RNA-sequencing for transcriptional profiling of whole blood in early stage and metastatic pancreatic cancer patients

Sima Kalantari1 | Bahram Kazemi1 | Raheleh Roudi2 | Hakimeh Zali3,4 | Alberto D'Angelo5 | Ashraf Mohamadkhani6 | Zahra Madjd7 | Akram Pourshams6,8

1 Cellular and Molecular Biology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2 Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, USA
3 Proteomics Research Center, Shahid Beheshti University of Medical Science, Tehran, Iran
4 Department of Tissue Engineering and Applied Cell Sciences, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
5 Department of Biology and Biochemistry, University of Bath, Bath, UK
6 Liver and Pancreatobiliary Diseases Research Center, Digestive Disease Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran
7 Oncopathology Research Center, Iran University of Medical Sciences, Tehran, Iran
8 Digestive Oncology Research Center, Digestive Diseases Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

Correspondence
Akram Pourshams, Liver and Pancreatobiliary Diseases Research Center, Digestive Disease Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran. Email: akrampourshams@gmail.com

Abstract
We investigated the transcriptional profile of whole blood in early and metastatic stages of pancreatic cancer (PaC) patients to identify potential diagnostic factors for early diagnosis. Blood samples from 18 participants (6 healthy individuals, 6 patients in early stage (I/II) PaC, and 6 patients in metastatic PaC) were analyzed by RNA-sequencing. The expression levels of identified genes were subsequently compared with their expression in pancreatic tumor tissues based on TCGA data reported in UALCAN and GEPIA2 databases. Overall, 331 and 724 genes were identified as differentially expressed genes in early and metastatic stages, respectively. Of these, 146 genes were shared by early and metastatic stages. Upregulation of PTCD3 and UBA52 genes and downregulation of A2M and ARID1B genes in PaC patients were observed from early stage to metastasis. TCGA database showed increasing trend in expression levels of these genes from stage I to IV in pancreatic tumor tissue. Finally, we found that low expression of PTCD3, A2M, and ARID1B genes and high expression of UBA52 gene were positively correlated with PaC patients survival. We identified a four-gene set (PTCD3, UBA52, A2M, and ARID1B) expressed in peripheral blood of early stage and metastatic PaC patients that may be useful for PaC early diagnosis.

KEYWORDS
blood profiling, early stage, metastasis, pancreatic cancer, RNA-sequencing

1 | INTRODUCTION

Pancreatic cancer (PaC), a highly fatal malignancy with a 5-year survival rate of only 7%–8%, is the seventh leading cause of global cancer death (Siegel et al., 2016). The high mortality rate of PaC is generally due to a delayed diagnosis at an advanced stage (Illés et al., 2020) when approximately 80% of patients have local spread or metastatic disease (Vareedayah et al., 2018). The stomach, spleen, colon, and left adrenal gland are potential sites of extra-pancreatic local disease invasion. In more advanced stages of PaC, liver, lymph nodes, and peritoneum are common sites of distant metastases (Yadav & Lowenfels, 2013). There are different types of PaC although the term “pancreatic cancer” is usually referred to pancreatic adenocarcinoma. Indeed, more than 95% of
pancreas malignant neoplasms involve the exocrine portion of the pancreas (duct system and acinar cells) with adenocarcinoma features (Hruban & Adsay, 2009).

Surgery, chemotherapy, and radiation therapy are treatment options for PaC patients with surgery the only potentially curative treatment (Vareedayah et al., 2018). Based on cancer stage, PaC patients are usually classified into four distinct groups (stages I-IV), which determine their treatment options. Patients in stage I have resectable tumors while stage II includes patients with “borderline resectable” tumors. In stage III, patients have locally advanced unresectable disease, with no metastasis. The final stage (stage IV) includes patients with metastatic disease, who are not eligible for surgical resection (Garcea et al., 2012; Karmazovsky et al., 2005). Unfortunately, the absence of specific symptoms in early stage of disease and the ability of pancreatic cancer to early metastasize limit the effectiveness of surgical resection, with only 15%-20% of patients eligible for curative pancreatectomy at the time of the diagnosis (Vareedayah et al., 2018).

Early diagnosis of PaC within the potentially resectable tumor window would represent a significant increase of patients’ survival and a decrease in the mortality rate. Challenges to the early PaC diagnosis include the high cost of imaging investigations—namely endoscopic ultrasound and cross-sectional imaging—and the lack of specific biomarkers. Currently, CA19-9 (carbohydrate antigen 19-9) is the only blood-based biomarker for pancreatic cancer diagnosis with 75% sensitivity and 90% specificity (Kim et al., 2004) which, due to the high rate of false-positive results (high values of CA19-9 in the absence of malignancy) (Ballehaninna & Chamberlain, 2012) it is not widely considered as an accurate test for PaC early diagnosis.

