Observational Study

Medicine®

OPEN

Genomic analysis of pancreatic cancer reveals 3 molecular subtypes with different clinical outcomes

Ji Woong Hwang, MD[a], Soo Kyung Jang, MD[b], Dong Jin Lee, MD, PhD[b,*]

Abstract

Pancreatic cancer has a very high mortality with a 5-year survival of <5%. The purpose of this study was to classify specific molecular subtypes associated with prognosis of pancreatic cancer using The Cancer Genome Atlas (TCGA) multiplatform genomic data.

Multiplatform genomic data (N = 178), including gene expression, copy number alteration, and somatic mutation data, were obtained from cancer browser (https://genome-cancer.ucsc.edu, cohort: TCGA Pancreatic Cancer). Clinical data including survival results were analyzed. We also used validation cohort (GSE50827) to confirm the robustness of these molecular subtypes in pancreatic cancer.

When we performed unsupervised clustering using TCGA gene expression data, we found three distinct molecular subtypes associated with different survival results. Copy number alteration and somatic mutation data showed different genomic patterns for these three subtypes. Ingenuity pathway analysis revealed that each subtype showed differentially altered pathways. Using each subtype-specific genes (200 were selected), we could predict molecular subtype in another cohort, confirming the robustness of these molecular subtypes of pancreatic cancer. Cox regression analysis revealed that molecular subtype is the only significant prognostic factor for pancreatic cancer ($P = .042$, 95% confidence interval 0.523–0.98).

Genomic analysis of pancreatic cancer revealed 3 distinct molecular subtypes associated with different survival results. Using these subtype-specific genes and prediction model, we could predict molecular subtype associated with prognosis of pancreatic cancer.

Abbreviations: BCCP = Bayesian compound covariate predictor, ROC = receiver operating characteristic, TCGA = The Cancer Genome Atlas.

Keywords: molecular subtypes, pancreatic cancer, prediction, RNAseq, TCGA

1. Introduction

Pancreatic cancer is a highly aggressive malignancy with a dismal prognosis.[1] This tumor is currently the fourth leading cause of cancer deaths in the United States,[2] and it is projected to be the second leading cause of cancer deaths by the year 2030.[3,4] Due to the lack of cancer-specific symptoms at the beginning stage of tumor initiation, <10% of pancreatic cancer patients are diagnosed at an early stage. Thus most patients are detected late and lose the opportunity of surgical resection. Unfortunately, even those who are able to receive surgery have a high probability of recurrence within the first 12 months[5] due to the aggressiveness of this tumor. In contrast with improved outcome of other solid tumors after treatments, the prognosis of pancreatic cancer remains poor during the past 2 decades. Pancreatic cancer patients have a median survival of 6 months and the 5-year survival rate is only 6% despite almost 50 years of research and therapeutic developments.[6] One of the reasons for such slow progress in treating pancreatic cancer is the lack of accurate prognostic markers which are essential for establishing individualized treatment strategies. Currently, traditional factors such as tumor grade and the TNM stage are used to determine treatment modality and predict the survival result of pancreatic cancer.[7] However, patients with the same TNM stage or other pathologic prognostic factors have diverse clinical courses and subsequently different prognoses. In addition, their responses to standard treatment options are heterogenous. Therefore, new prognostic factors need to be identified. The ability to establish an individualized treatment plan and prognosis prediction based

Editor: Naoshad Muhammad.

Funding: This study was funded by Hallym University Research Fund 2018 (HURF-2018-66).

Availability of data and material: All genomic data (N = 178) and clinical data of pancreatic cancer from TCGA project were obtained from TCGA data portal (https://tcga-data.nci.nih.gov) and cancer browser (https://genome-cancer.ucsc.edu). Gene expression data of validation cohort were obtained from Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/, accession number: GSE50827)

The authors report no conflicts of interest.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

[a] Department of Surgery, Kangnam Sacred Heart Hospital, Hallym University College of Medicine, [b] Department of Otolaryngology-Head and Neck Surgery, Division of Precision Medicine, Kangnam Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Korea.

[*] Correspondence: Dong Jin Lee, 1 Singil-ro Yeongdeungpo-gu, Seoul 150-950, Republic of Korea (e-mail: leesurgery@gmail.com).

