Integration of Protein-Protein Interaction Networks and Gene Expression Profiles Helps Detect Pancreatic Adenocarcinoma Candidate Genes

Lili Su[1†], Guang Liu[1†], Ying Guo[2†], Xuanping Zhang[1], Xiaoyan Zhu[1] and Jiayin Wang[1]*

[1]College of Electronics and Information Engineering, School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, China, [2]Department of Histology and Embryology, College of Basic Medical Sciences, Jilin University, Changchun, China

More and more cancer-associated genes (CAGs) are being identified with the development of biological mechanism research. Integrative analysis of protein-protein interaction (PPI) networks and co-expression patterns of these genes can help identify new disease-associated genes and clarify their importance in specific diseases. This study proposed a PPI network and co-expression integration analysis model (PRNet) to integrate PPI networks and gene co-expression patterns to identify potential risk causative genes for pancreatic adenocarcinoma (PAAD). We scored the importance of the candidate genes by constructing a high-confidence co-expression-based edge-weighted PPI network, extracting protein regulatory sub-networks by random walk algorithm, constructing disease-specific networks based on known CAGs, and scoring the genes of the sub-networks with the PageRank algorithm. The results showed that our screened top-ranked genes were more critical in tumours relative to the known CAGs list and significantly differentiated the overall survival of PAAD patients. These results suggest that the PRNet method of ranking cancer-associated genes can identify new disease-associated genes and is more informative than the original CAGs list, which can help investigators to screen potential biomarkers for validation and molecular mechanism exploration.

Keywords: cancer-associated genes, co-expression, protein-protein interaction, pancreatic adenocarcinoma, machine learning

INTRODUCTION

Abnormal gene regulation and uncontrolled growth of cells in pancreatic tissue can lead to pancreatic cancer (McGuigan et al., 2018; Yao et al., 2020). Pancreatic adenocarcinoma (PAAD) accounts for approximately 80% or more of all types of pancreatic cancer and is the most common type of pancreatic cancer. This tumour is one of the most common cancers and is the top 10 leading causes of cancer-related death (Siegel et al., 2020). Therefore, it is critical to screen and identify potential biomarkers that can be used to treat PAAD, which will significantly improve the quality of life and prognosis of PAAD patients.

Tumour tissue is highly heterogeneous, and its complexity is a significant obstacle to a comprehensive understanding of the molecular mechanisms underlying tumour development (Qian et al., 2020). Researchers have uncovered potential risk-causing genes in tumour

progression for many years (Bilici 2014; Le et al., 2016; Loosen et al., 2017). With advances in experimental tools and high-throughput sequencing technologies, an increasing number of genes are closely associated with cancer progression. The Cancer Gene Census (CGC) database (Sondka et al., 2018) screens candidate genes by searching for the presence of typical oncogene somatic mutation patterns, and determines the biological functions of candidate genes through literature curation, and has identified mutations affecting the molecular mechanism of gene dysfunction, which in turn explains the rationale for oncogenic transformation of the gene. The latest gene-wide screening described 719 genes that play essential functions in pan-cancer, providing researchers with great help in understanding cancer pathogenesis. However, whether these genes function in pancreatic adenocarcinoma tumour tissue and whether there are new pancreatic adenocarcinoma-specific biomarkers remain to be further investigated.

The development of computational methods for tumour candidate biomarker identification and ranking could fill the gap mentioned above. Alshahrani and Hoehndorf (2018) incorporated phenotypic similarity into network-based characterization learning and conveyed phenotypic association information through PPI networks to address the incompleteness of gene-phenotype knowledge. The features generated by their proposed SmuDGE algorithm can characterize gene-disease associations. By integrating information from various aspects such as PubMed abstracts, pathways, interactions, gene ontology, disease ontology, sequence similarity, and constructing bayesian ridge regression models, the pBRIT developed by Kumar et al. (2018) can prioritize disease genes. Recently, Li et al. (2019) constructed a graph convolutional neural network-based gene ranking algorithm PGCN by training both embedded learning models and association prediction models in an end-to-end manner and discovered some disease-associated candidate genes. These approaches have achieved a great success in exploring potential disease-associated genes, which also suggest that integrating multifaceted biological information can help us better delineate the gene-disease associations and thus screen key candidate genes. The PageRank algorithm is one of the most widely used page ranking algorithms, it is developed by Google and is named after Larry Page, Google's co-founder and president (Brin et al., 1998). When applied to biological networks, PageRank has a great stability can help evaluate important nodes and pathways in directed networks, such as metabolic networks (Iván and Grolmusz 2011).

With the development of high-throughput sequencing technology, tumour-related sequencing profiles are increasingly accumulated. Bioinformatics analysis based on expression profiling data has emerged in studying the molecular mechanisms of tumorigenesis and development. This study proposed a PPI network and co-expression integration analysis model (PRNet) to integrate PPI network and gene co-expression patterns to identify potential risk causative genes in pancreatic adenocarcinoma (PAAD). We scored the importance of candidate genes by PRNet to screen new key regulators based on existing cancer-related genes and PageRank algorithm is used for network scoring in PRNet. Moreover, the reliability of our

ranked candidate genes was further validated by various analytical tools. Overall, our data can give a guide to the study of the biological mechanisms of PAAD development and provide new insights for the future treatment of PAAD.

