HHS Public Access

Author manuscript

*Clin Cancer Res*. Author manuscript; available in PMC 2024 March 15.

Published in final edited form as:

*Clin Cancer Res*. 2023 September 15; 29(18): 3759–3770. doi:10.1158/1078-0432.CCR-23-0825.

# A Novel HOXA10-Associated 5 Gene-Based Prognostic Signature for Stratification of Short-Term Survivors of Pancreatic Ductal Adenocarcinoma

**Sophia G. Kisling**[1], **Pranita Atri**[1], **Ashu Shah**[1,*], **Jesse L. Cox**[2], **Sunandini Sharma**[2], **Lynette M. Smith**[3], **Dario Ghersi**[4], **Surinder K. Batra**[1,5,6,*]

[1]Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, NE, 68198, USA

[2]Department of Pathology and Microbiology, University of Nebraska Medical Center, NE, 68198, USA

[3]Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, NE, 68198, USA

[4]School of Interdisciplinary Informatics, College of Information Science & Technology, University of Nebraska Omaha, NE, 68182, USA

[5]Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, NE, 68198, USA

[6]Fred and Pamela Buffett Cancer Center, University of Nebraska Medical Center, NE, 68198, USA

## Abstract

**Purpose:** Despite the significant association of molecular subtypes with poor prognosis in pancreatic ductal adenocarcinoma (PDAC) patients, little effort has been made to identify the underlying pathway(s) responsible for this prognosis. Identifying a clinically relevant prognosis-based gene signature may be the key to improving patient outcomes.

**Experimental Design:** We analyzed the transcriptomic profiles of treatment-naïve surgically resected short- and long-term survivor tumors (GSE62452) for expression and survival, followed by validation in several datasets. These results were corroborated by immunohistochemical analysis of PDAC-resected short- and long-term survivor tumors. The mechanism of this differential survival was investigated using CIBERSORT and pathway analyses.

**Results:** We identified a short-surviving prognostic subtype of PDAC with a high degree of significance (p=0.018). One hundred thirty genes in this novel subtype were found to be regulated by a master regulator, HOXA10, and a five-gene signature derived from these genes, including BANF1, EIF4G1, MRPS10, PDIA4, and TYMS, exhibited differential expression in short-term

[*]**Correspondence:** Surinder K. Batra and Ashu Shah, Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, Omaha, Nebraska, 68198-5870, U.S.A. Tel: 402-559-5455, Fax: 402-559-6650, sbatra@unmc.edu; ashu.shah@unmc.edu.

**Conflicts of Interest**: The other authors declare no competing interests.

survivors (STS) and a strong association with poor survival. This signature was further associated with the proportion of T-cells and macrophages found in STS and long-term survivors (LTS), demonstrating a potential role in PDAC immunosuppression. Pathway analyses corroborated these findings, revealing that this HOXA10-driven signature is associated with immune suppression and enhanced tumorigenesis.

**Conclusions:** Overall, these findings reveal the presence of a HOXA10-associated prognostic subtype that can be used to differentiate between STS and LTS patients of PDAC and inform on the molecular interactions that play a role in this poor prognosis.

### Keywords

HOXA10; bioinformatics; signature; prognosis; pancreatic ductal adenocarcinoma

## Introduction

Cancer mortality is currently the second leading cause of death across the globe and is predicted to rise to the primary cause of death over the next few decades (1). Among the cancers whose projected deaths will increase dramatically by 2030, pancreatic ductal adenocarcinoma (PDAC) has the lowest five-year survival rate. Up 3% since 2014, this recalcitrant cancer has a stage-combined five-year survival rate of 11%. Moreover, PDAC is currently the third leading cause of cancer deaths, accounting for approximately 8% of all cancer deaths, and will be second in rank for deaths by 2030 (2–4).

For many PDAC patients, poor survival estimates are a consequence of late diagnosis. Due to the high occurrence of asymptomatic early stages, it is common that when patients are diagnosed, cancer has already metastasized to distant organs. At this stage in the disease, the five-year survival rate is about 3%, and treatment options are severely limited (3). Unfortunately, the challenges associated with the difficulties of successful treatment, such as asymptomatic early stages and tumor heterogeneity, further augment poor survival.

A large portion of PDAC research is dedicated to the identification of biomarkers for early diagnosis and subtyping (5,6). This has helped achieve some success in improving diagnosis and treatment strategies but with limited outcomes due to their inability to reliably inform on patient survival. Most notably, CA19–9 has been widely studied as a PDAC biomarker and, more recently, as a prognostic tool, but its efficacy is highly debated due to its startling lack of specificity and consistency (7,8). The Collisson et al. (9), Moffitt et al. (10), Bailey et al. (11), and Puleo et al. (12) PDAC classifications are also limited in that they offer alternative targets for therapeutic targeting but otherwise do not directly influence patient survival.

Due to the disparity observed in PDAC patient survival, there has been gathering interest in identifying prognostic markers for PDAC that play a role in survival. In addition to genes that are known to play a prominent role in developing PDAC, including mutated KRAS, TP53, SMAD4, CDKN2A, and ATM (13,14), over 50 genes have been correlated with poor prognoses in PDAC (15–20). Interestingly, some of these prognostically-relevant markers have been derived from subtypes of PDAC with differential survival, but no prognostic

markers of PDAC have been validated in clinical trials. Therefore, there is a need for reliable prognostic markers of PDAC.

In order to define these markers, there needs to be further investigation into subtyping on the basis of survival. By identifying biomarkers associated with short- and long-term PDAC survivors, we may uncover underlying molecules and pathways that are correlated with the mechanism of PDAC progression and metastasis. Consequently, the identification of these prognostic markers could lead to the development of a treatment modality to improve overall patient survival.

Since the significance of these short- and long-term surviving groups in PDAC is highly under-studied, it seemed relevant to ask if [1] there is any substantial difference in genes enriched in short-term survivors compared to long-term survivors, [2] we can identify a prognostic signature for PDAC using these differentially expressed genes (DEGs), and [3] we can identify pathways associated with enriched genes in short-term survivors. Using this approach, we aimed to define a prognostic signature that may provide insight into the underlying biology of the poor disease prognosis.

A key feature associated with tumor aggression and poor patient survival in PDAC, which has gained particular attention in the past few years, is the presence of an immunosuppressive tumor microenvironment. HOXA10, a hub gene that is strongly associated with survival, is also established as a key regulator of the immune suppression and stromal proliferation required for successful implantation in normal endometrium, as found in a study by Yao et al. using a Hoxa10-mutant mouse model (21). Further, HOXA10 is involved in multiple cellular processes, including proliferation, epithelial-mesenchymal transition (EMT), and chemotherapy resistance in various cancers (22), demonstrating its potential on multiple fronts.

Herein, we identified a HOXA10-driven prognostic signature that can be used to distinguish STS and LTS of PDAC and aimed to determine their associated pathways in the hope that it will lead to improved management of PDAC by characterizing novel pathways for therapeutic targeting.

## Materials and Methods

### Data collection

Gene Expression Omnibus (GEO; RRID:SCR_005012) (23,24) was perused for PDAC datasets using “homo sapiens”, “pancreatic cancer”, “tissue” and “expression array profiling by array” as filters. A second filter, the availability of survival information, helped us choose the dataset GSE62452 (25). This dataset comprised the gene expression profiles of 69 pancreatic tumor tissues and 61 normal adjacent tumor tissues from pancreatic ductal adenocarcinoma patients, respectively. Datasets GSE89997 (26), GSE15471 (27), GSE16515 (28), and GSE32676 (29) were also downloaded to validate the DEGs and hub genes.

### Data preprocessing and identification of enriched genes

The raw microarray data (.CEL files) from GSE62452 was downloaded and processed using the R package affy (RRID:SCR_012835) (30). The STS and LTS groups were determined by division from a median survival of 14.5 months; the overall survival range for STS was defined as 0.9–14.5 months, and the range for LTS was 14.5–70.8 months. The gene expression data was then analyzed using gene set enrichment analysis (GSEA; RRID:SCR_003199) with a two-class division of long-term and short-term survivors. A false discovery rate (FDR) q-value cutoff of 0.25 was utilized in GSEA to allow for the exploratory discovery of potentially significant datasets. The GSEA-enriched gene sets for STS were validated independently using graphical illustrations of microarray data from GSE89997, which includes extreme LTS and STS of PDAC post-surgery.

### Pathway enrichment analysis and PPI networks of enriched genes

To further explore key pathways associated with prognosis, Kyoto Encyclopedia of Genes and Genomes (KEGG; RRID:SCR_012773) pathway enrichment analysis (31,32), WikiPathways (RRID:SCR_002134) pathway enrichment analysis (33), and Gene Ontology (GO; RRID:SCR_002811) enrichment analysis (34,35) were utilized. These pathways were validated using National Center for Advancing Translational Sciences (NCATS; RRID:SCR_012857) BioPlanet pathway analysis web tool. Similarly, the ClueGO plugin in Cytoscape (RRID:SCR_005748) (36,37) and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; RRID:SCR_005223) (38,39) were used to visualize the functional grouping and interaction networks of the protein products of these enriched genes. Genes specifically upregulated in short- or long-term survivors were analyzed independently.