How to cite this article: Hwang JW, Jang SK, Lee DJ. Genomic analysis of pancreatic cancer reveals 3 molecular subtypes with different clinical outcomes. Medicine 2021;100:14(e24969).

Received: 16 April 2020 / Received in final form: 9 February 2021 / Accepted: 10 February 2021

http://dx.doi.org/10.1097/MD.0000000000024969

on molecular subtypes may extend the patient's survival time. Furthermore, defining the molecular subtypes of pancreatic cancer may contribute to the comprehensive understanding of genomic transition and cancer development.[8,9] Recently, several studies have reported on the molecular subtypes related to the prognosis of pancreatic cancer.[10–13] However, discrepancies among study results have been observed between molecular subtypes and their prognostic values. This could be due to limited number and regional differences of study populations.

The Cancer Genome Atlas (TCGA) has generated multiplatform genomic data for thousands of tumor samples across more than 25 cancer types, including pancreatic cancer.[14–16] Such a large number of available TCGA tumor datasets provide us an opportunity to study genomic profiles (including gene expression, copy number alteration, and somatic mutation data) of pancreatic cancer with increased statistical power. The first objective of this study was to determine the molecular subtypes of pancreatic cancer associated with clinical behavior by analyzing gene expression pattern using TCGA data. Second, this study aimed to develop a model to predict prognostic subgroups by using differentially expressed genes. The biological pathways of each molecular subtype were also explored.

2. Materials and methods

2.1. Genomic and clinical data sets

All genomic data (N = 178) of pancreatic cancer were obtained from TCGA Data Portal (https://tcga-data.nci.nih.gov) and the University of California Santa Cruz Cancer Browser (https://genome-cancer.ucsc.edu, cohort: TCGA Pancreatic Cancer). Gene expression data from mRNA-seq, copy number alteration data, somatic mutation data, and clinical data were included in the analyses. Clinical data included age, sex, TNM stage, survival data, and the primary anatomic site of pancreatic cancer (Supplement Data 1, http://links.lww.com/MD2/A23). In this study, approval of ethics committee or institutional review board was not necessary because we used publicly available genomic data.

2.2. Analysis of gene expression data and unsupervised clustering

To analyze gene expression data, BRB-ArrayTools software program (http://linus.nci.nih.gov/BRB-ArrayTools.html)[17] was used. After gene expression data were gene-median centered, gene variability was computed using median absolution deviation. A total of 6856 most variable genes were selected. ConsensusClusterPlus (Bioconductor)[18] was used to perform unsupervised clustering for these genes in the 178 pancreatic cancer cases. A heatmap was generated using Cluster and TreeView software programs[19] (Supplement Figure S1, http://links.lww.com/MD2/A21). Other statistical analyses were performed in R language (http://www.r-project.org).

2.3. Survival analysis

Association of each subtype with overall survival was evaluated using Kaplan-Meier plots and log rank test. Overall survival was defined as the time from surgery to death. Data were censored when a patient was alive without recurrence at the last follow-up. Statistically significant difference was considered at P value $<.05$. All statistical analyses were conducted in R language environment (http://www.r-project.org).

2.4. Analysis of copy number alteration and somatic mutation

Copy number alteration and somatic mutation data were analyzed and visualized using OncoPrint (https://cbioportal.org). Of the 127 most frequently mutated cancer genes in 12 cancer types identified in a previous study,[20] the 15 of the most frequently mutated genes in pancreatic cancer were selected for analysis.

2.5. Selection of specific gene signature in each cluster

Multiple 2-class t tests were performed for all possible combination of the three subtypes. To select genes that were differentially expressed between each cluster, we applied a stringent cutoff of $P < .001$ (Student t test) and 1.5-fold difference.

2.6. Significant canonical signaling pathways enriched in each cluster

Pathway analysis was carried using Ingenuity Pathways Analysis (Ingenuity, Redwood City, CA). Genes from the dataset that were associated with a canonical pathway in Ingenuity Pathways Knowledge Base were considered for analysis. Significance of the association between 200 genes of each cluster and the canonical pathway was measured using Fisher exact test ($P < .001$).

