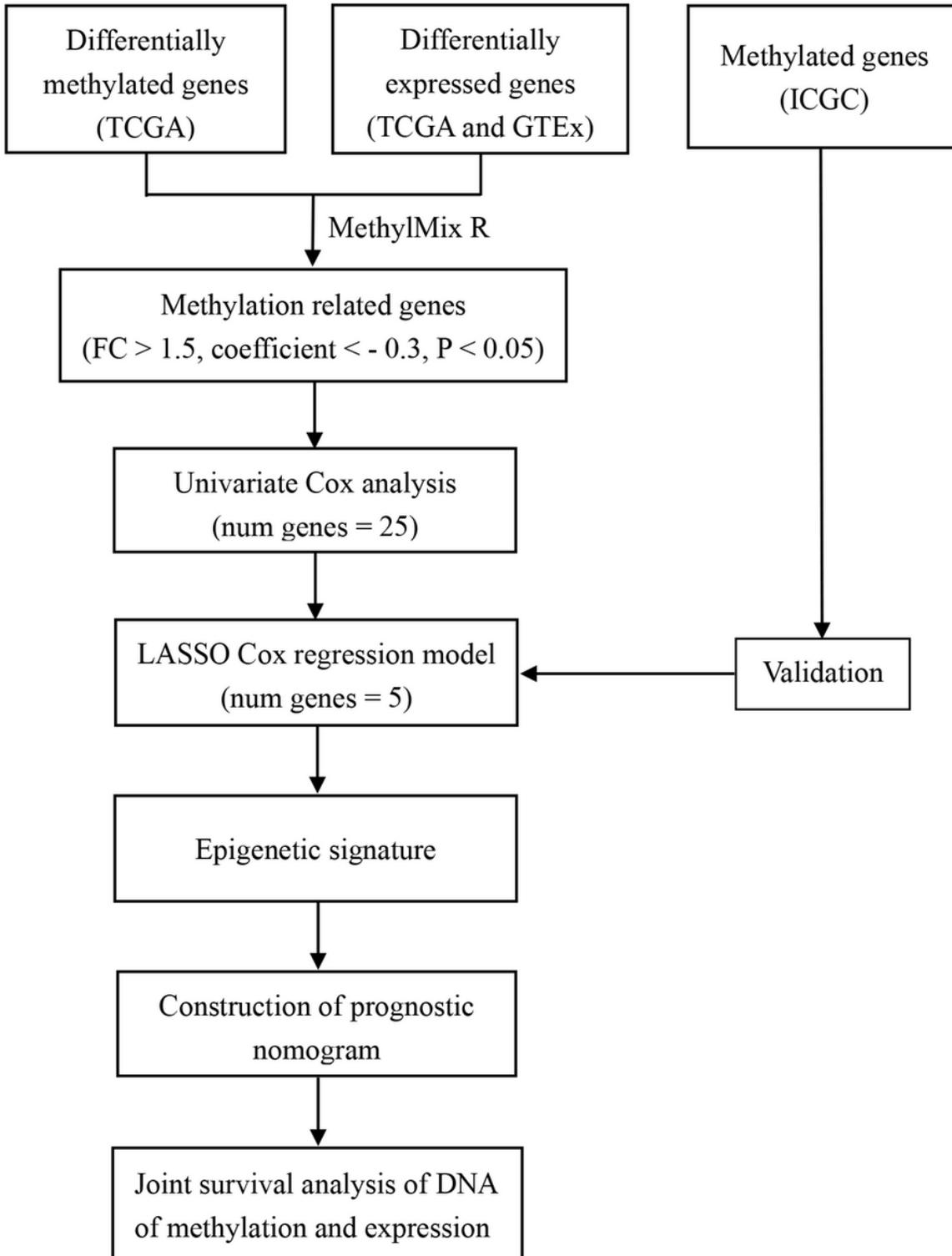


Figure 1

Flowchart of construction and validation of the signature.