### Validation through gene expression profiling

Gene association with overall survival and differential expression in PDAC was validated using survival analysis and expression analysis from the Gene Expression Profiling Interactive Analysis (GEPIA; RRID:SCR_018294) database (40,41), which utilizes The Cancer Genome Atlas (TCGA; tumor, N=179) and The Genotype-Tissue Expression (GTEx; normal, N=171) project data. Top 25$^{th}$ and bottom 25$^{th}$ quartiles were used as an expression threshold for survival analyses for each gene and the combined signature group. The gene signature was determined using graphical representations of the expression patterns for each STS-enriched gene in pancreatic tumor tissue compared to matched normal tissue in PDAC GEO datasets GSE15471, GSE16515, and GSE32676. Genes that had increased levels of expression in tumor tissues in each dataset were selected for further validation at the protein level in pancreatic tumor tissues using immunohistochemistry (IHC) data from the Human Protein Atlas (RRID:SCR_006710).

### CIBERSORT analysis

Data from GSE62452 was subjected to an *in silico* deconvolution by using cell type identification by estimating the relative subset of RNA transcripts (CIBERSORT; RRID:SCR_016955) and the LM22 signature matrix (42). The 22 immune cell types were then compared between high expression (top 25$^{th}$ percentile) and low expression (bottom

25th percentile) of HOXA10, EIF4G1, BANF1, PDIA4, TYMS, and MRPS10. These high and low expressing divisions reflected short-term and long-term surviving groups, respectively.

## Generation of HOXA10-specific antibodies

To generate HOXA10-specific antibodies, a peptide derived from human HOXA10 was selected based on antigenicity and immunogenicity scores predicated by Geneious (RRID:SCR_010519) and IEDB database (RRID:SCR_006604) analysis. The selected peptide (109–126aa) was sent for synthesis and rabbit polyclonal antisera generation (RS Synthesis). Briefly, the peptide was conjugated to Keyhole Limpet Hemocyanin (KLH), and routine immunization protocol was followed for the development of polyclonal antisera (43). The HOXA10-specific antibodies were purified using protein G based affinity chromatography and adsorption on KLH. These purified antibodies were run on a Western blot to verify the presence of the heavy and light chain domains and validated in IHC of PDAC tumor tissues (Supplementary Fig. S1).

## Immunohistochemical analysis

Immunohistochemical analyses were performed on paraffin-embedded STS and LTS PDAC tissue slides provided by the Rapid Autopsy Program at the University of Nebraska Medical Center. The slides were baked overnight at 58 °C and cooled to room temperature prior to paraffin removal using xylene washes (2 × 5 minutes). The slides were then rehydrated using an alcohol gradient (100, 90, 70, 50, 30, and 20%) for 5 minutes at each level, followed by a rinse with water to remove residual alcohol. Quenching of endogenous peroxidases was then performed by incubating the slides in a solution of 3% $H_2O_2$ in methanol in the dark for 1 hour. The slides were subsequently rinsed with water to remove residual peroxidase. Further, antigen retrieval was performed in 0.01M citrate buffer with 0.05% Tween 20 for 10 minutes using the microwave method. The slides were allowed to cool to room temperature, rinsed with water to remove residual citrate buffer, and then blocking was done using 2.5% horse serum (Impress Reagent Kit, Vector Laboratories; RRID:SCR_000821). The tissues were incubated with primary antibodies for HOXA10 (3ug/mL), BANF1 (2.5ug/mL; LSBio (LifeSpan) Cat# LS-B447; RRID:AB_2274440), TYMS (1:200; Thermo Fisher Scientific Cat# MA5–13512; RRID:AB_11004511), EIF4G1 (1:120; Sigma-Aldrich Cat# HPA028487; RRID:AB_10602219), MRPS10 (1:100; Sigma-Aldrich Cat# HPA029134; RRID:AB_10600058), and PDIA4 (1:200; Sigma-Aldrich Cat# HPA006140; RRID:AB_1848257) at 4 °C overnight. Following this, the tissue slides were washed with PBST (3 × 5 minutes) and PBS (1 × 5 minutes) and incubated with HRP-labelled universal secondary anti-mouse/rabbit IgG for 45 minutes. The slides were again washed with PBST (3 × 5 minutes) and then stained with DAB substrate kit (Vector Laboratories; RRID:SCR_000821) before counterstaining with hematoxylin, graded alcohol dehydration, and xylene washes. After drying, the slides were mounted with Permount mounting medium (Thermo Fisher Scientific; RRID:SCR_008452) and a coverslip. Blind slide scoring, which consisted of an intensity score (0–3; 0-negative, 1-weak, 2-moderate, 3-intense staining) and the percentage of positive cells (0–100%) for each marker, was performed by a pathologist. An H-score (intensity score x percentage of positive cells) for each marker in STS and LTS was calculated from these scores.

### Statistics

Data were analyzed using GraphPad Prism 9 (RRID:SCR_002798). Statistical analysis of CIBERSORT data for each marker was performed using the Mann-Whitney test. For IHC expression analysis, the F-test for sample variances was used for each marker, followed by the Mann-Whitney test, excluding any statistical outliers and samples of poor quality. Kaplan-Meier method was used to estimate overall survival distributions by low and high gene expression, categorized at the top $25^{th}$ and bottom $25^{th}$ quartiles and the logrank test was used to compare survival distributions between groups. A $p<0.05$ or $q<0.05$ was considered to indicate significance for each analysis.

### Data Availability

The expression profile data analyzed in this study were obtained from GEO at GSE62452, GSE89997, GSE15471, GSE16515, and GSE32676.

## Results

### Identification of hub genes associated with prognosis

Based on the availability of patient survival information in dataset GSE62452, we processed the raw data and divided the population into LTS and STS, as depicted in Fig. 1. GSEA was performed, and we identified several gene sets which are enriched in STS (Supplementary Table S1). The protein counterparts of the hub genes of these sets were identified as proteins that control cell cycle progression, interact with cell cycle regulators, mediate transcription, translation, and DNA replication, and contribute to cancer initiation and progression. However, two of these hub genes, MYBL2 and HOXA10, were especially noteworthy due to their involvement in cell differentiation and immune response. The gene set driven by MYBL2 (ES=0.59; p=0.002; q=0.206) and comprised of 177 genes (Fig. 2A), was distinctive due to the essential role of MYLB2 in hematopoietic stem cell and myeloid progenitor cell development (44). Similarly, the gene set driven by HOXA10 (ES=0.49; p=0.018; q=0.239) and comprised of 130 genes (Fig. 2B, Supplementary Table S2) was unique in that HOXA10 is associated with stromal cell proliferation and local immunosuppression, specifically as it relates to T lymphocytes (21). Further, our initial observations of HOXA10 expression and survival using GEPIA supported its increased expression in PDAC compared to the normal pancreas (Fig. 2C) and its association with poor survival (Fig. 2D). Since the immunosuppressive nature of the pancreatic tumor microenvironment is a key prognostic factor, this gene set was selected for future validation and exploration. Of these 130 HOXA10-driven genes, GSEA revealed 69 genes were enriched in STS (Supplementary Table S3). To validate this gene set, we performed a preliminary evaluation of their differential expression and association with overall survival in PDAC using their GEPIA profiles. Of the 69 enriched genes, 46 had significantly higher expression in cancer compared to the normal ($q<0.05$). Thirty-three of these upregulated genes were associated with short-term survival, and 11 were significant ($q<0.05$) (Fig. 1).

## Pathway analysis in short-term survivors

To investigate the biological pathways associated with the HOXA10 prognostic signature, we performed pathway analysis using different *in silico* tools (Fig. 1). The consensus of these analyses revealed the enrichment of genes and pathways involved in cell cycle regulation, protein homeostasis, immune response, and cellular maintenance in short-term survivors. Specifically, GO enrichment analysis revealed that the HOXA10-driven gene set was enriched in the cellular response to interleukin-1 and its mediated signaling pathway (gene ratio(s)≈0.0975, 0.0725; $q<0.005$), regulation of G2/M cell cycle transitions (gene ratio≈0.085; $q<0.05$), and antigen processing and presentation of the antigen (gene ratio≈0.0725; $q<0.05$), among others (Fig. 3A). Furthermore, we observed the enrichment of these biological pathways in the functional clusters derived from subsequent ClueGO and STRING analyses. Namely, pathways associated with protein homeostasis (ribosomal large subunit biogenesis, IRE1-mediated unfolded protein response), cellular metabolism (hexose catabolic process, negative regulation of RNA catabolic process, pyrimidine deoxyribonucleoside triphosphate metabolic process, regulation of RNA stability) and immune response (cellular response to interferon-beta) were associated with HOXA10 gene set suggesting its association with tumor progression and aggressiveness (Fig. 3B).