2.7. Prediction models with genomic signatures

Gene expression signature from TCGA cohort was used to stratify patients in a validation cohort from Gene Expression Omnibus database: GSE50827.[21] Expression data from 200 subtype-specific genes in TCGA data set were combined to form a classifier according to a Bayesian compound covariate predictor (BCCP), as described previously.[22–26] The BCCP classifier estimated the likelihood of an individual patient being in one of three subtypes. Briefly, gene expression data for each subtype gene signature from the TCGA cohort (ie, 200 significant genes for each subtype) were used to generate a Bayesian probability of each tissue sample belonging to a particular subtype. Therefore, 3 probability scores were generated for each tumor. Samples in validation cohorts were assigned to 1 of these 3 subtypes according to the highest probability scores.

2.8. Cox regression analysis and hazard ratio of each clinical variable and gene signatures

Cox regression analysis was performed using each clinical variable and selected gene signatures. Statistically significant difference was considered at P value $<.05$. All statistical analyses were conducted in R language environment (http://www.r-project.org).

3. Results

3.1. Analysis of TCGA genomic data showed novel molecular subtypes of pancreatic cancer associated with clinical behavior

To explore the molecular profiles of pancreatic cancer, we carried out unsupervised clustering analysis with mRNA expression data

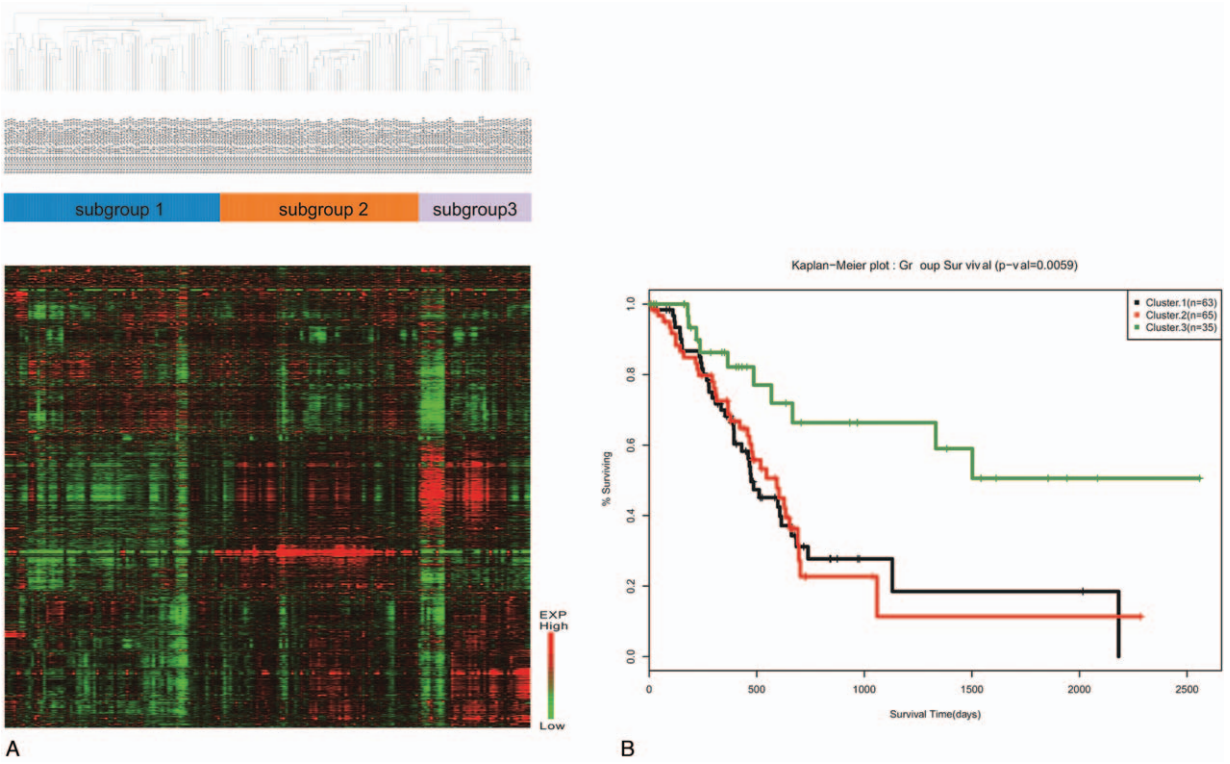

Figure 1. Unsupervised clustering of pancreatic cancers reveals 3 distinct molecular subtypes and Kaplan-Meier survival curve of 3 molecular subtypes of Pancreatic cancer. (A) A hierarchical clustering of gene expression data from 178 pancreatic cancer case in TCGA data. Genes with expression levels that were at least 2-fold different in at least 18 cases, relative to the median value across cases, were selected for hierarchical clustering analysis. The data are given in matrix format, in which rows represent individual genes and columns represent each patient. Each cell in the matrix represents the expression level of a gene feature in an individual patient. The color red or green in cell reflects relative high or low expression levels, respectively, as indicated in the scale bar. (B) Overall survival results of Pancreatic cancer patients according to each subtype. Cluster 3 showed better survival rate when compared to cluster 1 or cluster 2 ($P = .0059$).

from TCGA pancreatic cancer data. Interestingly, the gene expression profiles of pancreatic cancer were divided into three molecular subtypes (Fig. 1A). Case distribution according to each subgroup is shown in Table 1. χ^2 analysis showed that age or sex did not show any statistically significant difference among these three subtypes. Only nodal (N) stage showed a statistically significant difference ($P = .034$). Kaplan-Meier survival analysis revealed that the overall survival rate of patients with cluster 3 was significantly higher than that of subtype 1 or subtype 2 ($P = .0059$, Fig. 1B).