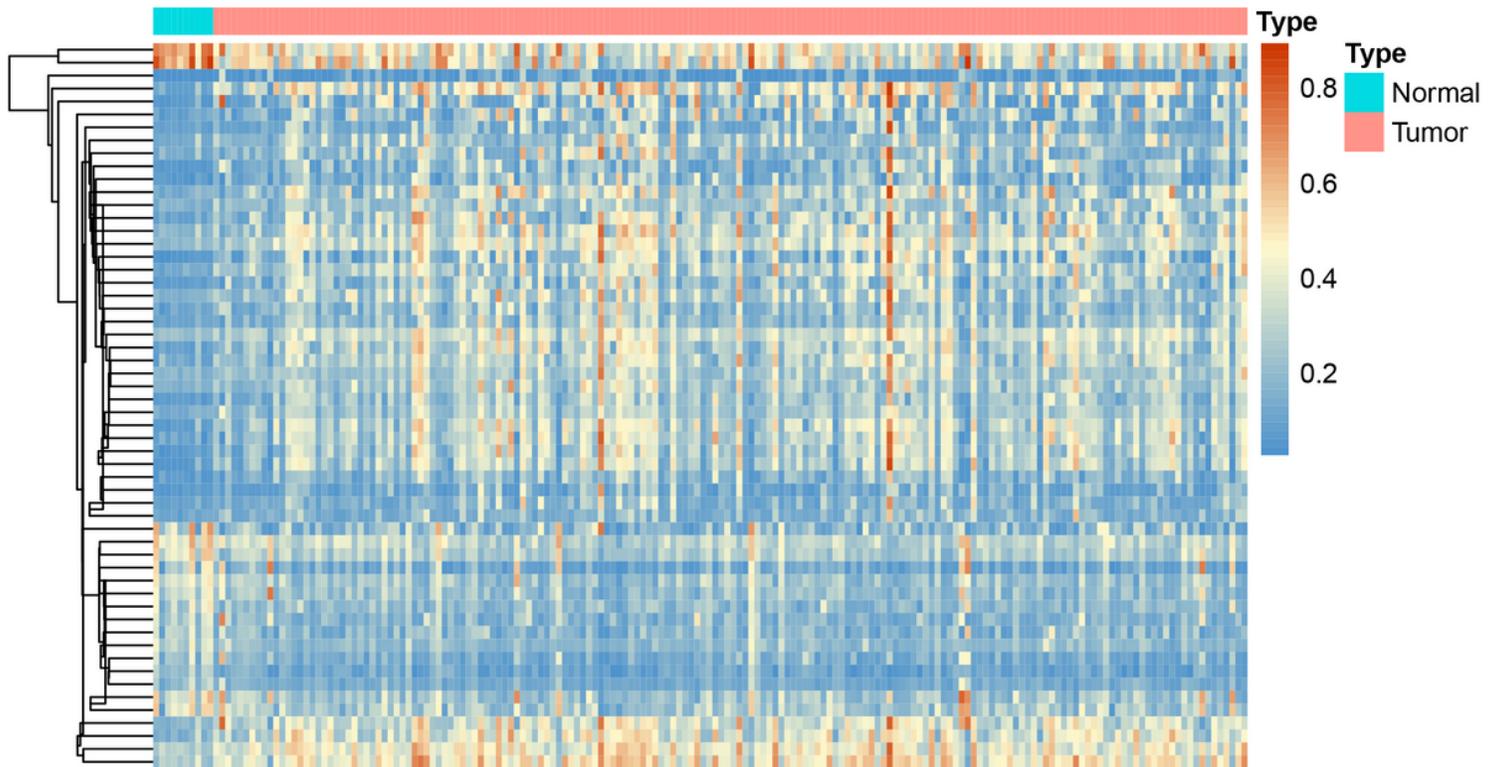
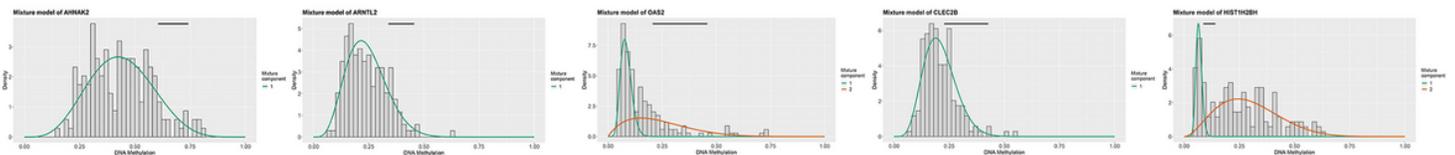


Figure 2

Heatmap of methylation related differentially expressed genes in pancreatic ductal adenocarcinoma). The color change from blue to red in the heatmap illustrates the trend from low to high methylation.