## Development and validation of the HOXA10-driven genes as a prognostic signature

To establish a subset of genes representative of the HOXA10 gene set, we evaluated their differential expression based on the normal pancreas and PDAC microarray data using several external datasets. From graphical illustrations of this data, we determined that 40 of the 69 enriched genes had increased expression in cancer cases compared to the normal pancreas (Supplementary Table S4). We further evaluated these genes and shortlisted the ones that contained a signal peptide, did not have any transmembrane domains, and had an association with poor survival (Fig. 1). Combined with the validation data from GEPIA, we were able to identify five enriched genes that were determined to have differential expression and association with survival for further investigation: BANF1, EIF4G1, MRPS10, PDIA4, and TYMS (Fig. 1). Each gene was upregulated in PDAC compared to the normal pancreas, and their expression was associated with decreased survival in PDAC; EIF4G1 ($p=0.0013$), MRPS10 ($p=0.013$), PDIA4 ($p=0.083$), TYMS ($p=0.00064$), and BANF1 ($p=0.053$) (Fig. 4A–E). These findings were corroborated by IHC of PDAC tissues from HPA (Supplementary Fig. S2), together demonstrating that these genes are expressed at both the RNA and tissue levels. An association with decreased survival was also observed for the newly identified prognostic genes using a Kaplan-Meier plot and log-rank test ($p=0.00076$) (Fig. 4F).

The HOXA10 and the five representative genes from the HOXA10-assocaited signature were validated using IHC of STS and LTS PDAC tissues. The expression of HOXA10 was significantly increased ($p=0.038$) in STS tissues compared to LTS tissues (Fig. 5A). Similarly, a trend of high expression for each of the five genes was observed in STS tissues, supporting the *in silico* findings: BANF1 (STS median=225; LTS median=160; $p=0.024$); EIF4G1 (STS median= 167.5; LTS median= 97.5; $p=0.045$); MRPS10 (STS median= 45; LTS median=57.5; $p=0.895$); PDIA4 (STS median= 65; LTS median= 35; $p=0.568$); TYMS (STS median=0; LTS median=12.6; $p=0.340$) (Fig. 5B–F).

### Immune cell enrichment analysis in HOXA10-driven gene set

Since we observed a significant correlation between HOXA10 expression and short-term survival of patients, we sought to investigate the associations between our HOXA10 gene set and immunosuppressive cell types. Our estimations using CIBERSORT and the LM22 signature matrix revealed a greater proportion of $T_{reg}$ cells and M2 macrophages, both immunosuppressive cell types, in patients with high expression of these HOXA10-driven genes (Fig. 6A–C, Supplementary Fig. S3A–C). Specifically, $T_{reg}$ cell signatures were increased in groups highly expressing HOXA10, PDIA4, EIF4G1, and BANF1 (Fig. 6A), while M2 macrophages were elevated in groups with high expression of MRPS10, EIF4G1, and TYMS (Fig. 6B). Conversely, $CD8^+$ T cells were proportionally lower in groups highly expressing MRPS10, PDIA4, and TYMS than in groups with low expression of these genes (Fig. 6C). Curiously, $CD8^+$ T cells were elevated in groups highly expressing HOXA10 compared to groups with low expression of HOXA10 (Fig. 6C). Upon further analysis, we found significantly higher expression of co-inhibitory receptors PD-1 ($p<0.0001$) and LAG-3 ($p<0.0001$), associated with T cell exhaustion, in the high HOXA10 expressing patients (Supplementary Fig. S4). In corroboration with our findings, previous studies have implicated HOXA10 in immune response in multiple physiological milieu (21,45). Taken together, these results indicate HOXA10 may play a crucial role in differential immune modulation and thus impact PDAC patient survival.

## Discussion

The lack of clinically relevant prognostic subtypes in PDAC is a widely recognized concern. Direct prognostication from these subtypes would allow for greater accuracy and eliminate any issues of consensus, as patient prognoses are currently inferred from existing molecular and histological subtypes, which do not always correspond with each other (46). Moreover, unlike molecular and histological subtypes, prognostic subtypes may offer an effective way to inform treatment strategies (47). Chiefly, using a prognostic signature to classify patients as STS or LTS will augment risk stratification and clinical decision-making after identifying a tumor, which is currently limited by the TNM staging system. Clinicians use the TNM staging system to report the size and spread of a tumor, which has minimal impact on the treatment plan and does not inform on individual patient prognoses (48). In PDAC, the combinatorial use of a prognostic signature and TNM staging would allow for further stratification of patients with either localized or metastatic disease into subgroups predisposed to long-term or short-term survival. With this knowledge, clinicians may design a more personalized treatment plan tailored to expected survival to preserve quality of life.

In this study, we analyzed the transcriptomics of short- and long-term survivor data of PDAC to reveal prognostically significant genes. Our GSEA analysis on a dataset of LTS and STS resulted in the enrichment of a HOXA10-driven cluster of 130 genes in STS, out of which 69 genes are associated with key biological pathways of cancer, including cell cycle regulation, protein homeostasis, DNA binding, and repair, and cytoskeletal maintenance. Specifically, these pathways are implicated in a cancer cell’s high level of proliferation, reduced apoptosis, altered DNA damage response, and immune evasion, among other factors, all contributing to the development and progression of cancer (49–52).

This highlights their potential as therapeutic targets and indicates that poor survival and the hallmarks of cancer may be deeply interconnected (50,53).

Five of the genes associated with these pathways, including BANF1, EIF4G1, MRPS10, PDIA4, and TYMS, stood out due to their strong association with survival. These genes have been implicated in several types of cancer, and several studies have also observed their prognostic potential (54–58). In fact, PDIA4 has previously been reported to predict reduced survival in cervical cancer (59). Furthermore, the driver gene of this 5-gene signature, HOXA10, is correlated with poor survival in multiple cancer types (60–63), and has recently been shown to play a role in PDAC development (64). The findings mentioned above are further validated by our immunohistochemistry staining and combined survival analysis. Our study emphasizes the importance of studying the molecular mechanisms of these genes and their association with survival in PDAC.

To further understand the biological significance of short-term survival, we performed pathway analyses of short-term survivor genes and identified several associated pathways. Most notably, high expression of interleukin-1 is associated with poor prognosis in multiple cancer types, and its upregulated expression is associated with immune suppression in PDAC (65,66). Similarly, inflammation induced by interleukin-1 is associated with carcinogenesis (67). The enrichment of these genes in metabolic pathways, homeostasis, and cell cycle regulation also suggests a potential association with chemoresistance and the development of cancer (68,69). In fact, the correlation of the HOXA10 signature with these biological pathways partly explains its association with the short-term survival of patients. This association is reinforced by the discovery of HOXA10-mediated immunosuppression by Yao et al. (21), as this supports the supposition that HOXA10 may play a key role in the pathogenesis of PDAC, which often lays claim to a highly immunosuppressive tumor stroma.

As immune response plays an important role in PDAC patient survival (70), it is pertinent to decipher the association of this newly identified signature with the immune profiles of PDAC patients. Interestingly, the expression of these HOXA10-driven genes were associated with several immunosuppressive cell types, including $T_{regs}$ and M2 macrophages, indicating the presence of an enhanced immunosuppressive tumor microenvironment in STS compared to LTS. This is supported by a recent meta-analysis that discovered an association between high levels of infiltrating M2 macrophages and poor survival in PDAC (71), and by a study that revealed that increased levels of peripheral $T_{regs}$ in PDAC is associated with poor patient survival (72). Though $CD8^+$ T cells seemed to be high in HOXA10-expressing patients, there have always been arguments on the quality versus quantity of T cells in cancer patients (26,73). Of note, the high expression level of PD-1 and LAG-3 in HOXA10 high patients hints towards immunosuppression and thus poor survival of those patients. These *in silico* findings are corroborated by previous studies that have also linked our HOXA10-driven genes with immune responses (74–77), particularly with suppressive roles (78,79). Further, recent research indicates that immune signatures may have prognostic value in PDAC (80), supporting the significance of these genes.