Previously, in a prospective study on 461 PaC cases in the Iranian population, we analyzed the clinical, pathological as well as therapeutic and survival features of these patients to provide reliable information of Iranian PaC patients. We observed that among 461 studied cases of PaC only 43 (9.5%) individuals were classified as stage I (with resectable tumor) at the time of diagnosis, and the 5-year survival of Iranian PaC patients was lower than the developed countries. We concluded that socioeconomic aspects alongside the stage and pathologic features of PaC may affect the probability of receiving timely treatment and, consequently, patient survival (Sheikh et al., 2020). In another study, published in 2021, we focused on the potential association between diabetes mellitus (DM) and PaC, to identify common molecular factors between these two conditions with the ultimate goal of PaC early diagnosis. In that study, we proposed the potential role of SPI1 and YY1 genes to investigate the relationship between DM and PaC and provide opportunities for PaC screening using blood tests (Kalantari et al., 2021). As further perspective of our investigation on pancreatic cancer, in the current project we focused on gene expression profiling of blood samples for early stage and metastatic PaC patients. The main goal of the study was the identification of differentially expressed genes (DEGs) in PaC patients compared to normal individuals with an emphasis on those shared DEGs between early and metastatic stages, to evaluate what molecular factors are suitable for the development of a blood test that could detect PaC in its early and therefore resectable stage.

2 | METHODS

2.1 | Study design summary

The current project was performed in three main steps including:

- Specimen collection
- RNA isolation and sequencing
- Bioinformatics analysis

2.2 | Specimen collection

2.2.1 | Study population

Patients recruited for the current study were selected among those who were doubtful to have a pancreatic mass and subsequently referred to the tertiary referral hospital (Shariati Hospital, Tehran, Iran) during 2019. First, informed written consent was obtained from every participant and a validated and reliable questionnaire was used to collect detailed demographic information, lifestyle as well as different exposures. The endoscopic ultrasonography was performed on every participant and in case of mass or a metastatic lesion identification, the patient underwent fine-needle aspiration (FNA). If the diagnosis of pancreatic cancer could not be achieved using hematoxylin/eosin staining, the samples were evaluated by immunohistochemistry staining. The collected pancreatic tissue samples were then reviewed by an expert pathologist who was blinded to the questionnaire data. Finally, the pancreatic mass and tumor staging were reconfirmed by an expert gastroenterologist. The pancreatic cancer cases at early (I/II) and metastatic (IV) stages were included in the study and blood samples were collected from all enrolled patients before surgery or any other medical intervention.

2.2.2 | Ethical statement

Sample collection procedures performed in this study involving human participants were conducted according to the principles expressed in the institutional and/or national research committee of the 2013 Helsinki declaration. This study was approved by the Institutional Review Board and Ethics Committees of the Digestive Disease Research Institute of Tehran University of Medical Sciences and the National Institute for Medical Research Development (approval number: IR NIMAD REC. I398.180).

2.3 | RNA isolation and sequencing

2.3.1 | Total RNA extraction, purification, and quantification

A volume of 600 μl peripheral whole blood (for each group) was collected into an EDTA tube and total RNA extraction was performed
using QIAzol Lysis Reagent (Qiagen) according to the manufacturers’ protocol. Consequently, DNase I, RNase-free Kit (Thermo Scientific) was used to remove potential DNA contamination from extracted RNA. Concentration, RNA integrity number (RIN), 28 S/18 S ratio and size of RNA were assessed using Agilent 2100 Bioanalyzer (Agilent RNA 6000 Nano Kit, Agilent Technologies) and samples with RIN ≥ 6 were used for sequencing. The purity of the samples was tested by NanoDrop ND-100 spectrophotometer (Nanodrop Technologies).

Library construction and RNA sequencing. Library construction was carried out according to the following steps:

a. mRNA isolation: 200 ng total RNA was purified using oligo-dT beads, followed by TruSeq RNA Sample Prep Kit v2 (Illumina) protocol for Library construction.

b. Fragmentation: The RNA was fragmented into small pieces with Elute, Prime, Fragment Mix.

c. cDNA synthesis: First-strand cDNA was generated by First Strand Master Mix and Super Script II (Invitrogen) reverse transcription, under reaction conditions of 25°C for 10 min, 42°C for 50 min, and 70°C for 15 min. The product was purified by Agencourt RNAClean XP Beads (AGENCOURT) and then Second Strand Master Mix and dATP, dCTP, dGTP, and dUTP mix were added to synthesize the second-strand cDNA (16°C for 1 h).