a



b

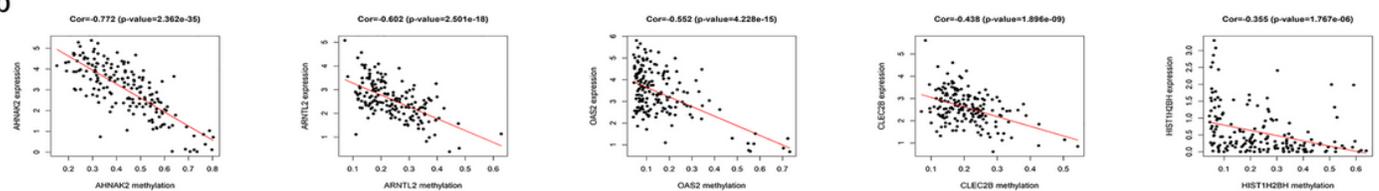


Figure 3

Summary of five methylated related genes in the signature. a The distribution map of methylated status of five genes. The histogram demonstrates the distribution of methylation in tumor samples. Horizontal black bars show the distribution of methylation in normal samples. b The correlation between methylation and gene expression in five genes.

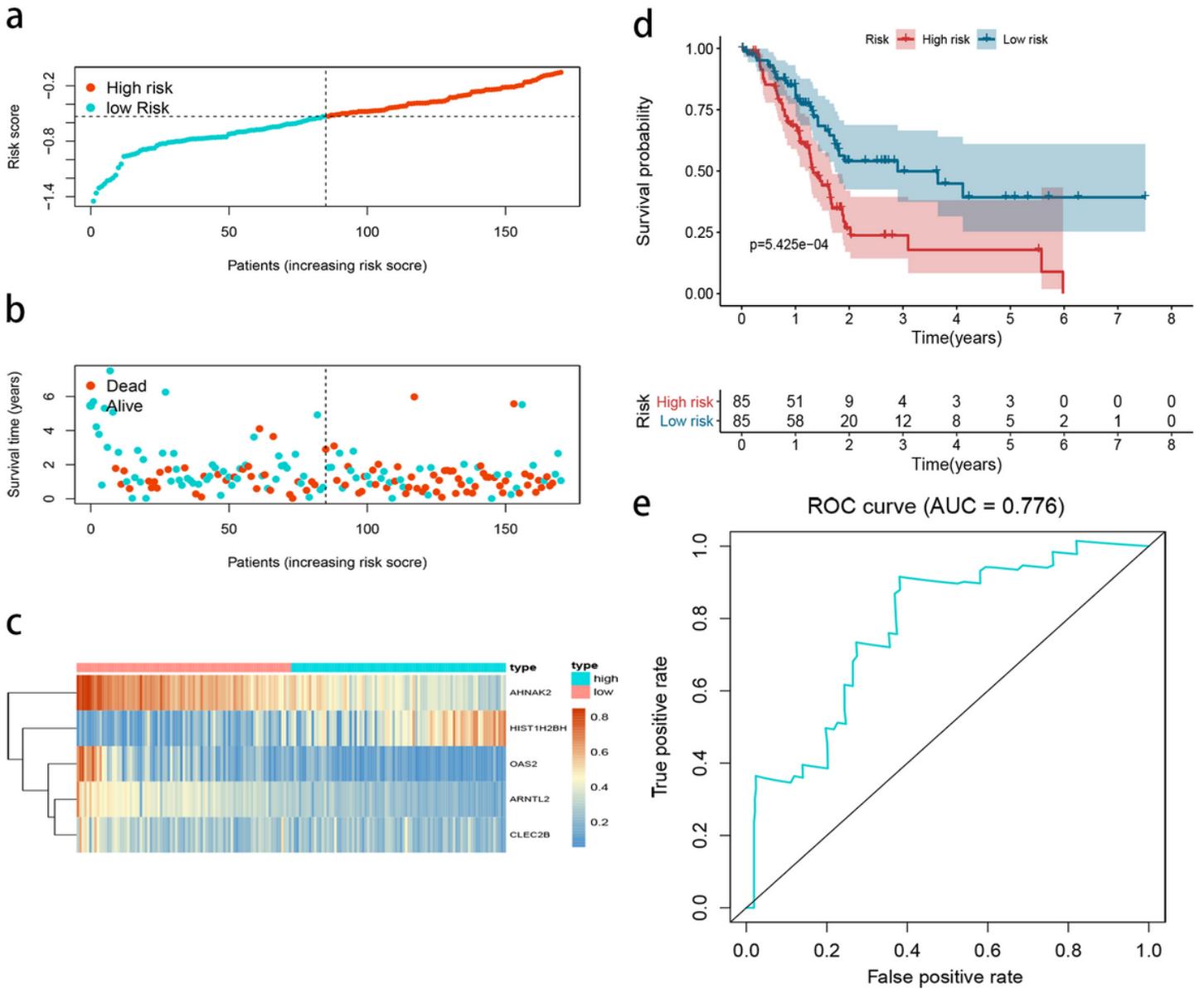


Figure 4

The epigenetic signature for overall survival prediction of pancreatic ductal adenocarcinoma patients in TCGA cohort. a The distribution of risk score derived from the signature. b The distribution of survival status of each patient. c The methylation level profile of the signature in the high-risk and low-risk subgroups. d Kaplan-Meier estimate of the OS. Patients were divided into high-risk and low-risk subgroup based on the median of risk score. The difference between the two curves was determined by the two-side log-rank test. e Time-dependent ROC curve of the signature.