The present study addresses the identification of a novel signature and its master regulator, which are associated with the pathogenesis of PDAC. To the best of our knowledge, this is the first attempt to elucidate the downstream effects of HOXA10, specifically as they relate to PDAC prognosis. This study details the design and validation of a 5-gene HOXA10-based prognostic signature using multiple state-of-art computational tools. The use of GSEA in combination with multi-database pathway analyses illustrates the robustness of the study and demonstrates a thorough examination of the significance of this signature. The innovation of this study lies in the experimentally validated signature on STS and LTS patient tumors, whose results were correlative. Further, the study has employed multiple *in silico* tools, including different GSE datasets and pathway analyses to validate the signature. However, further investigation is needed to determine the clinical utility of this 5-gene signature, including its validation using a larger cohort size with associated patient information. In addition, biological assessment of this signature using NanoString technologies and *in vitro* and *in vivo* models should be conducted to supplement our computational findings. Once we understand the biological significance of this signature and its driver gene, we may develop a clinically relevant HOXA10-based biomarker panel that will improve on the standard clinical staging parameters and identify pathways that can be targeted for therapeutic intervention in patients with unfavorable prognoses.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments:

The authors of this work were supported, in parts, by grants from the National Institute of Health (S.K. Batra; P01 CA217798, U01 CA210240, U01 CA200466, R01 CA228524, R01 CA247471, R01 CA254036, R01 CA256973, R01 CA263575, R01 CA273349, and R01 CA273319) and the University of Nebraska Medical Center Assistantship/Fellowship (S.G. Kisling.). We thank Dr. Sukhwinder Kaur and Dr. Sushil Kumar for providing valuable suggestions.

S.K. Batra is one of the co-founders of Sanguine Diagnostics and Therapeutics, Inc.

## References

1. Dagenais GR, Leong DP, Rangarajan S, Lanas F, Lopez-Jaramillo P, Gupta R, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. Lancet (London, England) 2020;395:785–94 [PubMed: 31492501]
2. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res 2014;74:2913–21 [PubMed: 24840647]
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022;72:7–33 [PubMed: 35020204]
4. Rawla P, Sunkara T, Gaduputi V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World J Oncol 2019;10:10–27 [PubMed: 30834048]
5. Yang M, Zhang CY. Diagnostic biomarkers for pancreatic cancer: An update. World J Gastroenterol 2021;27:7862–5 [PubMed: 34963749]
6. Xu Z, Hu K, Bailey P, Springfeld C, Roth S, Kurilov R, et al. Clinical Impact of Molecular Subtyping of Pancreatic Cancer. Front Cell Dev Biol 2021;9:743908

7. Yang J, Xu R, Wang C, Qiu J, Ren B, You L. Early screening and diagnosis strategies of pancreatic cancer: a comprehensive review. Cancer Commun (Lond) 2021;41:1257–74 [PubMed: 34331845]
8. Lautizi M, Baumbach J, Weichert W, Steiger K, List M, Pfarr N, et al. The limits of molecular signatures for pancreatic ductal adenocarcinoma subtyping. NAR Cancer 2022;4:zcac030
9. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med 2011;17:500–3 [PubMed: 21460848]
10. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet 2015;47:1168–78 [PubMed: 26343385]
11. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 2016;531:47–52 [PubMed: 26909576]
12. Puleo F, Nicolle R, Blum Y, Cros J, Marisa L, Demetter P, et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. Gastroenterology 2018;155:1999–2013 e3 [PubMed: 30165049]
13. Buscail L, Bournet B, Cordelier P. Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. Nat Rev Gastroenterol Hepatol 2020;17:153–68 [PubMed: 32005945]
14. Kim H, Saka B, Knight S, Borges M, Childs E, Klein A, et al. Having pancreatic cancer with tumoral loss of ATM and normal TP53 protein expression is associated with a poorer prognosis. Clin Cancer Res 2014;20:1865–72 [PubMed: 24486587]
15. Blackford A, Serrano OK, Wolfgang CL, Parmigiani G, Jones S, Zhang X, et al. SMAD4 gene mutations are associated with poor prognosis in pancreatic cancer. Clin Cancer Res 2009;15:4674–9 [PubMed: 19584151]
16. Zeng Y, Fan R. Identification and verification of CCNB1 as a potential prognostic biomarker by comprehensive analysis. Scientific reports 2022;12:16153 [PubMed: 36167975]
17. Pedersen RS, Nissen NI, Jensen C, Thorlacius-Ussing J, Manon-Jensen T, Olesen ML, et al. Plasma Kallikrein-Activated TGF-beta Is Prognostic for Poor Overall Survival in Patients with Pancreatic Ductal Adenocarcinoma and Associates with Increased Fibrogenesis. Biomolecules 2022;12 [PubMed: 36671397]
18. Xu X, Liang JH, Li JH, Xu QC, Yin XY. Values of a novel pyroptosis-related genetic signature in predicting outcome and immune status of pancreatic ductal adenocarcinoma. Gastroenterol Rep (Oxf) 2022;10:goac051
19. Kisling SG, Natarajan G, Pothuraju R, Shah A, Batra SK, Kaur S. Implications of prognosis-associated genes in pancreatic tumor metastasis: lessons from global studies in bioinformatics. Cancer Metastasis Rev 2021;40:721–38 [PubMed: 34591244]
20. Song C, Chen T, He L, Ma N, Li JA, Rong YF, et al. PRMT1 promotes pancreatic cancer growth and predicts poor prognosis. Cell Oncol (Dordr) 2020;43:51–62 [PubMed: 31520395]
21. Yao MW, Lim H, Schust DJ, Choe SE, Farago A, Ding Y, et al. Gene expression profiling reveals progesterone-mediated cell cycle and immunoregulatory roles of Hoxa-10 in the preimplantation uterus. Mol Endocrinol 2003;17:610–27 [PubMed: 12554760]
22. Yoshida H, Broaddus R, Cheng W, Xie S, Naora H. Deregulation of the HOXA10 homeobox gene in endometrial carcinoma: role in epithelial-mesenchymal transition. Cancer Res 2006;66:889–97 [PubMed: 16424022]
23. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013;41:D991–5 [PubMed: 23193258]
24. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10 [PubMed: 11752295]
25. Yang S, He P, Wang J, Schetter A, Tang W, Funamizu N, et al. A Novel MIF Signaling Pathway Drives the Malignant Character of Pancreatic Cancer by Targeting NR3C2. Cancer Res 2016;76:3838–50 [PubMed: 27197190]

26. Balachandran VP, Luksza M, Zhao JN, Makarov V, Moral JA, Remark R, et al. Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. Nature 2017;551:512–6 [PubMed: 29132146]
27. Idichi T, Seki N, Kurahara H, Yonemori K, Osako Y, Arai T, et al. Regulation of actin-binding protein ANLN by antitumor miR-217 inhibits cancer cell aggressiveness in pancreatic ductal adenocarcinoma. Oncotarget 2017;8:53180–93 [PubMed: 28881803]
28. Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, Lingle W, et al. FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. Cancer Cell 2009;16:259–66 [PubMed: 19732725]
29. Donahue TR, Tran LM, Hill R, Li Y, Kovochich A, Calvopina JH, et al. Integrative survival-based molecular profiling of human pancreatic cancer. Clin Cancer Res 2012;18:1352–63 [PubMed: 22261810]
30. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 2004;20:307–15 [PubMed: 14960456]
31. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30 [PubMed: 10592173]
32. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 1999;27:29–34 [PubMed: 9847135]
33. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. Nucleic Acids Res 2021;49:D613–D21 [PubMed: 33211851]
34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9 [PubMed: 10802651]
35. The Gene Ontology C Expansion of the Gene Ontology knowledgebase and resources. Nucleic Acids Res 2017;45:D331–D8 [PubMed: 27899567]
36. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091–3 [PubMed: 19237447]
37. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504 [PubMed: 14597658]
38. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021;49:D605–D12 [PubMed: 33237311]
39. Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res 2000;28:3442–4 [PubMed: 10982861]
40. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98–W102 [PubMed: 28407145]
41. Li C, Tang Z, Zhang W, Ye Z, Liu F. GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA. Nucleic Acids Res 2021;49:W242–W6 [PubMed: 34050758]
42. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453–7 [PubMed: 25822800]
43. Lee BS, Huang JS, Jayathilaka GD, Lateef SS, Gupta S. Production of antipeptide antibodies. Methods Mol Biol 2010;657:93–108 [PubMed: 20602209]
44. Shepard JL, Amatruda JF, Stern HM, Subramanian A, Finkelstein D, Ziai J, et al. A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. Proc Natl Acad Sci U S A 2005;102:13194–9 [PubMed: 16150706]
45. Wang T, Liu M, Jia M. Integrated Bioinformatic Analysis of the Correlation of HOXA10 Expression with Survival and Immune Cell Infiltration in Lower Grade Glioma. Biochem Genet 2023;61:238–57 [PubMed: 35836029]

