

ORIGINAL RESEARCH ARTICLE

Systemic drug repurposing for pancreatic cancer based on genetic and epigenetic network analysis via a systems biology approach and deep neural learning of drug-target interactions

Supplementary File

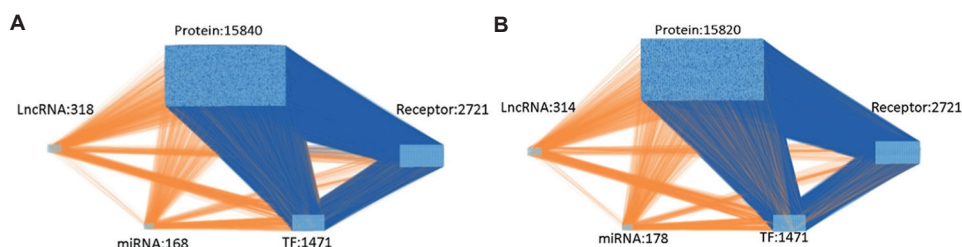


Figure S1. Real genome-wide genetic and epigenetic networks of pancreatic ductal adenocarcinoma and healthy control samples. (A) Real genome-wide genetic and epigenetic network for pancreatic ductal adenocarcinoma. (B) Real genome-wide genetic and epigenetic network for healthy controls. The numbers represent the amount of molecules in real GWGENs. Blue lines indicate protein-protein interactions, while orange lines denote regulatory interactions between genes. Statistical data for the nodes and edges in real genome-wide genetic and epigenetic networks are provided in [Tables S2](#) and [S3](#)

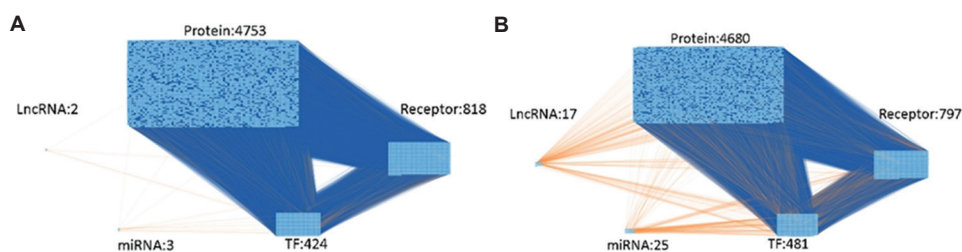


Figure S2. Core genome-wide genetic and epigenetic networks of pancreatic ductal adenocarcinoma and healthy control samples by the principal network projection method. (A) Core genome-wide genetic and epigenetic networks for pancreatic ductal adenocarcinoma. (B) Core genome-wide genetic and epigenetic networks for healthy controls. The numbers represent the amount of molecules in core genome-wide genetic and epigenetic networks. Blue lines indicate protein-protein interactions, while orange lines denote regulations between genes



Figure S3. The accuracy of training and validation during five-fold cross-validation

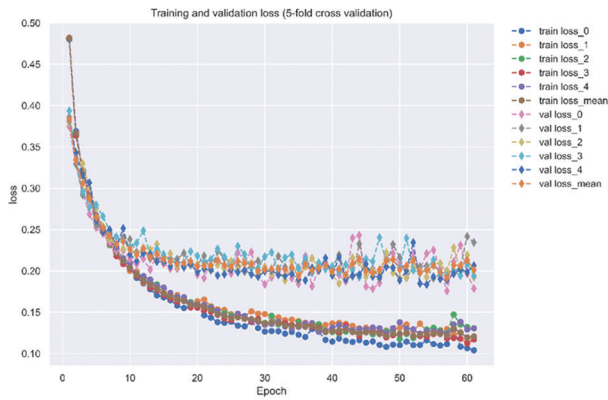


Figure S4. The loss of training and validation during five-fold cross-validation

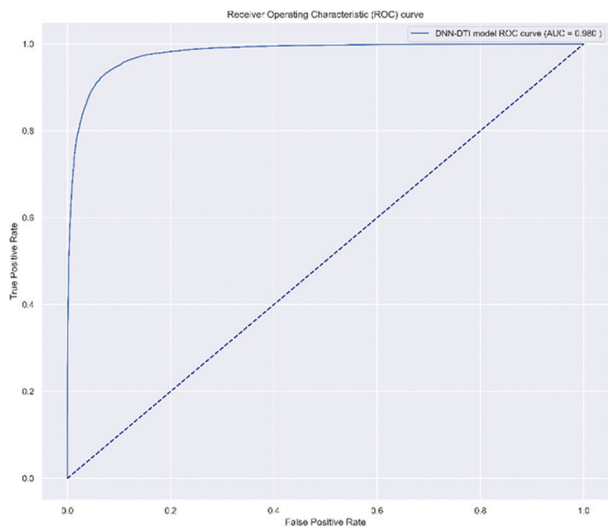


Figure S5. Receiver operating characteristic curve of the trained deep neural network-based drug-target interaction model. The area under the curve score is 0.980

Table S1. Statistical data for the predictive performance of the deep neural network-based drug-target interaction model, evaluated using five-fold cross-validation

Fold	Validation loss	Validation accuracy	Test loss	Test accuracy
1	0.17864	0.935993	0.198111	0.931188
2	0.234376	0.924918	0.191752	0.927867
3	0.19229	0.930341	0.191241	0.931147
4	0.194115	0.931186	0.256386	0.933015
5	0.206544	0.920866	0.197647	0.928905
Average	0.201193	0.928661	0.207027	0.930424
Standard deviation	0.018803	0.005248	0.024845	0.001826

Table S2. Statistical data for nodes in the real genome-wide genetic and epigenetic networks of pancreatic ductal adenocarcinoma and healthy controls

Nodes	Candidate GWGENs	Real PDAC	Real healthy
Rcp	2,721	2,721	2,721
TF	1,478	1,471	1,471
Proteins	18,012	15,840	15,820
miRNAs	203	168	178
LncRNAs	890	318	314
Total nodes (genes)	23,304	20,518	20,504

Abbreviations: GWGENs: Genome-wide genetic and epigenetic networks; LncRNAs: Long non-coding RNAs; miRNAs: MicroRNAs; PDAC: Pancreatic ductal adenocarcinoma; Rcp: Receptor; TF: Transcription factor.

Table S3. Statistical data table for edges in the real genome-wide genetic and epigenetic networks of pancreatic ductal adenocarcinoma and healthy controls

Edges	Candidate GWGENs	Real PDAC	Real healthy
PPIs	4,759,187	2,099,099	1,750,128
TF - Receptor	15,608	7,033	7,013
TF - TF	12,105	4,992	4,995
TF - Protein	84,709	38,269	39,524
TF - miRNA	1,275	428	442
TF - lncRNA	490	326	348
miRNA - Receptor	14,220	56	59
miRNA - TF	12,403	63	72
miRNA - Protein	83,098	564	679
miRNA - miRNA	12	6	6
miRNA - lncRNA	380	18	19
lncRNA - Receptor	222	110	100
lncRNA - TF	252	97	105
lncRNA - Protein	1,732	864	861
lncRNA - miRNA	0	0	0
lncRNA - lncRNA	6	5	5
Total edges	4,985,699	2,151,930	1,804,356

Abbreviations: GWGENs: Genome-wide genetic and epigenetic networks; LncRNA: Long non-coding RNA; miRNAs: MicroRNA; PDAC: Pancreatic ductal adenocarcinoma; PPI: Protein-protein interaction; TF: Transcription factor.