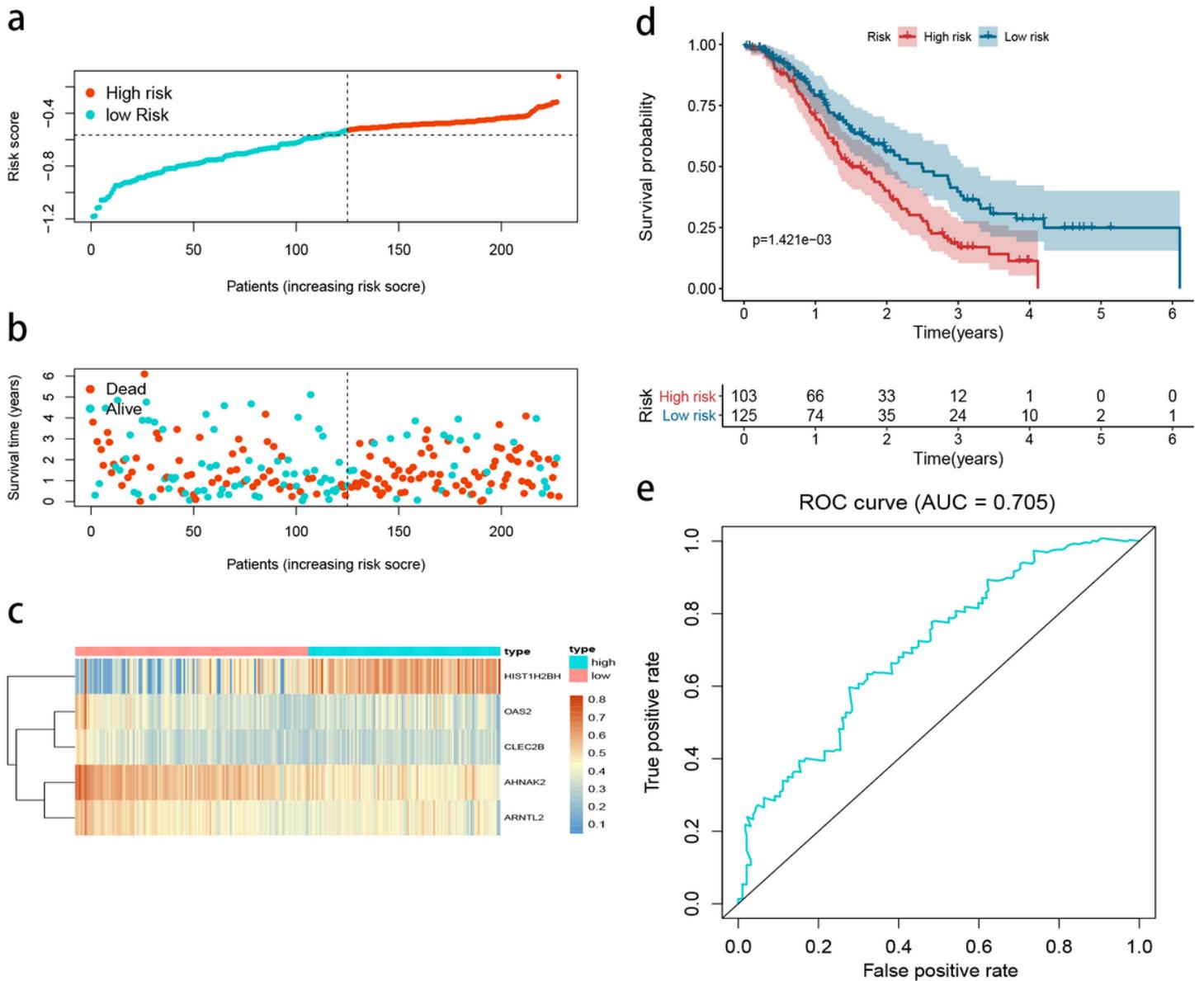


Figure 5

The epigenetic signature for overall survival prediction of pancreatic ductal adenocarcinoma patients in ICGC cohort. a The distribution of risk score derived from the signature. b The distribution of survival status of each patient. c The methylation level profile of the signature in the high-risk and low-risk subgroups. d Kaplan-Meier estimate of the OS. Patients were divided into high-risk and low-risk subgroup based on the median of risk