46. Martens S, Lefesvre P, Nicolle R, Biankin AV, Puleo F, Van Laethem JL, et al. Different shades of pancreatic ductal adenocarcinoma, different paths towards precision therapeutic applications. Ann Oncol 2019;30:1428–36 [PubMed: 31161208]
47. Collisson EA, Bailey P, Chang DK, Biankin AV. Molecular subtypes of pancreatic cancer. Nat Rev Gastroenterol Hepatol 2019;16:207–20 [PubMed: 30718832]
48. O'Sullivan B, Brierley J, Byrd D, Bosman F, Kehoe S, Kossary C, et al. The TNM classification of malignant tumours-towards common understanding and reasonable expectations. Lancet Oncol 2017;18:849–51 [PubMed: 28677562]
49. Campos A, Clemente-Blanco A. Cell Cycle and DNA Repair Regulation in the Damage Response: Protein Phosphatases Take Over the Reins. International journal of molecular sciences 2020;21
50. Yip HYK, Papa A. Signaling Pathways in Cancer: Therapeutic Targets, Combinatorial Treatments, and New Developments. Cells 2021;10
51. Aseervatham J Cytoskeletal Remodeling in Cancer. Biology (Basel) 2020;9
52. Bastola P, Leiserowitz GS, Chien J. Multiple Components of Protein Homeostasis Pathway Can Be Targeted to Produce Drug Synergies with VCP Inhibitors in Ovarian Cancer. Cancers (Basel) 2022;14
53. Hanahan D Hallmarks of Cancer: New Dimensions. Cancer Discov 2022;12:31–46 [PubMed: 35022204]
54. Expression Zhang G. and Prognostic Significance of BANF1 in Triple-Negative Breast Cancer. Cancer Manag Res 2020;12:145–50 [PubMed: 32021431]
55. Fu Z, Jiao Y, Li Y, Ji B, Jia B, Liu B. TYMS presents a novel biomarker for diagnosis and prognosis in patients with pancreatic cancer. Medicine (Baltimore) 2019;98:e18487
56. Wang X, Kuang W, Ding J, Li J, Ji M, Chen W, et al. Systematic Identification of the RNA-Binding Protein STAU2 as a Key Regulator of Pancreatic Adenocarcinoma. Cancers (Basel) 2022;14
57. Wang Z, Zhang H, Cheng Q. PDIA4: The basic characteristics, functions and its potential connection with cancer. Biomed Pharmacother 2020;122:109688
58. Lu Y, Yu S, Wang G, Ma Z, Fu X, Cao Y, et al. Elevation of EIF4G1 promotes non-small cell lung cancer progression by activating mTOR signalling. J Cell Mol Med 2021;25:2994–3005 [PubMed: 33523588]
59. Xing F, Song Z, Cheng Z. High expression of PDIA4 promotes malignant cell behavior and predicts reduced survival in cervical cancer. Oncol Rep 2022;48 [PubMed: 35014682]
60. Al-Kershi S, Bhayadia R, Ng M, Verboon L, Emmrich S, Gack L, et al. The stem cell-specific long noncoding RNA HOXA10-AS in the pathogenesis of KMT2A-rearranged leukemia. Blood Adv 2019;3:4252–63 [PubMed: 31867596]
61. Guo C, Ju QQ, Zhang CX, Gong M, Li ZL, Gao YY. Overexpression of HOXA10 is associated with unfavorable prognosis of acute myeloid leukemia. BMC Cancer 2020;20:586 [PubMed: 32571260]
62. Li B, Jin H, Yu Y, Gu C, Zhou X, Zhao N, et al. HOXA10 is overexpressed in human ovarian clear cell adenocarcinoma and correlates with poor survival. Int J Gynecol Cancer 2009;19:1347–52 [PubMed: 20009888]
63. Long Z, Li Y, Gan Y, Zhao D, Wang G, Xie N, et al. Roles of the HOXA10 gene during castrate-resistant prostate cancer progression. Endocr Relat Cancer 2019;26:279–92 [PubMed: 30667363]
64. Li J, Chang J, Wang J, Xu D, Yang M, Jiang Y, et al. HOXA10 promote pancreatic cancer progression via directly activating canonical NF-kappaB signaling pathway. Carcinogenesis 2022;43:787–96 [PubMed: 35553652]
65. Gelfo V, Romaniello D, Mazzeschi M, Sgarzi M, Grilli G, Morselli A, et al. Roles of IL-1 in Cancer: From Tumor Progression to Resistance to Targeted Therapies. Int J Mol Sci 2020;21 [PubMed: 33375030]
66. Das S, Shapiro B, Vucic EA, Vogt S, Bar-Sagi D. Tumor Cell-Derived IL1beta Promotes Desmoplasia and Immune Suppression in Pancreatic Cancer. Cancer Res 2020;80:1088–101 [PubMed: 31915130]

67. Karki R, Man SM, Kanneganti TD. Inflammasomes and Cancer. Cancer Immunol Res 2017;5:94–9 [PubMed: 28093447]
68. Ozaki T, Nakagawara A. Role of p53 in Cell Death and Human Cancers. Cancers (Basel) 2011;3:994–1013 [PubMed: 24212651]
69. Oh JH, Deasy JO. A literature mining-based approach for identification of cellular pathways associated with chemoresistance in cancer. Brief Bioinform 2016;17:468–78 [PubMed: 26220932]
70. Amin S, Baine M, Meza J, Alam M, Lin C. The impact of immunotherapy on the survival of pancreatic adenocarcinoma patients who do not receive definitive surgery of the tumor. Clin Transl Radiat Oncol 2020;24:34–40 [PubMed: 32613090]
71. Yu M, Guan R, Hong W, Zhou Y, Lin Y, Jin H, et al. Prognostic value of tumor-associated macrophages in pancreatic cancer: a meta-analysis. Cancer Manag Res 2019;11:4041–58 [PubMed: 31118813]
72. Liu C, Cheng H, Luo G, Lu Y, Jin K, Guo M, et al. Circulating regulatory T cell subsets predict overall survival of patients with unresectable pancreatic cancer. Int J Oncol 2017;51:686–94 [PubMed: 28714519]
73. Luksza M, Sethna ZM, Rojas LA, Lihm J, Bravi B, Elhanati Y, et al. Neoantigen quality predicts immunoediting in survivors of pancreatic cancer. Nature 2022;606:389–95 [PubMed: 35589842]
74. Ma H, Qian W, Bambouskova M, Collins PL, Porter SI, Byrum AK, et al. Barrier-to-Autointegration Factor 1 Protects against a Basal cGAS-STING Response. mBio 2020;11
75. Del Valle L, Dai L, Lin HY, Lin Z, Chen J, Post SR, et al. Role of EIF4G1 network in non-small cell lung cancers (NSCLC) cell survival and disease progression. J Cell Mol Med 2021;25:2795–805 [PubMed: 33539648]
76. Li H, Liu Q, Xiao K, He Z, Wu C, Sun J, et al. PDIA4 Correlates with Poor Prognosis and is a Potential Biomarker in Glioma. Onco Targets Ther 2021;14:125–38 [PubMed: 33447054]
77. Corselli M, Saksena S, Nakamoto M, Lomas WE, 3rd, Taylor I, Chattopadhyay PK. Single cell multiomic analysis of T cell exhaustion in vitro. Cytometry A 2022;101:27–44 [PubMed: 34390166]
78. Badura M, Braunstein S, Zavadil J, Schneider RJ. DNA damage and eIF4G1 in breast cancer cells reprogram translation for survival and DNA repair mRNAs. Proc Natl Acad Sci U S A 2012;109:18767–72 [PubMed: 23112151]
79. Chen TY, Yang CY, Yang MT, Kuo TF, Chang CL, Chen CL, et al. Protein disulfide isomerase a4 promotes lung cancer development via the Stat3 pathway in stromal cells. Clin Transl Med 2022;12:e606 [PubMed: 35170261]
80. Mi H, Sivagnanam S, Betts CB, Liudahl SM, Jaffee EM, Coussens LM, et al. Quantitative Spatial Profiling of Immune Populations in Pancreatic Ductal Adenocarcinoma Reveals Tumor Microenvironment Heterogeneity and Prognostic Biomarkers. Cancer Res 2022

## Translational Relevance

Current studies on identifying prognostic markers in pancreatic ductal adenocarcinoma (PDAC) have failed to define a clinically relevant panel, suggesting that an alternative approach may yield better results. Since the presence of short- and long-term surviving groups in PDAC is an established phenomenon, we analyzed the transcriptomics of these patients to delineate differentially expressed genes. This analysis yielded a novel HOXA10-associated 5 gene-based prognostic signature that is highly expressed in short-term survivors of PDAC. A deeper evaluation of this signature highlighted its connection with cancer-associated signaling pathways and immunosuppressive cell types, specifically $T_{regs}$ and M2 macrophages. These findings indicate that this gene signature may be utilized to stratify short- and long-term survivors of PDAC and provide therapeutic targets that directly influence patient survival.