MATERIALS AND METHODS

Dataset Preparation

TCGA PAAD transcriptome normalization data and clinical information were downloaded from UCSC Xena (https://xenabrowser.net/datapages/) for further co-expression analysis. The list of cancer-associated genes was obtained from the Cancer Gene Census (https://cancer.sanger.ac.uk/census). CERES-dependent scores were based on data from the Crispr technology to knockdown target genes and perform cell depletion assays (Ghandi et al., 2019). Target gene dependency scores were obtained from the DepMap web portal (https://depmap.org/portal/download/). A lower CERES score for a given gene in a given cell line indicates that the gene of interest plays an essential function in that cell line. A score of 0 indicates that the gene is not essential; accordingly, a score of -1 is the median of all pan-essential gene function scores. The experimentally validated protein interaction information was obtained by integrating BioGRID (Merriel et al., 2011), I2D (Brown and Jurisica 2007), BioPlex (Huttlin et al., 2017) and IntAct (Hermjakob et al., 2004).

Weighting Protein-Protein Interactions

Based on the TCGA PAAD transcriptome data, we can analyze the co-expression relationship between genes. The stronger co-expression relationship indicates that the two relationships are more functionally related. We first calculated the correlations of all possible gene pairs and filtered them based on the PPI network, keeping only the gene pairs with PPI present. Thus, the weight between each gene pair can be expressed as.

$$W_{(A,B)} = \begin{cases} corr\,(A,\,B), & Pair\,(A,\,B) = 1 \\ NA, & Pair\,(A,\,B) = 0 \end{cases}$$

Where $W_{(A,\,B)}$ denotes the weight of the linkage between gene A and gene B. $Pair\,(A,\,B) = 1$ indicates the presence of interaction between gene A and gene B, while the opposite does not exist. We only keep the gene pairs with interactions to calculate their weights. $corr\,(A,\,B)$ indicates the correlation between the expression levels of gene A and gene B.

Generating Disease-Specific Networks

After weighting the PPI network, we use the short random walks algorithm (Barnes and Feige, 1993) to decompose the network into a series of sub-networks. This step is implemented based on the cluster_walktrap function of the R package igraph. The core idea is that short random walks are more likely to be enriched in the same community. Therefore, we filtered the subnetworks based on the collected CAGs and only retained subnetworks containing at least one CAG for subsequent analysis. This means that PRNet filters sub-networks containing at least one CAG

instead of the entire weighted PPT network to construct disease-specific networks, which can better retain disease-specific information.

Ranking Candidate Genes

PRNet uses the PageRank algorithm (Xing and Ghorbani, 2004) to calculate all the final filtered sub-networks, which returns the PageRank value of each node, which is considered the importance level of the gene. The specific algorithm is as follows.

$$PR\left(gene_i\right) = \frac{1-q}{N} + q \sum_{gene_j}^{k} \left(\frac{PR\left(gene_j\right)}{L\left(gene_j\right)} \right)$$

Where $gene_i$ is the genes to be studied, and $gene_j$ stand for the genes interacting with $gene_i$. k is the number of genes interacting with the gene to be studied, $PR\left(gene_i\right)$ is the PageRank value of the gene, $L\left(gene_j\right)$ is the number of genes chained from $gene_j$. The 'N' means the total amount of genes used for analysis, and the 'q' is the damping factor, the meaning of which is the probability of reaching a gene and continues to navigate backwards. The default value of 'q' is 0.85.

Based on this, we can calculate the PageRank value of each gene and use it to indicate the relative importance of the gene in the disease-specific network and then potential screen biomarkers.
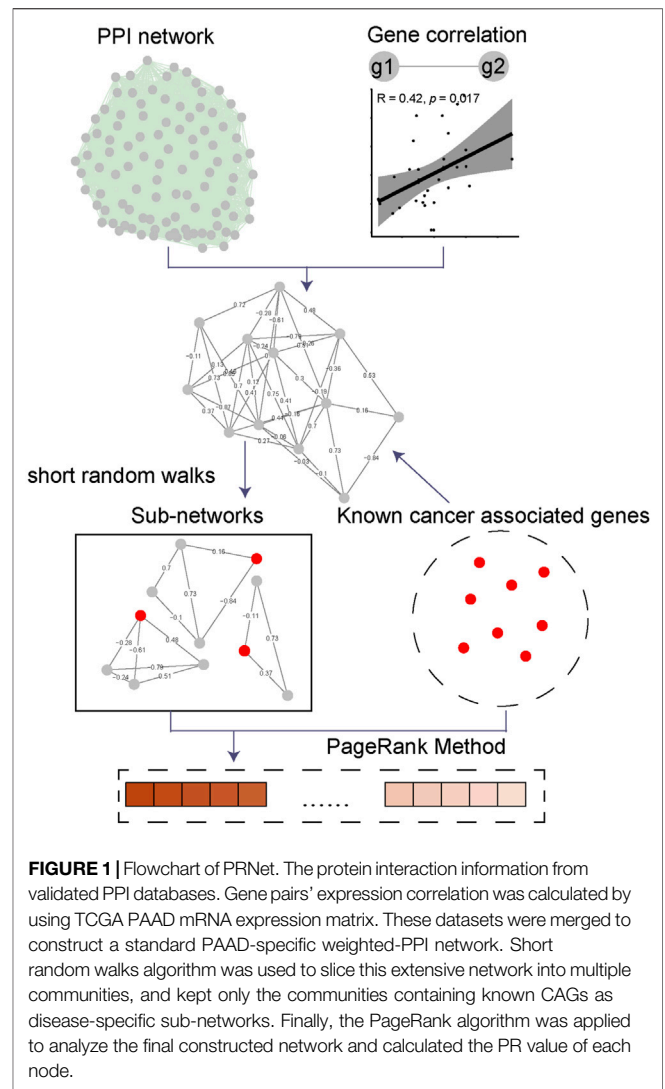
Bioinformatics Analysis