score. The difference between the two curves was determined by the two-side log-rank test. e Time-dependent ROC curve of the signature.

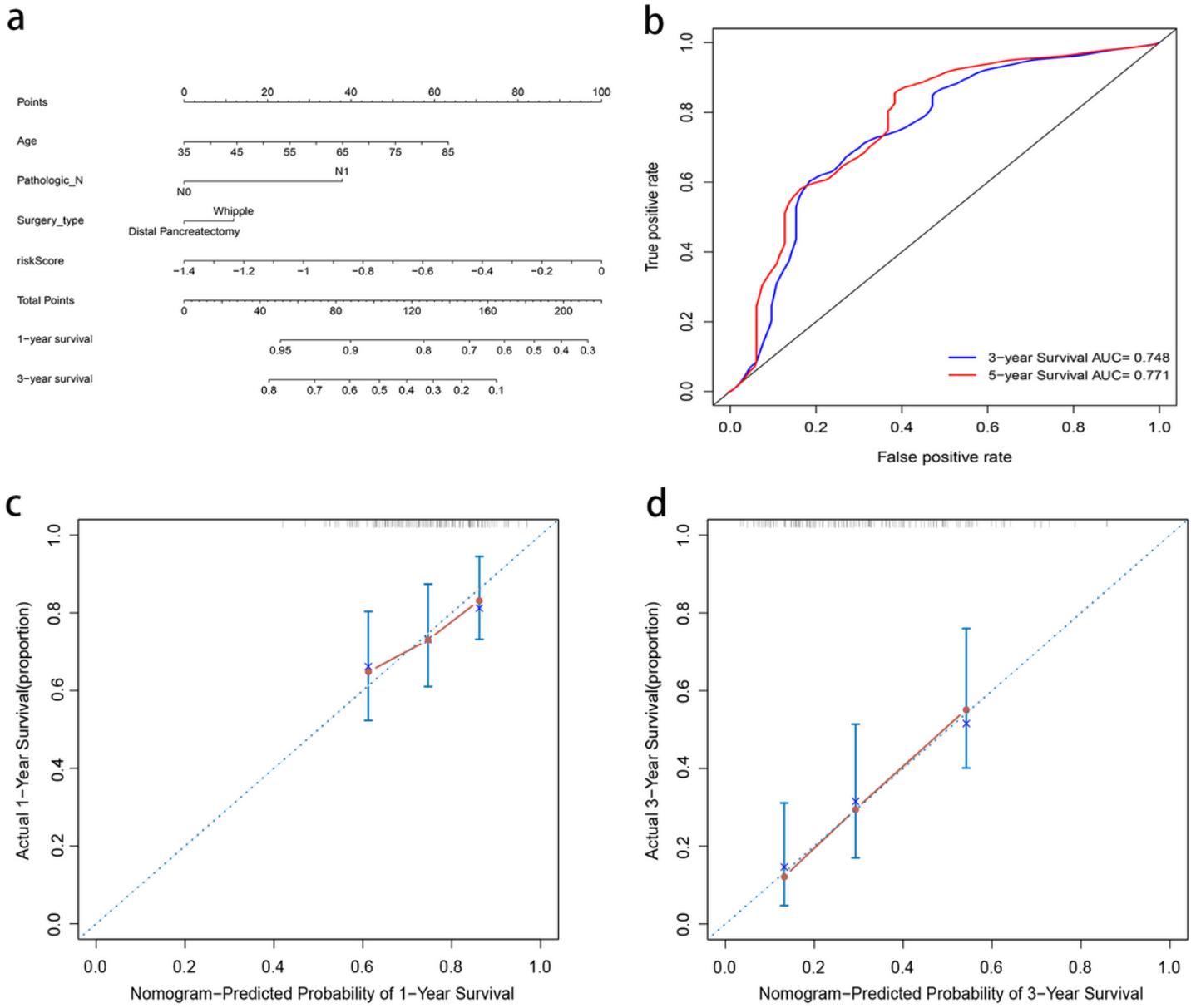


Figure 6

Nomogram for predicting overall survival in pancreatic ductal adenocarcinoma. a Nomogram including the epigenetic signature and clinicopathological traits. b AUC values of ROC predicted 1-year and 3-year overall survival rates of nomogram. c calibration curves for the probability of overall survival of 1- and 3-year.