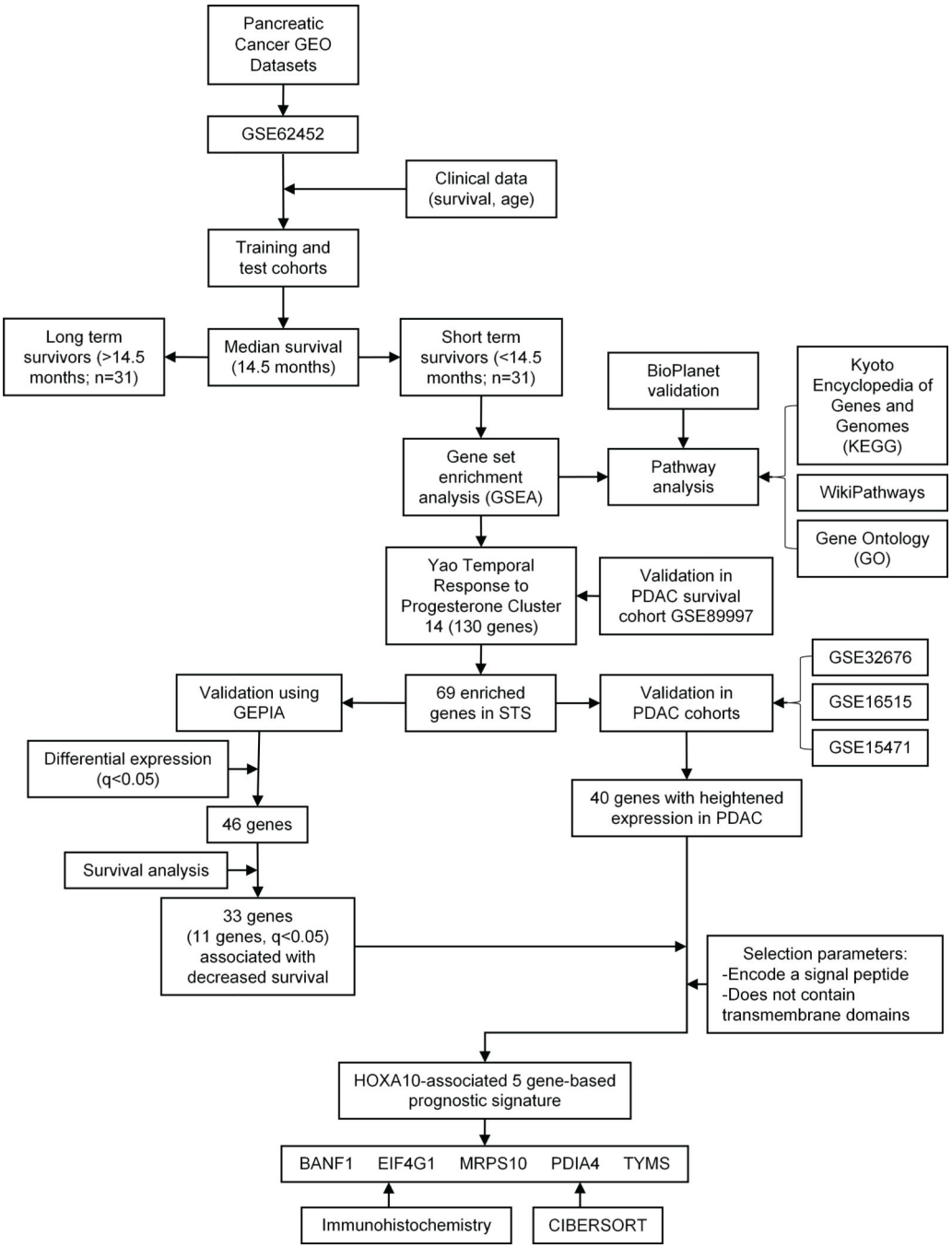


**Figure 1.**

Workflow for bioinformatics analysis and signature validation. GSE62542 was divided into STS and LTS based on a median survival of 14.5 months, and data for STS was subjected to GSEA and pathway enrichment analyses. A total of 130 genes were identified as part of the HOXA10 gene set, and 69 were specifically enriched in STS. Additional validation of gene expression data and association with survival determined five genes for final validation using IHC.

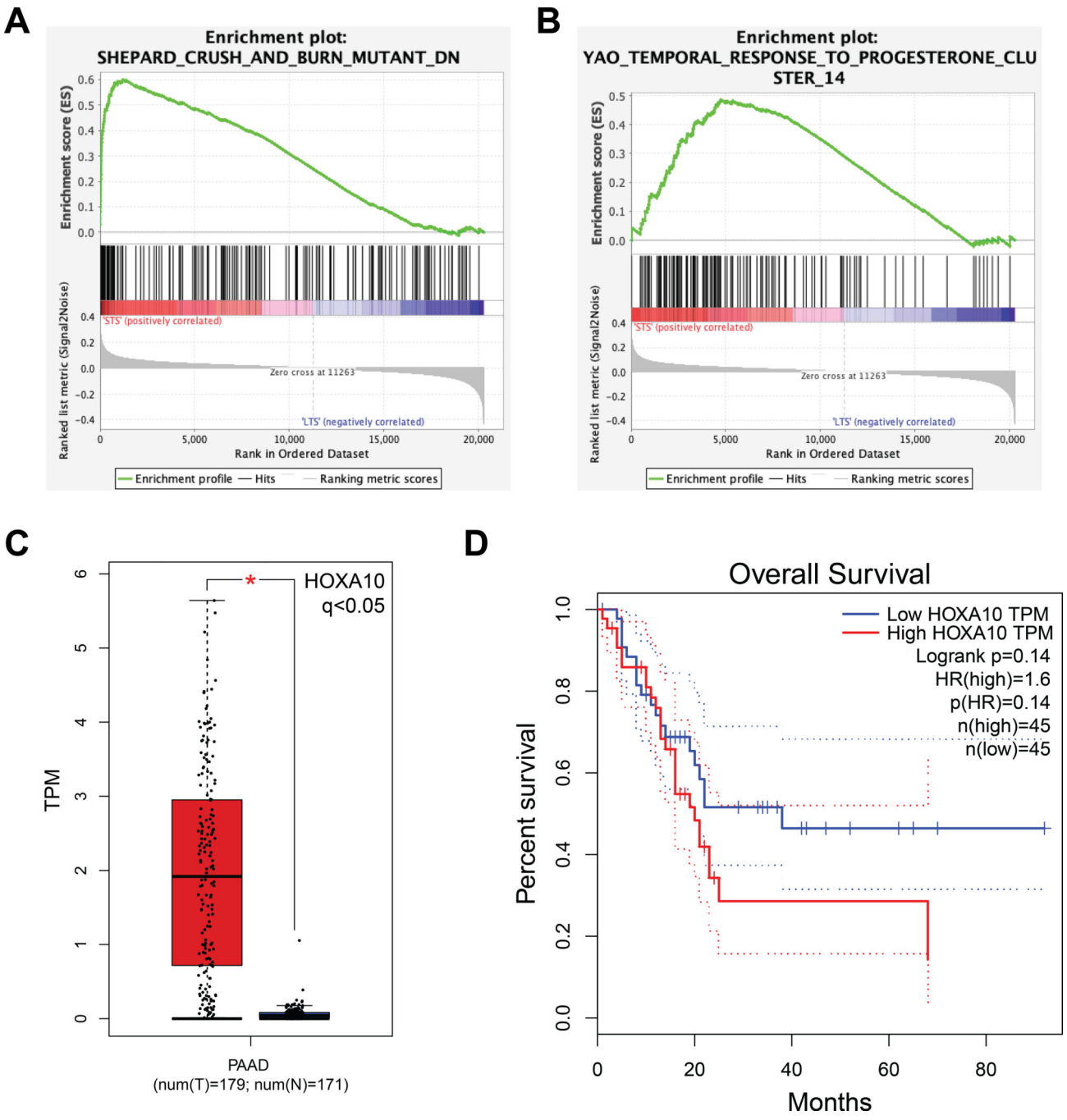


**Figure 2.**

Gene set enrichment analysis (GSEA) enrichment plots and HOXA10 expression profile. **A** and **B,** An enrichment score (ES) of 0.59 was observed for the gene set driven by MYBL2 (**A**), and an ES of 0.49 was determined for the HOXA10-driven gene set (**B**). **C** and **D,** HOXA10 is significantly upregulated in PDAC compared to the normal pancreas (**C**), and associated with decreased survival of PDAC patients (**D**).