Based on the TCGA PAAD gene expression normalized count matrix, we divided the samples into two subgroups by unsupervised hierarchical clustering using the R package "ConsensusClusterPlus" (Wilkerson and Hayes 2010). Then the matrix was processed as log2 (FPKM +1) and genes with zero expression in more than half of the samples were removed. The DESeq2 package (Love et al., 2014) was used to calculate the differential expression of all genes between the different groups, where adjusted p-values less than 0.05 and | log2FoldChange|>1 were considered differentially expressed genes. Principal component analysis (PCA) (Wold et al., 1987) revealed differences in expression patterns between different subgroups of PAAD. Gene set enrichment analysis (GSEA) (Shi and Walker 2007) was used to perform enrichment analysis of differential expression data. False discovery rate (FDR) q-value <0.05 was set as the cut-off criterion. Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway, we performed gene set enrichment analysis using the R package "clusterProfiler" v3.18.1. The enzyme, transcription factor, FDA approved drugs information were downloaded from the Human Protein Atlas (HPA) database. Survival analysis was performed using the R package "survival," and the overall survival curve was obtained by Kaplan-Meier estimation.
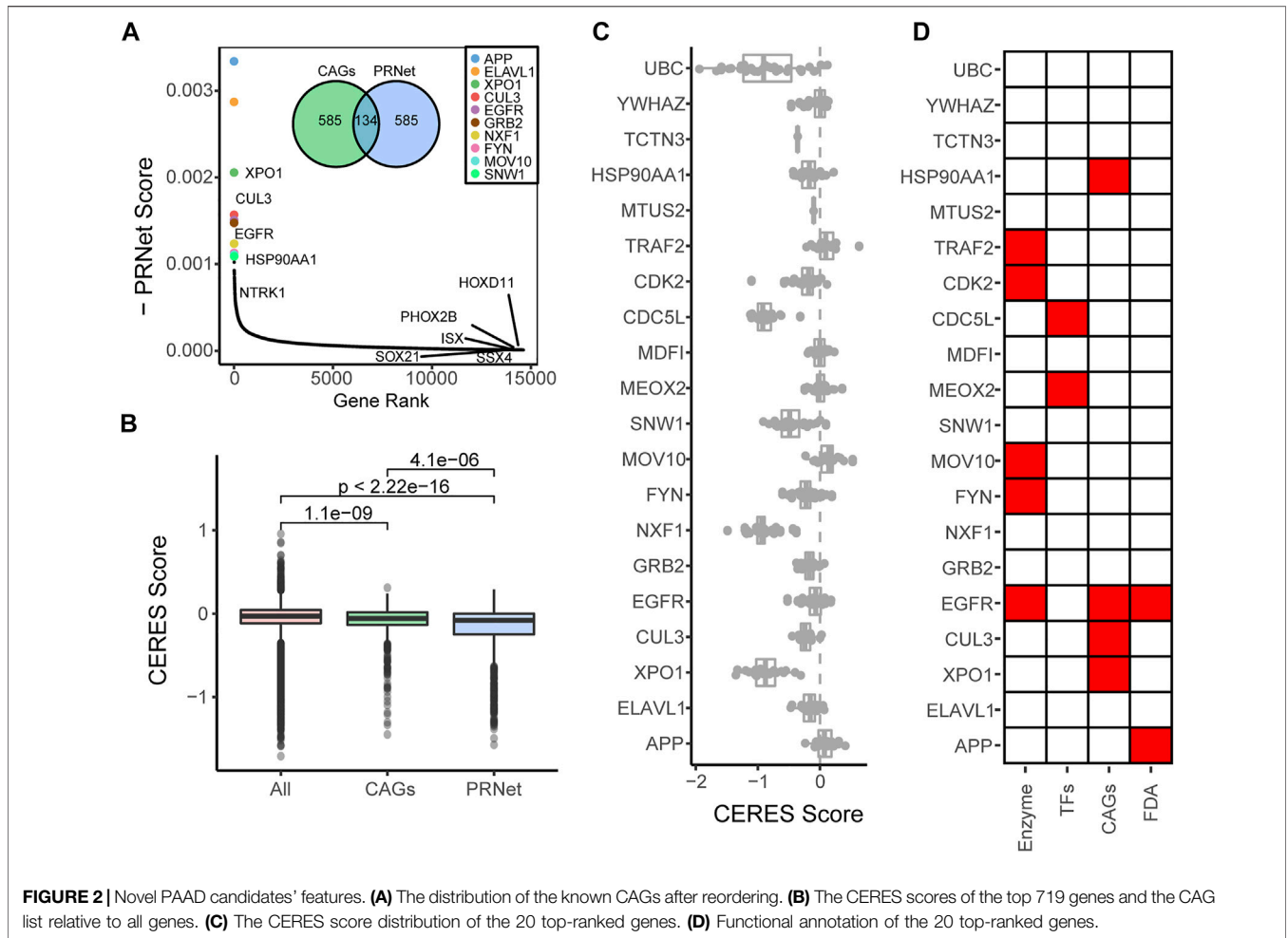
RESULTS

Prioritizing Disease Candidate Genes

The basic flow of model construction is shown in **Figure 1**. We curated the PPI information from several databases including

FIGURE 1 | Flowchart of PRNet. The protein interaction information from validated PPI databases. Gene pairs' expression correlation was calculated by using TCGA PAAD mRNA expression matrix. These datasets were merged to construct a standard PAAD-specific weighted-PPI network. Short random walks algorithm was used to slice this extensive network into multiple communities, and kept only the communities containing known CAGs as disease-specific sub-networks. Finally, the PageRank algorithm was applied to analyze the final constructed network and calculated the PR value of each node.