3.2. Association of three molecular subtypes with copy number alteration and somatic mutation

Oncoprint analysis revealed copy number alteration and somatic mutation patterns of three molecular subtypes of pancreatic cancer (Fig. 2). KRAS and TP53 mutations were more frequently found in subgroup 1 and subgroup 2 than those in subgroup 3. Frequency of TTN mutation was not significantly different among the 3 subgroups. When we analyzed copy number alteration, subgroup 1 showed amplification for TUBB8P7, DNM1P47, LINC00969, GTF2IRD2P1, TCF20, BTN2A3P, and PRSS3P2 with SNHG14 deletion. Subgroup 2 showed amplification for DNM1P47, GTF2IRD2P1, TCF20, and RRN3P2 with deletion of BAGE2 and LINC00969. Overall pattern revealed that subgroup 3 showed little alteration in copy number compared to subgroup 1 or subgroup 2 (Fig. 2).

3.3. Selecting specific genes for each cluster

To understand the difference in the underlying biology of the three subtypes, we sought to find genes whose expression was specific to each subtype by applying multiple 2-class *t* tests among these 3 subtypes ($P < .001$) (Supplement Data 2, http://links.lww.com/MD2/A26). Genes were then ranked according to fold-ratios and the top 200 genes were selected for each subtype (Fig. 3).

3.4. Activated signaling pathways in three subtypes

To uncover potential signaling pathways activated in each subtype, we carried out gene network analysis using analysis tool in IPA (Supplement Data 3, http://links.lww.com/MD2/A28). Results of analysis predicted that subtype 1 had alterations in ATM signaling, FXR/RXR activation pathway, and P53 signaling pathway while subtype 2 had alterations in nicotine degradation pathway, notch signaling, and PTEN signaling pathway. Analysis also predicted that subtype 3 had alterations in FAK signaling, ILK signaling, and PI3K signaling pathway (Table 2).

3.5. Validation for the presence of these three subtypes in independent cohorts

Having defined a gene expression signature that reflected the three molecular subtypes of pancreatic cancer significantly associated with prognosis, we next validated the presence of

Table 1

Case distribution according to each cluster after unsupervised clustering of The Cancer Genome Atlas Pancreatic cancers (N = 178).

	Cluster 1 (N = 71)	Cluster 2 (N = 69)	Cluster 3 (N = 38)	Total (N = 178)	*P*
Mean age, y	66.3	64.7	61.7		
Sex					.549
Male	34	38	19	91	
Female	29	28	17	74	
NA	8	3	2	13	
Stage					.065
IA	3	1	3	7	
IB	4	2	7	13	
IIA	14	8	3	25	
IIB	39	53	20	112	
III	2	1	1	4	
IV	1	1	1	3	
NA	8	3	3	14	
pT					.060
pT1	3	1	4	8	
pT2	3	9	7	19	
pT3	56	55	23	134	
pT4	1	1	1	3	
NA	8	3	3	14	
pN					.034
pN0	20	12	12	44	
pN1a	41	54	19	114	
pN1b	1	0	2	3	
NA	9	3	5	11	
Anatomic location					.086
Head of pancreas	48	58	24	130	
Body of pancreas	6	2	3	11	
Tail of pancreas	7	2	4	13	
Other location	2	4	5	11	
NA	8	3	2	13	