d. End repair and A-Tailing: The purified fragmented cDNA was combined with End Repair Mix and incubated at 30°C for 30 min. The end-repaired DNA was purified with Ampure XP Beads (AGENCOURT). Then A-Tailing Mix was added, mixed, and incubated at 37°C for 30 min.

e. Adapter ligation: The Adenylate 3'Ends DNA, RNA Index Adapter and Ligation Mix were combined, mixed and incubated at 30°C for 10 min. Then end-repaired DNA was purified with Ampure XP Beads (AGENCOURT).

f. UNG digestion: The Uracil-N-Glycosylase enzyme was added into the samples containing the DNA and incubated at 37°C for 10 min. The product underwent a second purification step with AGENCOURT Ampure XP Beads.

g. PCR: The PCR Primer Cocktail and PCR Master Mix were used to perform 15 cycles of PCR amplification to enrich the cDNA fragments. Finally, the PCR products were purified by Ampure XP Beads (AGENCOURT).

The constructed libraries were assessed by checking the distribution of the fragments size using the Agilent 2100 bioanalyzer (Agilent DNA 1000 Reagents, Agilent Technologies). The libraries then were amplified on cBot to generate the cluster on the flow cell (TruSeq PE Cluster Kit V3-cBot–HS Illumina), and therefore the amplified flow cell was sequenced using pair-end strategy (100 bp paired-end read) on the HiSeq. 4000 System (TruSeq SBS KIT-HS V3 Illumina). Quality control of FASTQ files was assessed using FastQC:Read QC and Trimmomatic tool was employed to check quality of trimming and adapter removal. The trimmed raw data were mapped to the reference genome (Ensembl GRCh38/hg19 assembly) via Hisat2 tool for transcripts alignment and quantification. Differentially expressed genes were identified using Cuffdiff (Cufflinks package version 2.1.1) tool and those with log fold change (logFc) ≥ ±1.5 and p-value ≤ 0.05 were considered for further bioinformatics analyses. Additionally, the Venn diagram online tool (https://bioinformatics.psb.ugent.be/webtools/Venn/) was used to show the shared DEGs between experimental groups.

2.4 Bioinformatics analysis

2.4.1 GEPIA2 data set

GEPIA2 (Gene Expression Profiling Interactive Analysis2) (http://gepia2.cancer-pku.cn/index.html) is a newly developed interactive platform to analyze the RNA sequencing expression data of tumors and normal samples from the TCGA (The Cancer Genome Atlas) and the GTEx (Genotype-Tissue Expression) projects, based on a standard processing pipeline (Tang et al., 2017). We used GEPIA2 data set for pan-cancer view and survival map of identified DEGs obtained from our RNA sequencing analysis.

2.4.2 UALCAN data set

UALCAN (http://ualcan.path.uab.edu/) is another interactive online platform for cancer OMICS data analysis. This platform allows researchers to perform in silico validation of potential genes of interest, identification of biomarkers, generation of graphs and plots of expression profile and patient survival information, with easy access to public available TCGA, CPTAC, and MET500 cancer OMICS data (Chandrashekar et al., 2017). We used UALCAN data set to show expression level of identified DEGs in pancreatic cancer tissue compared to normal tissue, as well as expression levels of DEGs according to pancreatic tumor grade, individual cancer stages, and nodal metastasis. We also used UALCAN data set to generate Kaplan-Meier survival curves to compare the high and low expression of DEGs in pancreatic cancer and correlation of DEGs with patients overall survival.

3 RESULTS

3.1 Demographics

The whole blood samples collected from 18 participants (12 patients with histologically confirmed PaC and 6 healthy individuals) were divided into three groups as follows: early stage (I/II) PaC, metastatic PaC, and normal control. For each group, six blood samples were pooled and RNA-sequencing was performed on these pooled samples. The demographic information and clinical characteristics for each group are detailed in Table 1.
3.2 RNA-sequencing analysis

The RNA-sequencing data analysis and reference genome mapping for differentially expressed genes (DEGs) of each group was carried out. Compared to normal control, the number of DEGs identified in early stage and metastatic PaC patients were 13,309 and 19,030, respectively. After the exclusion of those DEGs with not significant logFc (= expression values near to zero), 331 genes of the early stage group and 724 genes of the metastatic group were considered statistically significant DEGs based on logFc and p-value parameters. Tables with genes ID, logFc, and p-value for each experimental group are reported in Online Supplementary Files S1 and S2. As we aimed to identify the shared DEGs between early stage and metastatic PaC patients, using the Venn diagram tool we identified 146 genes shared between these two experimental groups. The result is shown in Figure 1. These 146 genes are categorized as up and downregulated genes in Online Supplementary File S3. Considering logFc ≥ ±1.5 and p-value ≤ .05 (as the cutoff criteria), we identified PTCD3, A2M, UBA52, and ARID1B as statistically significant shared DEGs between early and metastatic stages. Hence, we performed further bioinformatics analysis on these identified genes.