204,399 interactions from BioGRID, 16,334 interactions from I2D (Brown and Jurisica 2007), 28,382 interactions from BioPlex (Huttlin et al., 2017) and 109,220 interactions from IntAct (Hermjakob et al., 2004). These protein interactions collected were all experimentally validated with high confidence, so we directly merged the results from all databases. A total of 358,335 interactions among 20,379 proteins were finally collected, and after removing duplicate protein interaction pairs, 309,321 experimentally identified interactions among 20,379 proteins were obtained for further analysis. Next, we downloaded TCGA PAAD transcriptome data, calculated gene pairs' expression correlation in the human-derived PPI network. We merged the correlation coefficients into the PPI network as the weights of gene pairs to construct a standard PAAD-specific weighted-PPI network. Next, we analyzed this network using short random walks, sliced this extensive network into multiple communities, and kept only the communities containing known CAGs as disease-specific sub-networks. A total of 147 communities were generated as a result and only 19

FIGURE 2 | Novel PAAD candidates' features. **(A)** The distribution of the known CAGs after reordering. **(B)** The CERES scores of the top 719 genes and the CAG list relative to all genes. **(C)** The CERES score distribution of the 20 top-ranked genes. **(D)** Functional annotation of the 20 top-ranked genes.

communities contain the known CAGs were obtained for further analysis. We used the PageRank algorithm to analyze the final constructed network and calculated the PR value of each node. Higher PR values indicate that the gene plays a more regulatory role in PAAD, i.e., more important for PAAD development. Lower PR values indicate that the gene is more isolated in the gene regulatory network of PAAD and less likely to act as a candidate in PAAD.

Novel PAAD Candidate Genes

Based on the above method, we calculated the PR values of 14,615 genes in PAAD, and the ranking and PR values of these genes were shown in **Supplementary Table S1**. We selected three subnetworks containing the most genes and performed KEGG enrichment analysis (**Supplementary Figure S1** and **Supplementary Table S2**). The largest subnetwork contains many pathways related to cell proliferation and division, and maintenance of life activities such as cell cycle, DNA replication, ubiquitin-mediated proteolysis, etc. The pathways involved in the second subnetwork are related to malignant tumour progression, such as MAPK signaling pathway, ERBB signaling pathway, VEGF signaling pathway, etc. The third major sub-network is related to ribosomes.

We further compared the distribution of the known CAGs after reordering. As shown in **Figure 2A**, there are XX identical genes in the top 719 genes and the CAG list, XPO1, CUL3, EGFR, HSP90AA1, NTRK1, etc., were clearly ranked at the top, further indicating their essential functions in the malignant progression PAAD. However, we also found that some CAGs such as HOXD11, PHOX2B, SSX4, ISX, SOX21 ranked significantly lower. Our selected CAG list is all about pan-cancer markers, but not all markers play essential functions in PAAD. To further confirm the reliability of our screening results, we analyzed the CERES scores of the top 719 genes and the CAG list relative to all genes (**Figure 2B**). The results showed that the CERES scores of PRNet screened genes were significantly lower than those of the known CAG list, indicating that these genes play more critical functions in PAAD. Then, we showed the top 20 genes, plotted their CERES score distribution individually, and annotated their functions (**Figures 2C,D**). Some genes are known as enzymes, transcription factors or tumour therapeutic targets. We checked the genetic alterations of the top 20 genes in TCGA PAAD cohorts and less mutations were found, in addition to the fact that most of the genes had samples that showed amplification in samples, especially YWHAZ, with a high frequency (7%) (**Supplementary Figure S2A**). Further functional enrichment
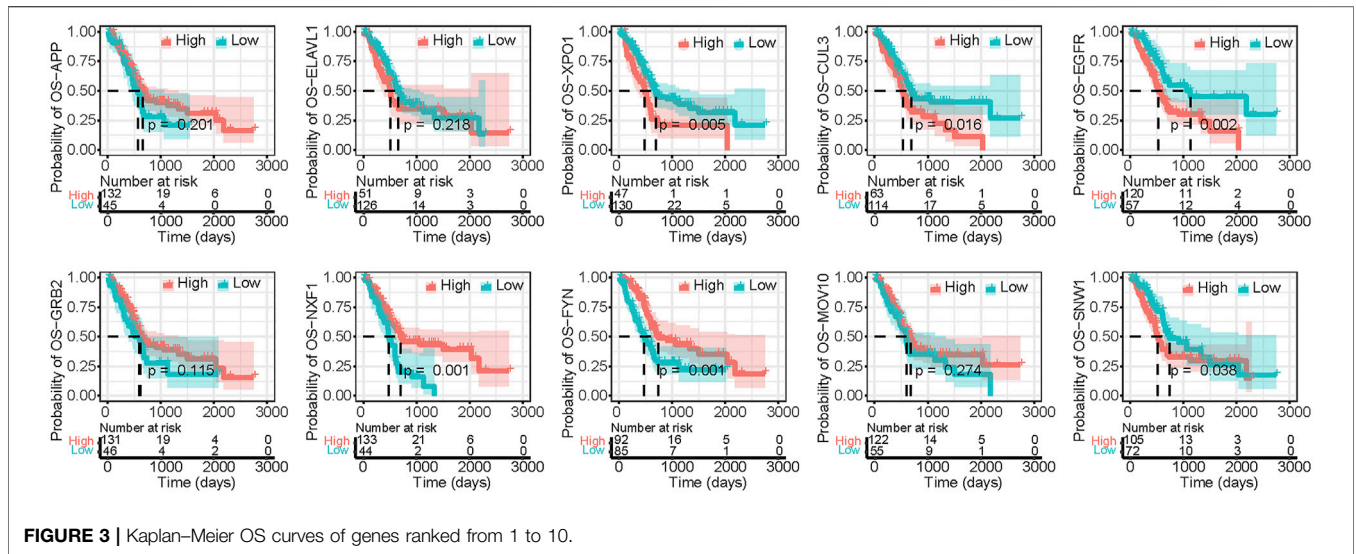
FIGURE 3 | Kaplan–Meier OS curves of genes ranked from 1 to 10.
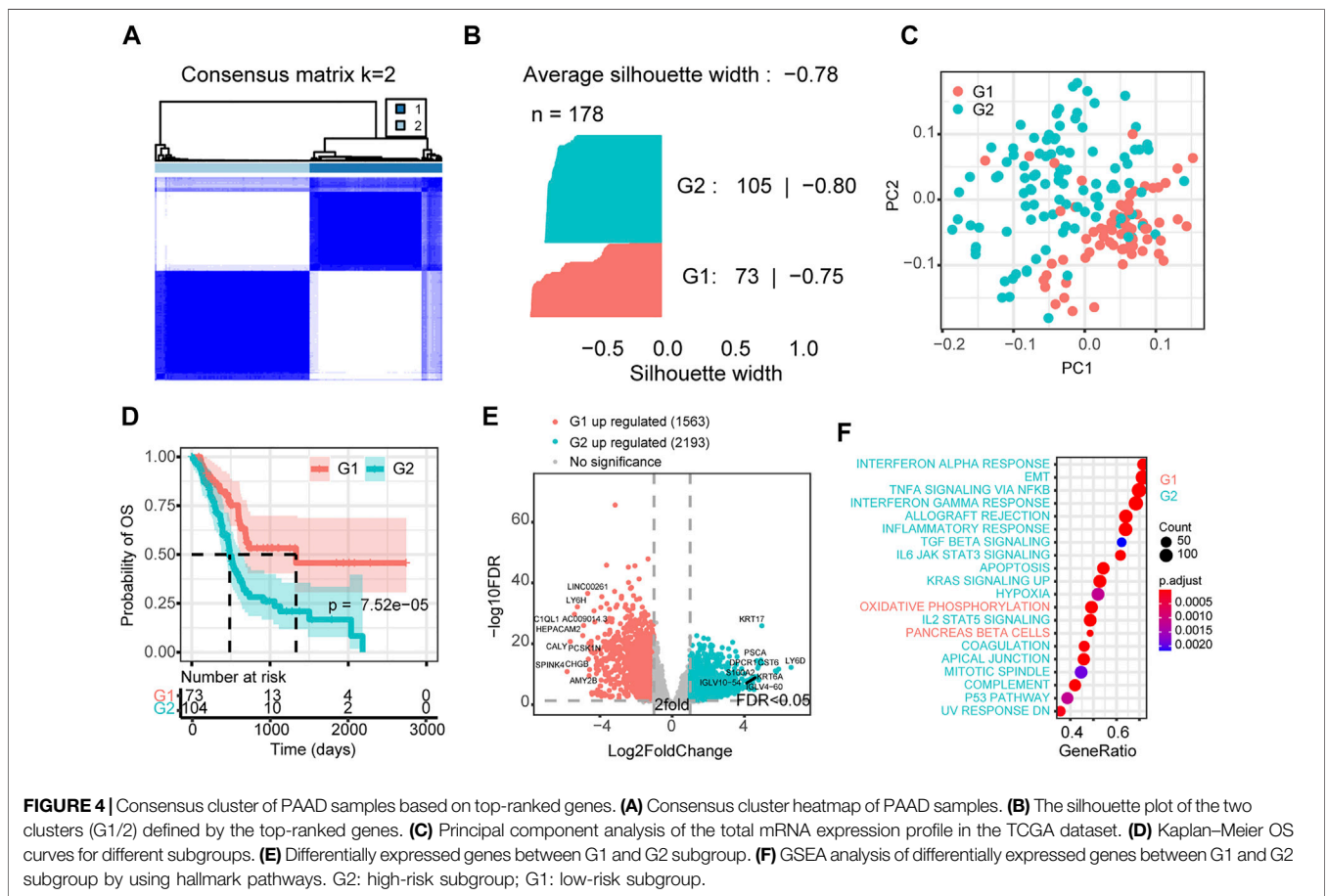