NA = data not available, pN = pathologic N stage, pT = pathologic T stage.

these three molecular subtypes in independent cohorts (GSE50827). To construct a subtype prediction model, we adapted a previously developed model using BCCP algorithms. When the prediction model was applied to data from validation cohort (GSE50827), validation cohort was divided into the 3 molecular subtypes (Supplement Figure S2, http://links.lww.com/MD2/A21, and Supplement Figure S3, http://links.lww.com/MD2/A21).

3.6. Cox regression analysis and hazard ratio of each clinical variable and gene signatures

Results of cox regression with univariate analysis revealed that node positive and gene signatures for molecular subtype showed statistically significant *P* value (Table 3). However, in multivariate analysis, molecular subtype is the only significant prognostic factor for pancreatic cancer ($P = .042$, 95% confidence interval 0.523–0.98) (Fig. 4).

4. Discussion

The lack of prognostic factors needed to precisely determine a patient's individualized treatment plan has contributed to the poor results of pancreatic cancer. Current prognostic markers for pancreatic cancer include clinicopathological features such as depth of invasion, tumor grade, TNM stage, and histologic differentiation. However, most of these markers only become available after surgical resection and inconsistencies between these predictors and survival is often observed.[27,28] Serum carbohydrate antigen 19–9 (CA 19–9) is still the only biomarker

Figure 2. Most frequently altered genes in pancreatic cancer. Mutated gene cprofiles were different between each cluster. In cluster 2 and 3, Tp53 showed missense mutations in most of cases but not in cluster 1. Also CDKN2A mutations were found in cluster 2 and cluster 3 but not in cluster 1. In contrast, PIK3CA mutation was more frequent in cluster1 than other clusters. NOTCH1 mutation was frequently found in cluster 2 when compared to other clusters. *NFE2L2* gene which is related with oxidative stress response pathway was frequently mutated in cluster 3.