3.3 Pancancer view and survival map of shared DEGs

Using GEPIA2 data set analysis, we compared the transcription levels of PTCD3, A2M, UBA52, and ARID1B genes between different types of cancers and normal tissue. The resulting heatmap showed the different expression levels of these four shared DEGs in almost all types of cancers compared to normal tissue. The survival map showed the impact of four shared DEGs on the overall survival rate of patients in different types of cancers, with UBA52 gene showing the most significant role in the overall survival of patients with adrenocortical carcinoma (ACC), kidney chromophobe (KICH), and colon adenocarcinoma (COAD). The heatmap and survival map of the four shared DEGs are represented in Figure 2.

3.4 Shared DEGs analysis based on TCGA data

UALCANCAN data set analysis, which is based on TCGA data, revealed a decrease in expression levels of PTCD3, A2M, and ARID1B genes in pancreatic cancer tissue compared with normal pancreatic tissue. The differential expression was not considerable for the UBA52 gene between tumor and normal pancreas. The box plot in Figure 3 represents the different transcription levels of these four shared DEGs.

We subsequently investigated PTCD3, A2M, UBA52, and ARID1B gene expression levels in different PaC stages. The stage plot of these shared DEGs reported a decrease of gene expression levels for PTCD3, A2M, and ARID1B genes in PaC stages I–IV when compared to healthy individuals but UBA52 gene did not show a significant differential expression. The box plot in Figure 4 shows the expression levels of these four shared DEGs based on individual pancreatic cancer stages. Based on pancreatic tumor grade, TCGA data analysis showed low expression of PTCD3, A2M, and ARID1B genes although no significant expression of UBA52 gene in grades 1–4 of pancreatic tumour compared to normal samples. Moreover, we investigated the expression levels of these four shared DEGs in nodal metastasis and we observed that expression levels of PTCD3, A2M, and ARID1B for both metastatic and no-metastatic regional lymph nodes were lower than normal samples. However, expression levels were almost equal for metastatic and no-metastatic regional lymph node samples. In this analysis, UBA52 expression was not significant. TCGA data analysis of four shared DEGs based on pancreatic tumor grade and nodal metastasis are represented in Figures 5 and 6, respectively.

Using the Kaplan–Meier plot, we compared the overall survival rate of PaC patients according to the high and low expression of PTCD3, A2M, UBA52, and ARID1B genes. The results indicated that low expression of PTCD3, A2M, and ARID1B genes has a positive effect on patient survival while UBA52 gene high expression level was positively correlated with survival of
FIGURE 2  The pancancer view and survival map of four shared DEGs (shared between experimental groups) (the full name of the cancers has been mentioned in the abbreviation section). (a) The heatmap plot represents the transcript expression levels of four shared DEGs in different types of cancers compared with normal samples. T: tumor, N: normal. (b) The survival map represents the impact of four obtained DEGs on overall survival rate of patient in different types of cancers. DEGs, differentially expressed genes; HR, hazardous ratio. (by GEPIA2 data set analysis).

FIGURE 3  The box plot represents the different transcription levels of four shared DEGs in pancreatic cancer and normal samples. Blue box: normal tissue (n = 4). Red box: tumor tissue (n = 178) (by UALCAN data set analysis). DEGs, differentially expressed genes.
PaC patients. The Kaplan–Meier plot of four shared DEGs is represented in Figure 7.

### 4 | DISCUSSION

It is commonly agreed that early diagnosis of pancreatic cancer will significantly improve the success of treatment and, consequently, increase the overall survival rate of patients. However, early diagnosis of PaC is not straightforward and current biomarkers have limited utility to detect PaC in the early stage (Ballehaninna & Chamberlain, 2012). At present, biomarkers identified as potential compounds for the early detection of PaC include proteins (Brand et al., 2011; Chan et al., 2014; Gold et al., 2010; Mirus et al., 2015; Pan et al., 2011; Wingren et al., 2012), miRNAs (Cote et al., 2014; Li et al., 2013; Schultz et al., 2014), metabolites (Mayers et al., 2014; Ritchie et al., 2013; Xie et al., 2015), aberrantly glycosylated antigens (CA19-9 and Tn antigens) (Haab et al., 2015; Kim et al., 2002; Nanashima et al., 1999), autoantibodies (Capello et al., 2013; Nagayoshi et al., 2014), and exosomes (Capello et al., 2017; Slater et al., 2013). Some studies have also used the microarray technique to investigate the gene expression profile of peripheral blood mononuclear cells (PBMCs) in PaC patients (Baine et al., 2011; Caba et al., 2014; Huang et al., 2010). In contact with tumor tissue PBMCs can alter their gene expression profile and induce a complex transcriptional response (Liotta et al., 2003; Twine et al., 2003; Whitney et al., 2003). Therefore, studies on blood gene expression profiling may contribute to identify reliable cancer biomarkers.