FIGURE 4 | Consensus cluster of PAAD samples based on top-ranked genes. **(A)** Consensus cluster heatmap of PAAD samples. **(B)** The silhouette plot of the two clusters (G1/2) defined by the top-ranked genes. **(C)** Principal component analysis of the total mRNA expression profile in the TCGA dataset. **(D)** Kaplan–Meier OS curves for different subgroups. **(E)** Differentially expressed genes between G1 and G2 subgroup. **(F)** GSEA analysis of differentially expressed genes between G1 and G2 subgroup by using hallmark pathways. G2: high-risk subgroup; G1: low-risk subgroup.

showed that these genes were significantly enriched in some pathways which play important roles in PAAD progression, such as PI3K/AKT/mTOR signaling pathway (Kennedy et al., 2011), EGFR signaling pathway (Williams et al., 2017) and NF-κB

signaling pathway (Prabhu et al., 2014), etc. (**Supplementary Figure S2B**). Survival analysis showed that the expression of most genes (12 of 20) correlated with the survival of PAAD patients (**Figure 3** and **Supplementary Figure S2C**). However, many

genes are less studied, which means their functions in PAAD deserve more exploration.

Comparison of Top-Ranked Genes and CAGs

To further elaborate the superiority of our screened model in PAAD, we selected the top 719 genes and CAG list for unsupervised hierarchical clustering (**Figure 4A** and **Supplementary Figure S3A**). For comparison, we defined both as two classes (**Supplementary Table S3**). Compared with the CAG list derived classification system, the classification system of the PRNet-based signature definition was significantly better with a silhouette width of −0.78 (**Figure 4B** and **Supplementary Figure S3B**). Although there were significant transcriptome differences in both signature-defined subgroups (**Figure 4C** and **Supplementary Figure S3C**), interestingly, the subgroup obtained based on the CAG list did not suggest survival differences (**Supplementary Figure S3D**). The subgroups classified based on the PRNet-defined signature could significantly distinguish survival (**Figure 4D**). We further investigated the differences in gene expression and pathway activity between these two groups. As shown in **Figure 4E**, 1,563 genes were significantly overexpressed in the low-risk group, and 2,193 genes were significantly elevated in the high-risk group (**Supplementary Table S4**). We then analyzed the pathways affected by these genes. Tumour hallmark signaling pathway enrichment analysis suggested that the high-risk group was significantly enriched in interferon response, epithelial-mesenchymal transition, TGFβ and other immune response and oncogenic pathways. In contrast, the low-risk group showed normal pancreatic cells' molecular characteristics, overexpressed some markers of normal pancreatic cells, and showed higher oxidative phosphorylation and ribosomal signaling pathway activities (**Figure 4F**, **Supplementary Figure S4** and **Supplementary Table S5**).