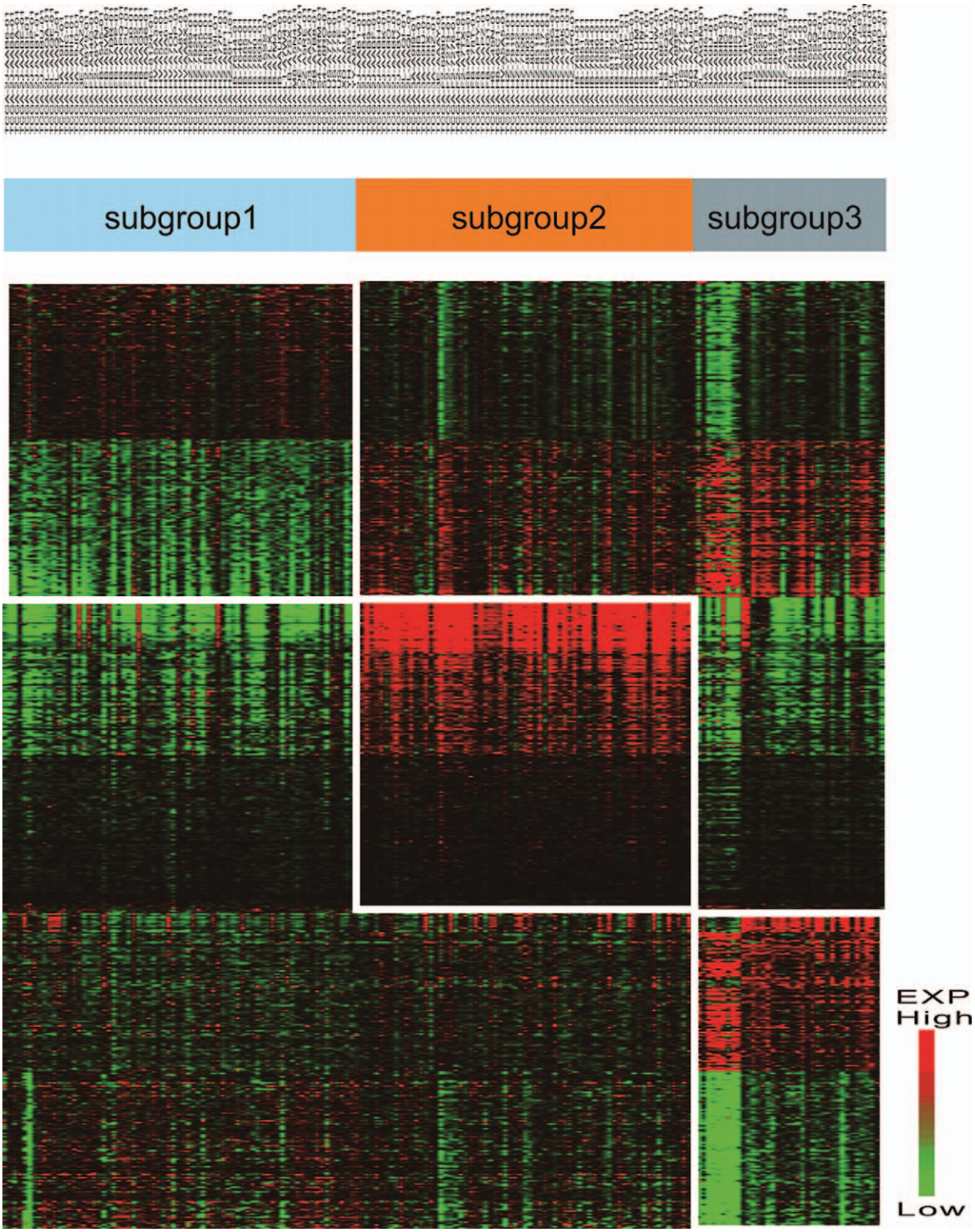

Figure 3. Each cluster-specific gene expression patterns showed by heatmap. The data are given in the matrix represents the expression level of a gene feature in an individual patient. We selected each-cluster specific 200 gens and draw the heatmap. The color red or green in cell reflects relative high or low expression levels.

able to monitor disease progression during pancreatic cancer treatment,[29] but poor specificity and false-positive elevation in the presence of obstruction jaundice restrict its role in clinical practice.[30] Thus, finding prognostic biomarkers that might improve clinical outcome through patient classification with further understanding of the mechanisms of pancreatic cancer is highly desirable. Recently, several studies have reported on the molecular subtypes related to the prognosis of pancreatic

Table 2

Subgroup-specific altered canonical pathways and related molecules.

Ingenuity canonical pathways	-log (*P*)	Ratio	Molecules
Subgroup 1			
ATM signaling	3.06	0.0625	RAD51, CDC25C, TP73, CCNB2, CDK1
FXR/RXR activation	2.95	0.0476	G6PC2, TTR, SDC1, APOH, VTN, GC
p53 signaling	1.71	0.036	TP73, E2F1, BIRC5, KLB
Subgroup 2			
Nicotine degradation II	1.65	0.0411	UGT2A3, NADP, FMO4
Notch signaling	1.41	0.0526	NOTCH2, FURIN
PTEN signaling	1.11	0.0248	MAGI1, INSR, RAC3
Subgroup 3			
FAK signaling	1.24	0.0303	CAPN8, CAPN9, ACTC1
ILK signaling	1.01	0.0204	RHOV, MUC1, ITGB6, ACTC1
PI3K signaling	.978	0.0234	BLK, CD19, CR2

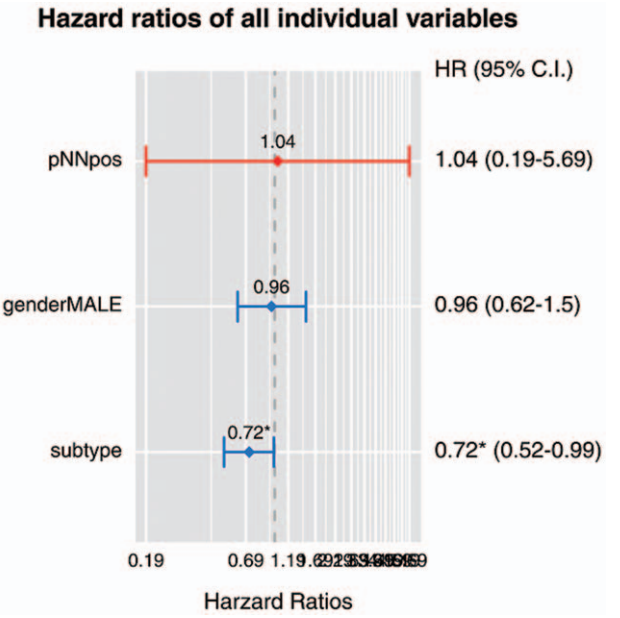

Figure 4. Cox regression result and hazard ratio of each clinical variable and molecular subtype. In multivariate analysis, molecular subtype is the only significant prognostic factor for pancreatic cancer.