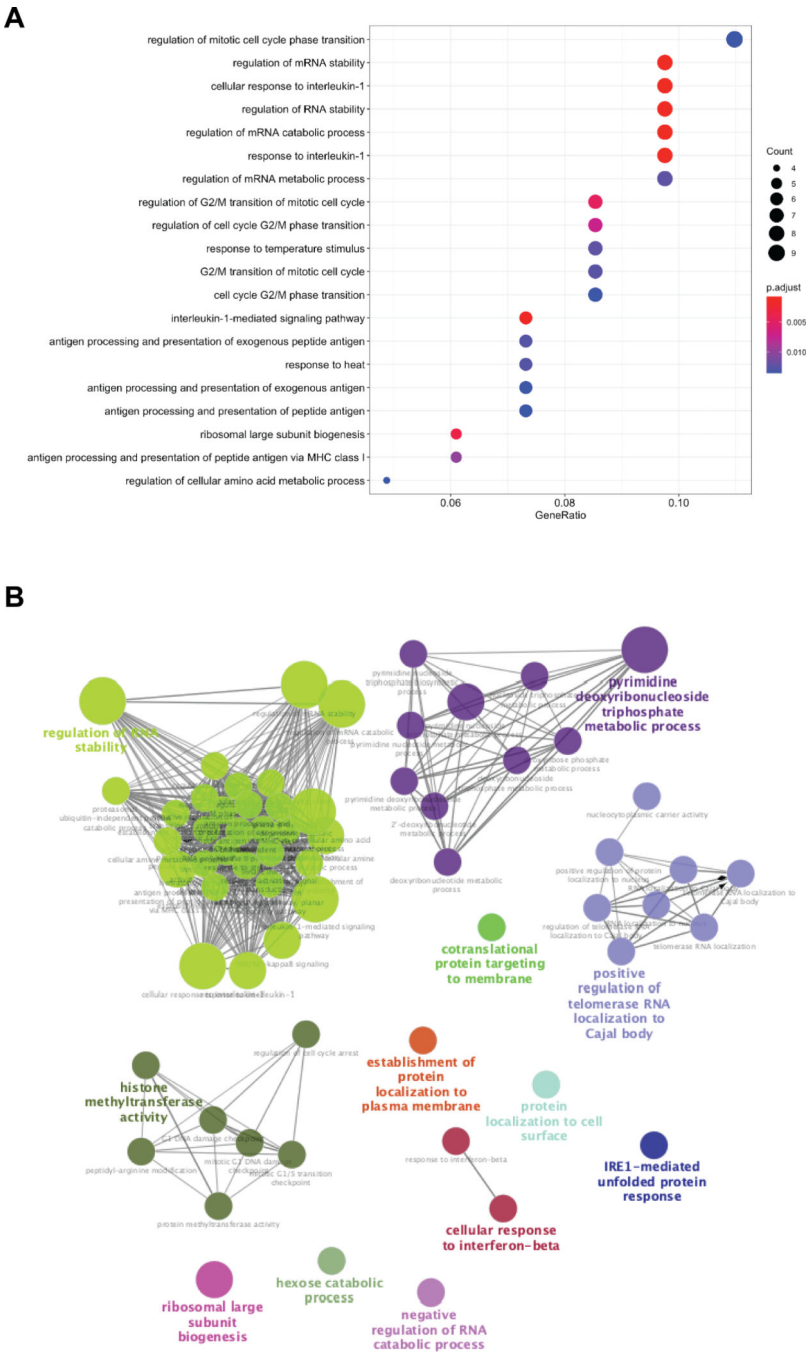


**Figure 3.**

Geno Ontology (GO) enrichment and ClueGO analysis of the HOXA10 gene set. **A,** Pathways involved in protein homeostasis (regulation of mRNA stability, ribosomal large subunit biogenesis), cellular metabolism (regulation of mRNA catabolic/metabolic process, regulation of cellular amino acid metabolic process), immune response (cellular response to interleukin-1, response to interleukin-1, interleukin-1-mediated signaling pathway, antigen processing and presentation of exogenous peptide antigen/via MHC class I), and cell cycle (regulation of mitotic cell cycle phase transition, regulation of G2/M transition of mitotic

cell cycle/phase transition) were enriched in the HOXA10 gene set. The count depicts the number of genes included in each GO term, and the gene ratio demonstrates the percentage of HOXA10 gene set genes enriched in each GO term. All terms were significantly associated with the HOXA10 gene set. **B,** ClueGO analysis revealed that the HOXA10 gene set is enriched in several biological networks, namely those associated with cellular metabolism, protein homeostasis, and immune response.

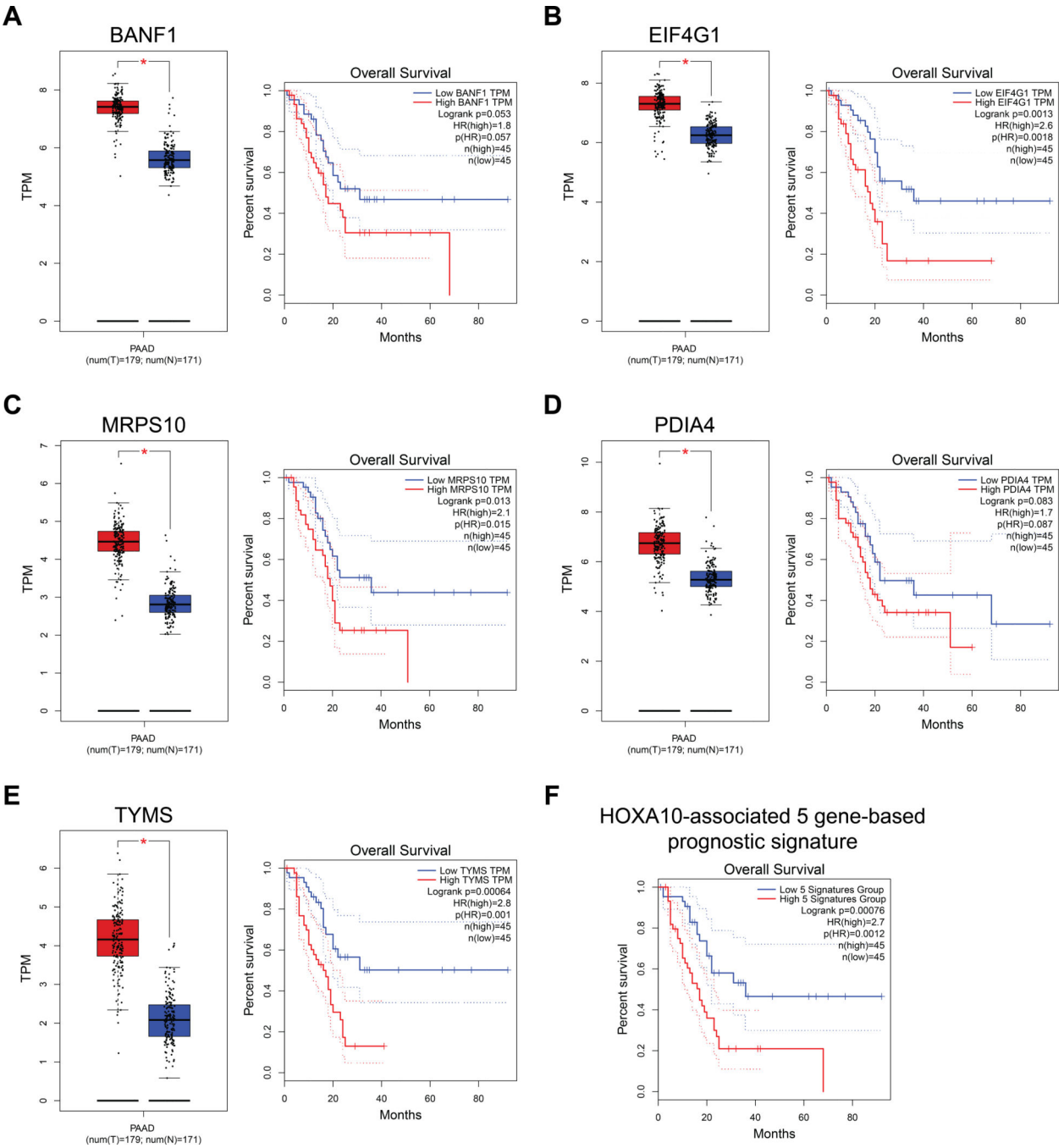


**Figure 4.**

GEPIA profiles for the five signature genes, including relative mRNA expression in tumor (red) compared to the normal pancreas (blue) and Kaplan-Meier plots of OS in patients with high expression of these five genes. **A-E,** BANF1 is significantly upregulated in PDAC compared to the normal pancreas and associated with decreased survival (**A**); EIF4G1 is significantly upregulated in PDAC compared to the normal pancreas and significantly associated with decreased survival (p=0.0013) (**B**); MRPS10 is significantly upregulated in PDAC compared to the normal pancreas and significantly associated with decreased

survival (p=0.013) (**C**); PDIA4 is significantly upregulated in PDAC compared to the normal pancreas and associated with decreased survival (**D**); TYMS is significantly upregulated in PDAC compared to the normal pancreas and significantly associated with decreased survival (p=0.00064) (**E**). **F,** As a panel, high expression of these five genes is significantly associated with decreased survival (p=0.0012).