We also compared the association of high and low-risk groups with clinical factors to investigate a significant survival difference between the two groups (**Supplementary Figure S5**). Interestingly, no significant differences were found at either the age or sex level. However, there were more samples of lymph node metastases and distal metastases in the high-risk group, and they were more enriched in stage III and stage IV tumours. This corresponds to a worse prognosis for survival at later clinical stages. The worse overall survival in the high-risk group compared to the low-risk group may be related to their greater susceptibility to lymph node metastasis and distal metastasis.

DISCUSSION

Pancreatic adenocarcinoma is a kind of tumour of highly heterogeneous, which leads a big issue for researchers to get a comprehensive understanding in the study of tumour molecular mechanisms (McGuigan et al., 2018; Yao et al., 2020). Attempts have been made for many years to discover the potential risk-causing genes working in PAAD progression. In our study, an analysis model integrating PPI network and co-expression patterns is proposed, which can identify potential risk causative genes in pancreatic

adenocarcinoma. The importance of candidate genes was scored by PRNet to screen new pivotal regulators. Furthermore, the reliability of the screened candidate genes was validated by kinds of analytical tools.

Based on the PRNet method, we constructed the disease-specific subnetworks. As shown in **Supplementary Figure S1** and **Supplementary Table S2**, pathways related to cell cycle and life activity maintenance are mainly in the subnetwork with the most significant number of genes. And pathways related to tumour progression, such as MAPK and ERBB signaling pathway, etc., presents in the second-largest subnetwork. These results indicate that the networks we screened are closely related to tumours. Cancer is characterized by uncontrolled cell proliferation, which results from the abnormal expression and activities of various cell cycle proteins, making the cell cycle a typical hotspot in tumour research (Sherr 1996). In addition, MAPK, VEGF and other signaling pathways have been proved closely associated with malignant tumour progression in many studies (Ferrara 2005; Wagner and Nebreda 2009; Gargalionis et al., 2018). Therefore, analysis based on these networks may help identify core regulatory proteins and explore their potential as tumour biomarkers.

Looking into the top 20 rank score potential biomarkers in **Figure 2**, some of them have been proved the functions in PAAD, such as XPO1 and FYN. XPO1 mRNA expression is heterogeneous in pancreatic adenocarcinoma and is associated with progression stage and shorter survival (Birnbaum et al., 2019). Nuclear export Inhibitors based on study of XPO1 are used to therapy pancreatic cancer (Muqbil et al. 2018). FYN is one of Src family kinases. Its upregulation is associated with pancreatic cancer metastasis (Chen et al., 2010). Interestingly, our study found some proteins that are not annotated as enzymes, transcription factors, and FDA-validated targets, such as APP, ELAVL1 and YWHAZ. As a potential biomarker with the highest PageRank scores, amyloid precursor protein (APP) is a substrate of proteases and secretase, which is mainly focused on pathogenesis of Alzheimer's disease, was reported been involved in pancreatic cancer (Hansel et al., 2003). And further study figured out that the major enzyme mediating APP cleavage in pancreatic cancer is ASAM10 (Woods and Padmanabhan 2013). It has been suggested that ELAVL1 is an RNA-binding protein, which may play a translational modification regulatory sway in prostate cancer progression (Melling et al., 2016). In addition, ELAVL1 may influence the efficacy of glioma heterogeneous targeted drugs by affecting cell fusion (Filippova and Nabors 2020). ELAVL1 also plays an essential function in tumours such as colorectal cancer (Gu et al., 2019) and breast cancer (Chou et al., 2015; Luo et al., 2020), suggesting that ELAVL1 may indeed serve as a potential marker for PAAD. In contrast, YWHAZ, despite its low gene expression and not being easily detectable, is involved in tumorigenesis and progression in many tumours (Nishimura et al., 2013; Guo et al., 2018; Zhao et al., 2018), including pancreatic cancer (Xue et al., 2018). These results suggest that the potential genes we screened effectively-identified novel markers not found in the standard gene set.

Overall, our proposed model is superior in reconstructing potential tumour markers and identifying new therapeutic targets and revealing more tumour heterogeneity than the traditional CAG list. These results offer a way to study the pancreatic adenocarcinoma's pathogenesis and provide new ideas for the

future drug development and clinical treatment of pancreatic adenocarcinoma.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The Cancer Genome Atlas (TCGA), Datasets link: https://xenabrowser.net/datapages/.

AUTHOR CONTRIBUTIONS

JYW designed and supervised the experiments. LLS, GL and YG performed the experiments and data analysis, XPZ and XYZ contributed to data analysis and predictor development. LLS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.854661/full#supplementary-material

REFERENCES

Alshahrani, M., and Hoehndorf, R. (2018). Semantic Disease Gene Embeddings (SmuDGE): Phenotype-Based Disease Gene Prioritization without Phenotypes. *Bioinformatics* 34 (17), i901–i907. doi:10.1093/bioinformatics/bty559

Barnes, G., and Uriel, F. (1993). *Short Random Walks on Graphs*. San Diego, CA: Association for Computing Machinery.