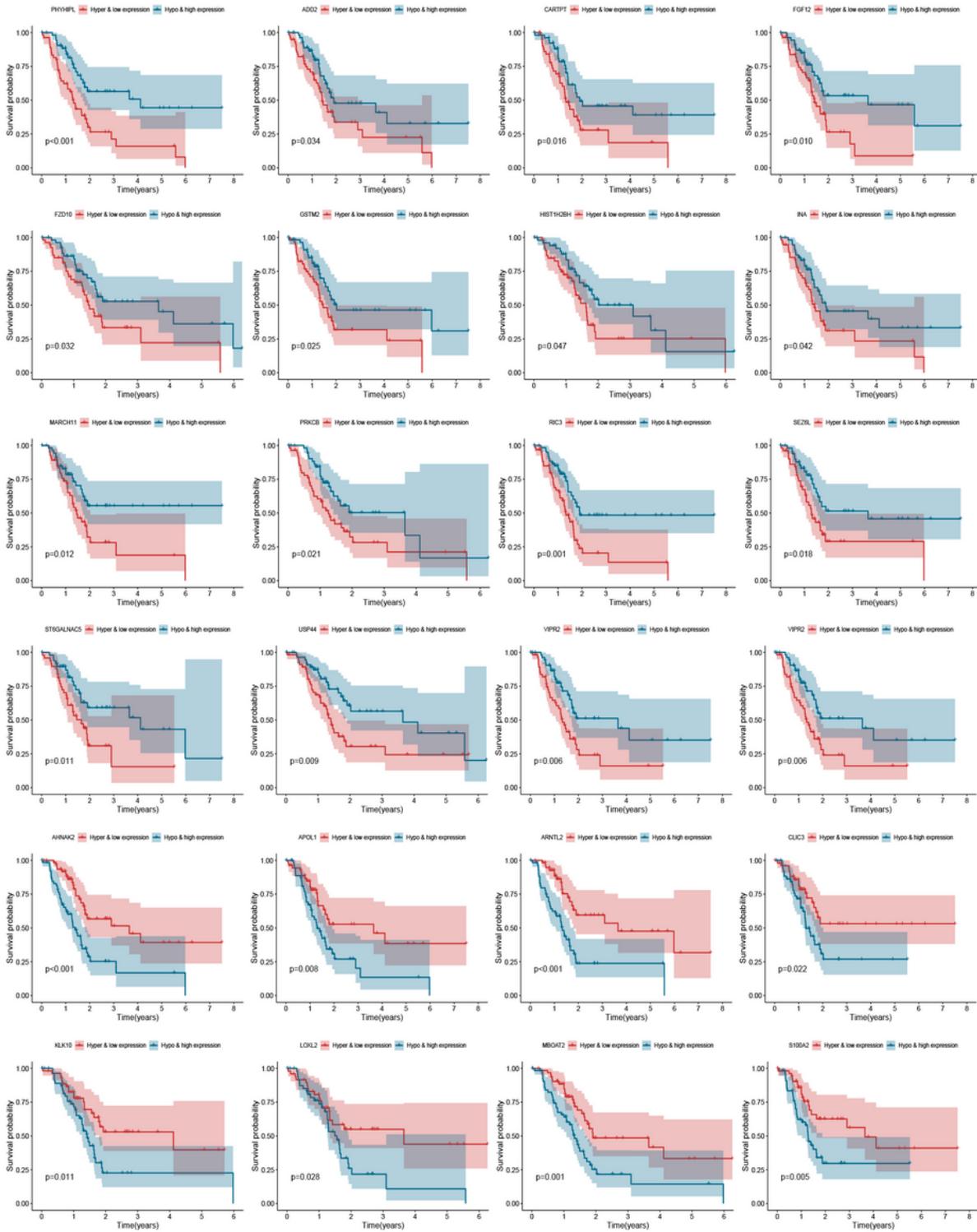


Figure 7

Kaplan-Meier survival curves of joint analysis of methylation and expression data in pancreatic ductal adenocarcinoma.