In the current project, a transcriptome study of differentially expressed genes in blood samples of PaC patients was carried out. We aimed to identify differentially expressed genes from early stage to metastatic PaC, considering the following three inclusion criteria for patients selection: whole blood sample, the participants’ age between 50–80 years old and patient’s samples collected before surgery or any other medical intervention. Using RNA-sequencing technique, we first identified differentially expressed genes in early stage PaC and metastatic PaC cases compared with normal individuals (significant DEGs of early stage group = 331, metastatic group = 724). Then using the Venn diagram tool, 146 genes were identified as shared DEGs between early stage and metastatic PaC patients; 103 genes were upregulated during cancer progression (from early stage to metastasis) while remaining genes were found downregulated. Based on logFc ≥
±1.5 and p-value ≤ .05 as the cutoff criteria, we finally identified PTCD3 and UBA52 genes upregulate while A2M and ARID1B genes were downregulated (see Online Supplementary File S3).

The AT-rich interaction domain-containing protein 1B (ARID1B) is a member of the human switch/sucrose nonfermentable (SWI/SNF) chromatin remodeling complex (Aso et al., 2015). SWI/SNF is known as a tumor suppressor and mutations of this complex are widespread across a variety of human cancers (Oike et al., 2013; Shain & Pollack, 2013). The loss of ARID1B has been reported in breast (Oike et al., 2013), colorectal (Cajuso et al., 2014), and pancreatic (Shain et al., 2012) cancers as well as malignant melanoma and hepatocellular carcinoma (Oike et al., 2013). Khursheed et al. carried out a study on the tumor-suppressor activity of ARID1B gene and reported its reduced expression level in advanced-stage pancreatic tumors when compared with normal pancreas (Khursheed et al., 2013). In line with this report, we observed a decreased expression level of the ARID1B gene in blood samples of PaC patients compared to normal sample and decreased expression level pattern of the ARID1B gene from early stage to metastatic PaC (logFc: early stage = −4.50163, metastatic = −6.18771).

Low expression levels of A2M were also observed in blood samples of PaC cases compared to normal samples with decreased pattern from early stage to metastatic PaC (logFc: early stage = −3.13408, metastasis = −4.01882). Alpha-2-macroglobulin (A2M) plays a key role as a universal proteinase inhibitor (Granger et al., 2005) by covalently binding to different proteinase, such as carboxyl, thiol, serine, and metalloprotease (Zhu et al., 2021). This glycoprotein can also bind to various growth factors, cytokines and other non-proteolytic molecules (Burgess et al., 2008). Due to its unique molecular structure, A2M can bind to proteinases and plays its anti-inflammatory function, especially for proteinase released from granulocytes and other inflammation-related cells (Zhu et al., 2021).

Overexpression matrix metalloproteinase 9 (MMP9) in blood has been correlated with cell-growth, invasion, and tumor metastasis in breast, colorectal, and pancreatic cancers (Hurst et al., 2007; Kapischke et al., 2008; Somiari et al., 2006). In contrast, inhibition of the MMP9 activity has been associated to primary tumor mass shrinkage and metastasis probability (Kapischke et al., 2008). Huang et al reported that MMP9 increased level in PBMCs might support PaC development (Huang et al., 2010). At early cancer stages,
macrophages are recruited into the tumor microenvironment for their antitumour immune activity but can differentiate into tumor-associated macrophages due to specific signals from the tumor microenvironment, leading to the endothelial to mesenchymal transition (EMT) and metastasis in tumor cells (Adamska et al., 2017; Weizman et al., 2014). Tekin et al study showed that, protease-activated receptor 1 (PAR1) plays a pivotal role in driving macrophage recruitment into the pancreatic tumor microenvironment. This recruited macrophage can secrete MMP9, undergoing tumor cells to EMT process. They finally speculated that the recruited naive macrophages within the tumor microenvironment can differentiate into tumor-associated macrophages, thus supporting MMP9-PAR1-EMT dependent process and facilitating tumor progression (Tekin et al., 2020). These studies support our results of downregulated A2M (as proteinase inhibitor) gene expression levels in the blood of PaC patients compared to normal individuals, as well as its decreased gene expression pattern from early to metastatic stage.