Bilici, A. (2014). Prognostic Factors Related with Survival in Patients with Pancreatic Adenocarcinoma. *World. J. Gastroenterol.* 20 (31), 10802. doi:10.3748/wjg.v20.i31.10802

Birnbaum, D. J., Pascal, F., Birnbaum, D., Mamessier, E., and Bertucci, F. (2019). XPO1 Expression Is a Poor-Prognosis Marker in Pancreatic Adenocarcinoma. *J. Clin. Med.* 8 (5), 596. doi:10.3390/jcm8050596

Brin, S. (1998). Lawrence %J Computer Networks Page, and ISDN Systems. 1998. "The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Comput. Netw. ISDN Syst.* 30 (1-7), 107

Brown, K. R., and Jurisica., I. (2007). Unequal Evolutionary Conservation of Human Protein Interactions in Interologous Networks. *Genome Biol.* 8 (5), R95–R11. doi:10.1186/gb-2007-8-5-r95

Chen, Z-Y., Cai, L., Bie, P., Wang, S-G., Jiang, Y., Dong, J-H., et al. (2010). Roles of Fyn in Pancreatic Cancer Metastasis. *J. gastroenterology* 25 (2), 293–301. doi:10.1111/j.1440-1746.2009.06021.x

Chou, S.-D., Murshid, A., Eguchi, T., Gong, J., and Calderwood, S. K. (2015). HSF1 Regulation of β-catenin in Mammary Cancer Cells through Control of HuR/elavL1 Expression. *Oncogene* 34 (17), 2178–2188. doi:10.1038/onc.2014.177

Ferrara, N. (2005). VEGF as a Therapeutic Target in Cancer. *Oncology* 69 Suppl 3 (Suppl. 3), 11–16. doi:10.1159/000088479

Filippova, N., and Nabors, L. B. (2020). ELAVL1 Role in Cell Fusion and Tunneling Membrane Nanotube Formations with Implication to Treat Glioma Heterogeneity. *Cancers* 12 (10), 3069. doi:10.3390/cancers12103069

Gargalionis, A. N., Basdra, E. K., and Papavassiliou, A. G. (2018). Mechanosignalling in Tumour Progression. *J. Cell Mol. Med.* 22 (2), 704–705. doi:10.1111/jcmm.13452

Ghandi, M., Huang, F. W., Jané-Valbuena, J., Kryukov, G. V., Lo, C. C., McDonald, E. R., et al. (2019). Next-generation Characterization of the Cancer Cell Line Encyclopedia. *Nature* 569 (7757), 503–508. doi:10.1038/s41586-019-1186-3

Gu, C., Zhang, M., Sun, W., and Dong, C. (2019). Upregulation of miR-324-5p Inhibits Proliferation and Invasion of Colorectal Cancer Cells by Targeting ELAVL1. *Oncol. Res.* 27 (5), 515–524. doi:10.3727/096504018x15166183598572

Guo, F., Jiao, D., Sui, G. Q., Sun, L. N., Gao, Y. J., Fu, Q. F., et al. (2018). Anticancer Effect of YWHAZ Silencing via Inducing Apoptosis and Autophagy in Gastric Cancer Cells. *Neoplasma* 65 (5), 693–700. doi:10.4149/neo_2018_170922N603

Hansel, D. E., Rahman, A., Wehner, S., Herzog, V., Yeo, C. J., and Maitra, A. (2003). Increased Expression and Processing of the Alzheimer Amyloid Precursor Protein in Pancreatic Cancer May Influence Cellular Proliferation. *Cancer Res.* 63 (21), 7032

Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., et al. (2004). IntAct: an Open Source Molecular Interaction Database. *Nucleic Acids Res.* 32 (Suppl. l_1), D452–D455. doi:10.1093/nar/gkh052

Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Ting, L., Baltier, K., Colby, G., et al. (2017). Architecture of the Human Interactome Defines Protein Communities and Disease Networks. *Nature* 545 (7655), 505–509. doi:10.1038/nature22366

Iván, G., and Grolmusz, V. (2011). When the Web Meets the Cell: Using Personalized PageRank for Analyzing Protein Interaction Networks. *Bioinformatics* 27 (3), 405–407. doi:10.1093/bioinformatics/btq680

Kennedy, A. L., Morton, J. P., Manoharan, I., Nelson, D. M., Jamieson, N. B., PawlikowskiPawlikowski, J. S., et al. (2011). Activation of the PIK3CA/AKT Pathway Suppresses Senescence Induced by an Activated RAS Oncogene to Promote Tumorigenesis. *Mol. Cell* 42 (1), 36–49. doi:10.1016/j.molcel.2011.02.020

Kumar, A. A., Van Laer, L., Alaerts, M., Ardeshirdavani, A., Moreau, Y., Laukens, K., et al. (2018). pBRIT: Gene Prioritization by Correlating Functional and Phenotypic Annotations through Integrative Data Fusion. *Bioinformatics* 34 (13), 2254–2262. doi:10.1093/bioinformatics/bty079

Le, N., Sund, M., Vinci, A., Beyer, G., Ashan Javed, M., Krug, S., et al. (2016). Ashan Javed, Sebastian Krug, Albrecht Neessee, and Marvin SchoberPrognostic and Predictive Markers in Pancreatic Adenocarcinoma. *Dig. Liver Dis.* 48 (3), 223–230. doi:10.1016/j.dld.2015.11.001

Li, Y., Kuwahara, H., Yang, P., Song, L., and Gao, X. (2019). PGCN: Disease Gene Prioritization by Disease and Gene Embedding through Graph Convolutional Neural Networks. *bioRxiv* 00 (00), 1–19. doi:10.1101/532226