cancer.[10,11] Sanjeev et al demonstrated that restoration of miR-200 resulted in the reversal of drug resistance and sensitizes pancreatic cancer cells to gemcitabine cytotoxicity.[10] Ioannis et al reported the expressions of miR-21 and miR-155 were associated with tumor stage and poor prognosis.[11] However, discrepancies among study results have been observed between molecular subtypes and their prognostic values. Such discrepancies could be due to the limited number of study populations with regional differences.

To explore molecular profiles of pancreatic cancer, we carried out unsupervised clustering analysis with mRNA expression data from TCGA pancreatic cancer data. Interestingly, gene expression profiles of pancreatic cancer were divided into 3 molecular subtypes (Fig. 1A). Consensus cluster confirmed the robustness of these three molecular subtypes. When we used Kaplan-Meier survival analysis, the overall survival rate of patients in subtype 3 was significantly higher than that in subtype 1 or subtype 2. Although pancreatic cancer is a disease with poor survival results, subtype 3 showed >60% of 5-year overall survival rate (Fig. 1B). In this respect, pancreatic cancer patients in subgroup 3 should receive a more aggressive treatment plan. Oncoprint analysis revealed different copy number alteration patterns among these three molecular subtypes of pancreatic cancer (Fig. 2). Subgroup 1 and subgroup 2 showed more alterations of copy number at different gene levels. However, subgroup 3 showed little alterations of copy number compared to subgroup 1 or subgroup 2 (Fig. 2). Regarding somatic mutation, KRAS and TP53 mutations were more frequently found in subgroup 1 and subgroup 2 than those in subgroup 3. Frequencies of TTN mutation were not significantly different among the three subgroups. All these findings indicate that chromosome instability might be the reason for the poor prognosis of pancreatic cancer patients, especially for those in subgroup 1 and subgroup 2.

Among subgroup-specific genes, *COL7A1* gene was expressed higher in subgroup 1 and PNLIP gene was expressed higher in subgroup 2. Both *COL7A1* and *PNLIP* genes have been reported as progression-related genes in pancreatic cancer.[31,32] Also, *CNDN18* gene was low-expressed in subgroup 3. *CLDN18* gene has been reported as transcriptional regulator via specific protein kinase C signaling pathway and modification of DNA methylation in human pancreatic cancer cells.[33]

In this study, we carried out gene network analysis by using analysis tool in IPA to uncover potential signaling pathways activated in each subtype (Supplement Data 2, http://links.lww.com/MD2/A23). Analysis predicted that subtype 1 had alterations in ATM signaling, FXR/RXR activation pathway, and P53 signaling pathway while subtype 2 had alterations in nicotine degradation pathway, notch signaling, and PTEN signaling pathway. Analysis also predicted that subtype 3 had alterations in ILK signaling, FAK signaling, and PI3K signaling pathway. Different altered pathways and target molecules should be considered depending on each molecular subtype when planning target therapy using specific target molecules to block certain pathways in pancreatic cancer.

Lastly, having defined gene expression signature that reflected the three molecular subtypes of pancreatic cancer that were significantly associated with prognosis, we tried to validate the presence of these three molecular subtypes in independent cohorts (GSE50827). To construct a subtype prediction model,

Table 3

Uivariate and multivariate analysis of clinicopathologic factors and gene signatures using Cox regression analysis.