Our study indicated that pentatricopeptide repeat domain protein 3 (PTCD3) was downregulated in blood samples of PaC patients compared with normal individuals whereas its expression level increased from early to metastatic stage (logFc: early stage = -5.73478, metastasis = -4.13016). The oncogenic role of PTCD3 was reported in breast and prostate cancers as well as lymphoma (Huang et al., 2018), and can corroborate our results given its upregulation during the tumor progression. High UBA52 (Ubiquitin A-52) expression level was also observed in our experimental groups compared to normal sample and its expression level was found upregulated during tumor progression (logFc: early stage = 2.99128, metastasis = 5.04053). Noteworthy, UBA52 activates RPL40-MDM2-p53 pathway and, consequently, induces cell cycle arrest and apoptosis in tumor tissue (e.g., colorectal cancer, as reported by Zhou et al. (2019). We can therefore speculate that UBA52 plays a key role in tumor inhibition, although its upregulation in our study was in contrast with this report. Hence, we suppose that UBA52, when expressed in blood cells, may be involved in other pathways resulting in its upregulation during tumor progression.

Lastly, we compared the expression levels of PTCD3, UBA52, A2M, and ARID1B in blood and pancreatic tumor tissue according to individual cancer stages, pancreatic tumor grade, and nodal metastasis, reported in TCGA database. Similarly, we observed low expression of PTCD3, A2M, and ARID1B genes in blood and pancreas tumor tissue compared to normal individuals. But UBA52 showed high
expression in blood samples although its expression was not significant in tumor tissue compared to normal control (see Figure 3 and Online Supplementary File S3). Similar result was observed for grades 1–4 of pancreatic tumors and in metastatic and no-metastatic regional lymph nodes (see Figures 5 and 6). The stage plot illustrated sinusoidal changes in expression levels of PTCD3, A2M, and ARID1B genes from stage I to IV although their increasing trend was observed during cancer progression. While UBA52 expression showed a slight increase from stage I to IV (see Figure 4). The sinusoidal changes in expression levels of PTCD3, A2M, and ARID1B were also observed in Grade 1–4 of pancreatic tumors. The PTCD3 and ARID1B genes showed a decreasing trend, but A2M gene showed an increasing trend and UBA52 showed no significant expression during tumor progression (see Figure 5). Additionally, we observed that UBA52 expression levels were not significant in metastatic and no-metastatic regional lymph nodes tumors compared with normal control. While expression levels of PTCD3, A2M, and ARID1B genes in metastatic and no-metastatic regional lymph nodes were lower than normal samples, although these expression levels were almost equal in metastatic and no-metastatic regional lymph node tumors (see Figure 6). Using Kaplan–Meier plot we finally found that low expression of PTCD3, A2M, and ARID1B genes in PaC patients has positive effect on PaC patients overall survival while high expression level of UBA52 gene was in positively correlated with survival of PaC patients (see Figure 7).

4.1 Conclusion and future perspective

We hypothesized that the identification of genes differentially expressed from early stage to metastatic PaC patients in peripheral blood could be helpful to design noninvasive and cost-effective diagnostic tests for early diagnosis of pancreatic tumor in respectable stages. Using RNA-sequencing analysis, we identified four DEGs in the blood transcriptome profile of pancreatic cancer patients that were shared between early and metastatic stages and we further investigated their expression trend during cancer progression. Overall, 146 DEGs were identified although we subsequently focused on those statistically significant DEGs according to jointly agreed

**FIGURE 7** The Kaplan–Meier plot represents the impact of four shared DEGs on overall survival rate of patients with pancreatic cancer. In all graphs, blue and red lines demonstrate low and high expression levels of transcript, respectively. High expression level \(n = 45\), low expression level \(n = 132\) (by UALCANN data set analysis). DEGs, differentially expressed genes; PaAD, pancreatic adenocarcinoma.
cut-off criteria. The small number of samples was the main limitation of the current study. We believe that study including more samples could draw more accurate conclusions. Of note, there is growing evidence demonstrating that platelets might have an important role in the development and progression of cancer and their mRNA profile is currently emerging as a new potential source in cancer biomarker research. Unlike other studies — that mainly focus on peripheral blood mononuclear cells — we studied whole blood samples of PaC cases. We, therefore, suggest investigating on our identified DEGs, with particular emphasis on their expression levels in platelets. Finally, as extended investigation, we propose validation of our four DEGs on RNA and protein level as well as the investigation on their regulatory networks, to shed light on important molecular pathways of blood cells in pancreatic cancer progression.

AUTHOR CONTRIBUTIONS
Sima Kalantari: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; writing — original draft; writing — review & editing. Bahram Kazemi: Investigation; supervision; writing — review & editing. raheleh roudi: investigation; methodology; project administration; writing — review & editing. Hakimeh Zali: Formal analysis; methodology; writing — review & editing. Alberto D’Angelo: Methodology; writing — review & editing. ashraf mohamadkhani: formal analysis; methodology; writing — review & editing. Zahra Madjd: Methodology; supervision; writing — review & editing. Akram PourshAMS: Conceptualization; investigation; methodology; project administration; supervision; validation; visualization; writing — review & editing.