Loosen, S. H., Neumann, U. P., Trautwein, C., Roderburg, C., and Luedde, T. (2017). Current and Future Biomarkers for Pancreatic Adenocarcinoma. *Tumour Biol.* 39 (6), 1010428317692231. doi:10.1177/1010428317692231

Love, M., Anders, S., and Huber, W. (2014). Differential Analysis of Count Data–The DESeq2 Package. *Genome Biol.* 15 (550), 10–1186. doi:10.1186/s13059-014-0550-8

Luo, N., Zhang, K., Li, X., and Hu, Y. (2020). ZEB1 Induced-upregulation of Long Noncoding RNA ZEB1-AS1 Facilitates the Progression of Triple Negative Breast Cancer by Binding with ELAVL1 to Maintain the Stability of ZEB1 mRNA. *J. Cell Biochem.* 121 (10), 4176–4187. doi:10.1002/jcb.29572

McGuigan, A., Kelly, P., Turkington, R. C., Jones, C., Coleman, H. G., and McCain, R. S. (2018). Pancreatic Cancer: A Review of Clinical Diagnosis, Epidemiology, Treatment and Outcomes. *World. J. Gastroenterol.* 24 (43), 4846–4861. doi:10.3748/wjg.v24.i43.4846

Melling, N., Taskin, B., Hube-Magg, C., Kluth, M., Minner, S., Koop, C., et al. (2016). Cytoplasmic Accumulation of ELAVL1 Is an Independent Predictor of Biochemical Recurrence Associated with Genomic Instability in Prostate Cancer. *Prostate* 76 (3), 259–272. doi:10.1002/pros.23120

Merriel, R. B., Gibbs, P., O'BrienO'Brien, T. J., and Hibbert, M. (2011). BioGrid Australia Facilitates Collaborative Medical and Bioinformatics Research across Hospitals and Medical Research Institutes by Linking Data from Diverse Disease and Data Types. *Hum. Mutat.* 32 (5), 517–525. doi:10.1002/humu.21437

Muqbil, I., Azmi, A., and Mohammad, R. (2018). Nuclear Export Inhibition for Pancreatic Cancer Therapy. *Cancers* 10 (5), 138. doi:10.3390/cancers10050138

Nishimura, Y., Komatsu, S., Ichikawa, D., Nagata, H., Hirajima, S., Takeshita, H., et al. (2013). Overexpression of YWHAZ Relates to Tumor Cell Proliferation and Malignant Outcome of Gastric Carcinoma. *Br. J. Cancer* 108 (6), 1324–1331. doi:10.1038/bjc.2013.65

Prabhu, L., Mundade, R., Korc, M., LoehrerLoehrer, P. J., and Lu, T. (2014). Critical Role of NF-Kb in Pancreatic Cancer. *Oncotarget* 5 (22), 10969–10975. doi:10.18632/oncotarget.2624

Qian, J., Olbrecht, S., Boeckx, B., Vos, H., Laoui, D., Etlioglu, E., et al. (2020). A Pan-Cancer Blueprint of the Heterogeneous Tumor Microenvironment Revealed by Single-Cell Profiling. *Cell Res.* 30 (9), 745–762. doi:10.1038/s41422-020-0355-0

Sherr, C. J., and Charles, J. (1996). Cancer Cell Cycles. *Science* 274 (5293), 1672–1677. doi:10.1126/science.274.5293.1672

Shi, J., and Walker, M. (2007). Gene Set Enrichment Analysis (GSEA) for Interpreting Gene Expression Profiles. *Curr. Bioinfo.* 2 (2), 133–137. doi:10.2174/157489307780618231

Siegel, R. L., Miller, K. D., and Jemal., A. (2020). Cancer Statistics, 2020. *CA A Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590

Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I., and Forbes, S. A. (2018). The COSMIC Cancer Gene Census: Describing Genetic Dysfunction across All Human Cancers. *Nat. Rev. Cancer* 18 (11), 696–705. doi:10.1038/s41568-018-0060-1

Wagner, E. F., and Nebreda, Á. R. (2009). Signal Integration by JNK and P38 MAPK Pathways in Cancer Development. *Nat. Rev. Cancer* 9 (8), 537–549. doi:10.1038/nrc2694

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Williams, M., Lomberk, G., and Urrutia, R. (2017). EGFR (ErbB) Signaling Pathways in Pancreatic Cancer Pathogenesis. *J. Pancreat. Cancer*, 1–26. doi:10.1007/978-1-4939-6631-8_15-2

Wold, S., Kim, E., and Paul, G. (1987). Principal Component Analysis. *Chemom. Intelligent Laboratory Syst.* 2 (1-3), 37–52. doi:10.1016/0169-7439(87)80084-9

Woods, N., and Padmanabhan, J. (2013). Role of ADAM10-Mediated APP Processing in Pancreatic Cancer. *AACR*.

Xing, W., and Ali, G. (2004). *Weighted Pagerank Algorithm*. Fredericton, New Brunswick: IEEE Computer Society.

Xue, R., Lin, H., Xu, W., Gao, Y., Pang, Y., and Jianyu, H. (2018). Derivation and Validation of the Potential Core Genes in Pancreatic Cancer for Tumor-Stroma Crosstalk. *BioMed Res. Int.* 2018, 4283673. doi:10.1155/2018/4283673

Yao, W., Maitra, A., and Ying, H. (2020). Recent Insights into the Biology of Pancreatic Cancer. *EBioMedicine* 53, 102655. doi:10.1016/j.ebiom.2020.102655

Zhao, J.-F., Zhao, Q., Hu, H., Liao, J.-Z., Lin, J.-S., Xia, C., et al. (2018). The ASH1-miR-375-YWHAZ Signaling axis Regulates Tumor Properties in Hepatocellular Carcinoma. *Mol. Ther. - Nucleic Acids* 11, 538–553. doi:10.1016/j.omtn.2018.04.007