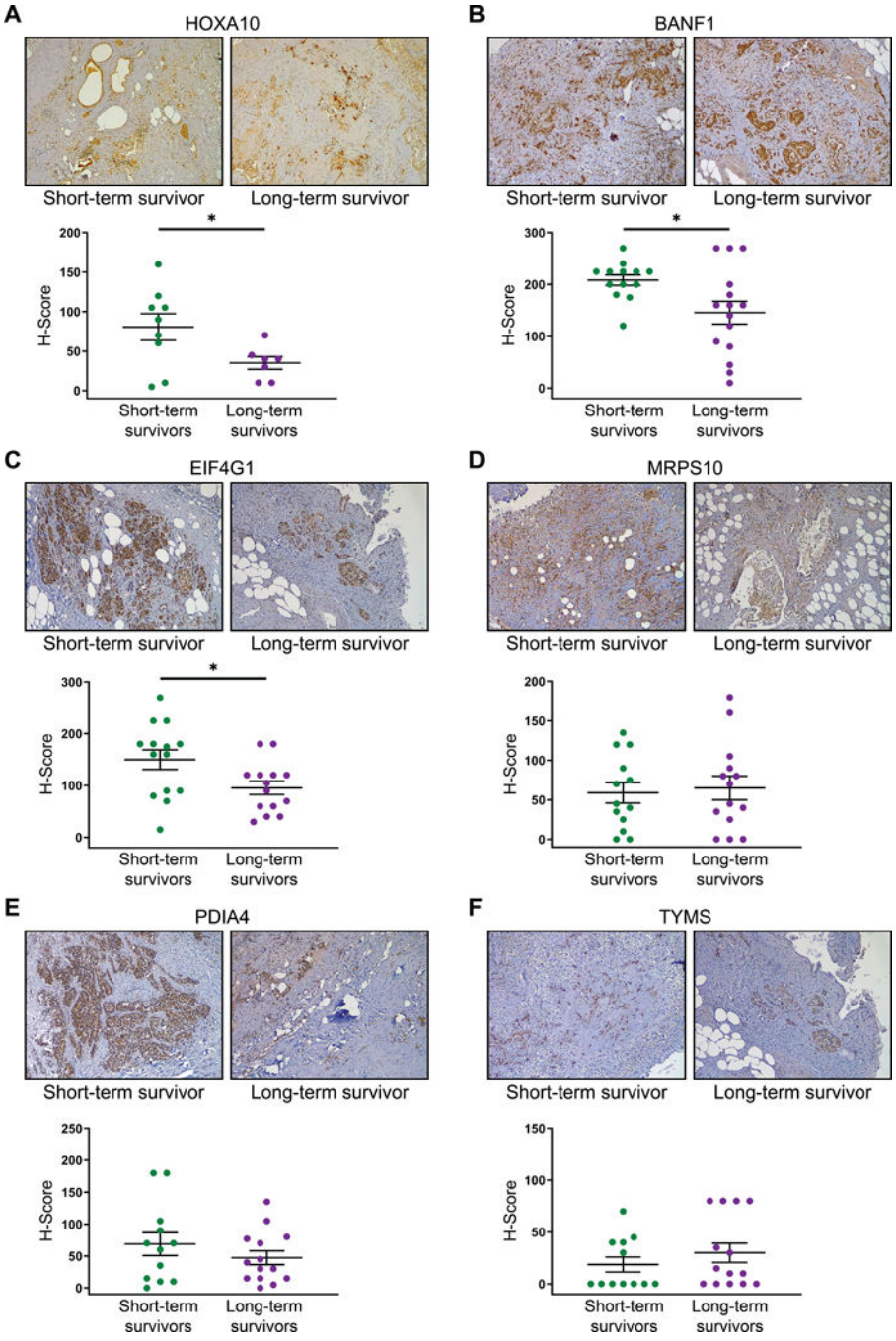


**Figure 5.**

Immunohistochemical analysis of HOXA10 and the five signature genes. **A-E,** In a cohort of short (HOXA10 n=9; BANF1 n=13; EIF4G1 n=14; MRPS10 n=13; PDIA4 STS n=12; TYMS n=12) and long surviving (HOXA10 n=7; BANF1 n=15; EIF4G1 n=14; MRPS10 n=14; PDIA4 STS n=14; TYMS n=14) patients, a trend of increased expression was observed in short-term survivors for the five signature genes (*, $p<0.05$). Panels C and F contain serial sections from LTS patients, stained with two different antibodies for EIF4G1 and TYMS.

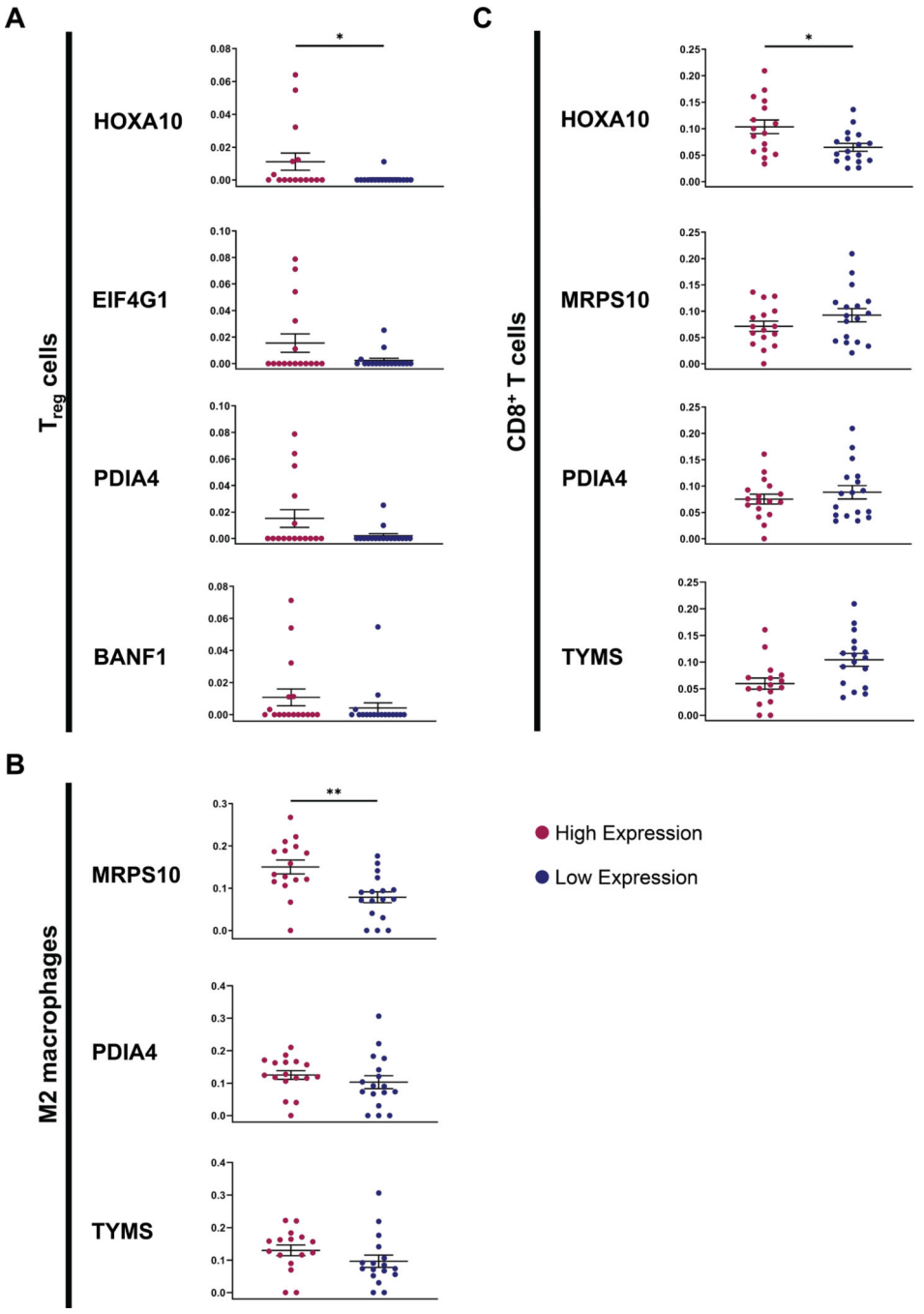


**Figure 6.**

CIBERSORT analysis for HOXA10 and its driven signature. Heightened expression of HOXA10 and its driven signature is associated with altered levels of immunosuppressive cell types. **A,** $T_{reg}$ cells were elevated in groups that highly expressed HOXA10 (p=0.0163), EIF4G1, PDIA4, and BANF1. **B,** M2 macrophages were elevated in groups highly expressing MRPS10 (p=0.0017), PDIA4, and TYMS. **C,** $CD8^+$ T cells were elevated in

patients with high expression of HOXA10 (p=0.0207), and decreased in patients with high expression of MRPS10, PDIA4, and TYMS.