ACKNOWLEDGEMENTS
We thank all the patients and healthy donors who participated in the present study. We are grateful to Dr. Mojgan Bandehpour for RNA isolation procedure consulting, Beijing Genomics Institute (BGI.Group) for RNA-sequencing, Dr. Babak Arefnejad for RNA-sequencing analysis, and all those who helped in this project. This project was funded by the National Institute for Medical Research Development (grant number: 982950, grant recipient: Akram Pourshams).

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
All data generated or analyzed during this study are included in this published article and its Supplementary Information files.

ORCID
Akram Pourshams @ http://orcid.org/0000-0002-7950-3983

REFERENCES
Adamska, A., Domenichini, A., & Falasca, M. (2017). Pancreatic ductal adenocarcinoma: current and evolving therapies. International Journal of Molecular Sciences, 18(7), 1338.

Aso, T., Uozaki, H., Morita, S., Kumagai, A., & Watanabe, M. (2015). Loss of ARID1A, ARID1B, and ARID2 expression during progression of gastric cancer. Anticancer Research, 35(12), 6819–6827.

Baine, M. J., Chakraborty, S., Smith, L. M., Mallya, K., Sasson, A. R., Brand, R. E., & Batra, S. K. (2011). Transcriptional profiling of peripheral blood mononuclear cells in pancreatic cancer patients identifies novel genes with potential diagnostic utility. PLoS One, 6(2), e17014.

Ballehaninna, U. K., & Chamberlain, R. S. (2012). The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. Journal of Gastrointestinal Oncology, 3(2), 105–119.

Brand, R. E., Nolen, B. M., Zeh, H. J., Allen, P. J., Eloubeidi, M. A., Goldberg, M., Elton, E., Arnoletti, J. P., Christein, J. D., Vickers, S. M., Langmead, C. J., Landsittel, D. P., Whitcomb, D. C., Grizzle, W. E., & Lokshin, A. E. (2011). Serum biomarker panels for the detection of pancreatic cancer. Clinical Cancer Research, 17(4), 805–816.

Burgess, E. F., Ham, A. J. L., Tabb, D. L., Billheimer, D., Roth, B. J., Chang, S. S., Cookson, M. S., Hinton, T. J., Cheek, K. L., Hill, S., & Piettenpol, J. A. (2008). Prostate cancer serum biomarker discovery through proteomic analysis of alpha-2 macroglobulin protein complexes. Proteomics - Clinical Applications, 2(9), 1223–1233.

Caba, O., Prados, J., Ortiz, R., Jiménez-Luna, C., Melguizo, C., Álvarez, P. J., Delgado, J. R., Irigoyn, A., Rojas, I., Pérez-Florido, J., Torres, C., Perales, S., Linares, A., & Aráneaga, A. (2014). Transcriptional profiling of peripheral blood in pancreatic adenocarcinoma patients identifies diagnostic biomarkers. Digestive Diseases and Sciences, 59(11), 2714–2720.

Cajuso, T., Hánninen, U. A., Kondelin, J., Gyfle, A. E., Tanskanen, T., Katainen, R., Pitkänen, E., Ristolainen, H., Kaasinen, E., Taipale, M., Taipale, J., Böhm, J., Renkonen-Sinisalo, L., Mecklin, J. P., Järvinen, H., Tuupanen, S., Klipivaara, O., & Vaheristo, P. (2014). Exome sequencing reveals frequent inactivating mutations in ARID1A, ARID1B, ARID2 and ARID4A in microsatellite unstable colorectal cancer. International Journal of Cancer, 133(3), 611–623.

Capello, M., Bantis, L. E., Scoio, G., Zhao, Y., Li, P., Dhillon, D. S., Patel, N. J., Kundnani, D. L., Wang, H., Abbruzzese, J. L., Maitra, A., Tempero, M. A., Brand, R., Brennan, L., Feng, E., Taguchi, I., Janout, V., Firpo, M. A., Mulvihill, S. J., ... Hanash, S. M. (2017). Sequential validation of blood-based protein biomarker candidates for early-stage pancreatic cancer. Journal of the National Cancer Institute, 109(4), dwj266.

Capello, M., Cappello, P., Linty, F. C., Chiarle, R., Sperduti, I., Novarino, A., et al. (2013). Autoantibodies to Ezrin are an early sign of pancreatic ductal adenocarcinoma: An evidence based appraisal. Journal of Hematology & Oncology, 6(1), 1–13.

Chan, A., Prassas, I., Dimitromanolakis, A., Brand, R. E., Serra, S., Diamandis, E. P., & Blasutig, I. M. (2014). Validation of biomarkers that complement CA19. 9 in detecting early pancreatic cancer. Clinical Cancer Research, 20(22), 5787–5795.

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarti, B. V. S. K., & Varambally, S. (2017). UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia, 19(8), 649–58.

Cote, G. A., Gore, J. A., McElvey, S. D., Heathers, L. E., Xu, H., Sherman, S., & Korc, M. (2014). A pilot study to develop a diagnostic test for pancreatic ductal adenocarcinoma based on differential expression of select miRNA in plasma and bile. American Journal of Gastroenterology, 109(12), 1942–1952.

Garcea, G., Cairns, V., Berry, D. P., Neal, C. P., Metcalfe, M. S., & Dennison, A. R. (2012). Improving the diagnostic yield from staging laparoscopy for periampullary malignancies: The value of pre-operative inflammatory markers and radiological tumor size. Pancreas, 41(2), 233–237.
detection of familial pancreatic cancer. *Translational Oncology*, 6(2), 99–103.

Somiari, S. B., Somiari, R. I., Heckman, C. M., Olsen, C. H., Jordan, R. M., Russell, S. J., & Shriver, C. D. (2006). Circulating MMP2 and MMP9 in breast cancer—potential role in classification of patients into low risk, high risk, benign disease and breast cancer categories. *International Journal of Cancer*, 119(6), 1403–1411.

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, 45(W1), W98–W102.

Tekin, C., Aberson, H. L., Waasdorp, C., Hooijer, G. K. J., de Boer, O. J., Dijk, F., Bijlsma, M. F., & Spek, C. A. (2020). Macrophage-secreted MMP9 induces mesenchymal transition in pancreatic cancer cells via PAR1 activation. *Cellular Oncology*, 43(6), 1161–74.

Twine, N. C., Stover, J. A., Marshall, B., Dukart, G., Hidalgo, M., Stadler, W., Logan, T., Dutcher, J., Hudes, G., Dorner, A. J., Slonim, D. K., Trepicchio, W. L., & Burczynski, M. E. (2003). Disease-associated expression profiles in peripheral blood mononuclear cells from patients with advanced renal cell carcinoma. *Cancer Research*, 63(18), 6069–6075.

Vareedayah, A. A., Alkaade, S., & Taylor, J. R. (2018). Pancreatic adenocarcinoma. *Missouri Medicine*, 115(3), 230–235.

Weizman, N., Krelin, Y., Shabtay-Orbach, A., Amit, M., Binenbaum, Y., Wong, R. J., & Gil, Z. (2014). Macrophages mediate gemcitabine resistance of pancreatic adenocarcinoma by upregulating cytidine deaminase. *Oncogene*, 33(29), 3812–3819.

Whitney, A. R., Diehn, M., Popper, S. J., Alizadeh, A. A., Boldrick, J. C., Relman, D. A., & Brown, P. O. (2003). Individuality and variation in gene expression patterns in human blood. *Proceedings of the National Academy of Sciences*, 100(4), 1896–1901.

Wingren, C., Sandström, A., Segersvärd, R., Carlsson, A., Andersson, R., Löhr, M., & Borrebaeck, C. A. K. (2012). Identification of serum biomarker signatures associated with pancreatic cancer. *Cancer Research*, 72(10), 2481–2490.

Xie, G., Lu, L., Qiu, Y., Ni, Q., Zhang, W., Gao, Y.-T., Risch, H. A., Yu, H., & Jia, W. (2015). Plasma metabolite biomarkers for the detection of pancreatic cancer. *Journal of Proteome Research*, 14(2), 1195–1202.

Yadav, D., & Lowenfels, A. B. (2013). The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology*, 144(6), 1252–1261.

Zhou, Q., Hou, Z., Zuo, S., Zhou, X., Feng, Y., Sun, Y., & Yuan, X. (2019). LUCAT1 promotes colorectal cancer tumorigenesis by targeting the ribosomal protein L40-MDM2-p53 pathway through binding with UBA 52. *Cancer Science*, 110(4), 1194–207.

Zhu, M., Zhao, B., Wei, L., & Wang, S. (2021). alpha-2-macroglobulin, a native and powerful proteinase inhibitor, prevents cartilage degeneration disease by inhibiting majority of catabolic enzymes and cytokines. *Current Molecular Biology Reports*, 7, 1–7.

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---

How to cite this article: Kalantari, S., Kazemi, B., Roudi, R., Zali, H., D’Angelo, A., Mohamadkhani, A., Madjd, Z., & Pourshams, A. (2023). RNA-sequencing for transcriptional profiling of whole blood in early stage and metastatic pancreatic cancer patients. *Cell Biology International*, 47, 238–249. https://doi.org/10.1002/cbin.11924