Factors	Exp (β)	SE	95.0% CI	*P*
Univariate analysis				
Gene signatures	0.707	0.150	(0.526–0.948)	.021*
Sex (male)	0.894	0.221	(0.579–1.38)	.61
Node-positive	2.031	0.274	(1.186–3.476)	.0098*
Stage IV	4.101	1.007	(0.569–29.53)	.161
T4	1.613	1.226	(0.145–17.8)	.697
Multivariate analysis				
Gene signatures	0.719	1.39	(0.523–0.98)	.042*
Node-positive	1.04	0.8669	(0.190–5.690)	.963

CI = confidence interval, Exp (β) = odds ratio, SE = standard error.
* Means statistically significant *P* value.

we adapted a previously developed model using BCCP algorithms. The BCCP classifier estimated the likelihood that an individual patient would be included in 1 of 3 clusters according to a BCCP *P* value of .5. The cut-off was set by maximal point of sum of sensitivity and specificity. When the prediction model was applied to the validation cohort (GSE50827), robustness of the 3 molecular subtypes was confirmed in the validation cohort (Supplement Figure S2, http://links.lww.com/MD2/A21 and Supplement Figure S3, http://links.lww.com/MD2/A21), although we could not compare survival results among these three molecular subtypes due to the lack of survival data.

In summary, using TCGA multiplatform data, we were able to identify three molecular subtypes of pancreatic cancer that were associated with clinical behavior. We suggested gene signatures to predict the prognosis of pancreatic cancer and this could be utilized as a useful prognostic marker of pancreatic cancer. However, some limitations should be considered. First, this study was a retrospective study using TCGA data. A prospective clinical study is needed in the future to apply this result of translational study to a clinical setting. Second, the validation cohort (GSE50827) did not have survival results. Therefore, we were only able to validate the robustness of these 3 molecular subtypes in the validation cohort. Total 265 clinical cohorts of pancreatic cancer in Gene Expression Omnibus database and 3 other cohorts in TCGA data did not include survival data. For this reason, we were unable to examine the association between subtype 3 and better survival results. Further evaluation with more cases and survival data are warranted.

In conclusion, multiplatform genomic analysis of pancreatic cancer revealed 3 distinct molecular subtypes associated with different survival results. Cox regression analysis revealed that molecular subtype is the only significant prognostic factor for pancreatic cancer. Using these subtype-specific genes and prediction model, we could predict the molecular subtypes associated with prognosis in pancreatic cancer.

Acknowledgment

The authors appreciated patients and their families who generously donated their tissues to TCGA, as well as the members of TCGA who collected and disclosed valuable data.

Author contributions

Administrative support: Dong Jin Lee, Ji Woong Hwang.
Collection and assembly of data: Ji Woong Hwang.
Conception and design: Dong Jin Lee.
Conceptualization: Dong Jin Lee, Soo Kyung Jang.
Data analysis and interpretation: Dong Jin Lee, Ji Woong Hwang.
Data curation: Ji Woong Hwang, Dong Jin Lee, Soo Kyung Jang.
Formal analysis: Ji Woong Hwang.
Funding acquisition: Dong Jin Lee.
Investigation: Ji Woong Hwang.
Manuscript writing: Dong Jin Lee, Ji Woong Hwang.
Methodology: Ji Woong Hwang, Dong Jin Lee.
Provision of study materials or patients: Dong Jin Lee, Ji Woong Hwang.
Supervision: Dong Jin Lee.
Validation: Ji Woong Hwang.
Writing – original draft: Ji Woong Hwang.
Writing – review & editing: Dong Jin Lee.
Final approval of manuscript: All authors

References

[1] Hidalgo M. Pancreatic cancer. N Engl J Med 2010;362:1605–17.
[2] Singh D, Upadhyay G, Srivastava RK, et al. Recent advances in pancreatic cancer: biology, treatment, and prevention. Biochim Biophys Acta 2015;1856:13–27.
[3] Waddell N, Pajic M, Patch AM, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 2015;518:495–501.
[4] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013;63:11–30.
[5] Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. Gastroenterology 2013;144:1252–61.
[6] Thompson MJ, Rubbi L, Dawson DW, et al. Pancreatic cancer patient survival correlates with DNA methylation of pancreas development genes. PLoS One 2015;10:e0128814.
[7] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471–4.
[8] Stathis A, Moore MJ. Advanced pancreatic carcinoma: current treatment and future challenges. Nat Rev Clin Oncol 2010;7:163–72.
[9] Yue Y, Osipov A, Fraass B, et al. Identifying prognostic intratumor heterogeneity using pre- and post-radiotherapy 18F-FDG PET images for pancreatic cancer patients. J Gastrointest Oncol 2017;8:127–38.
[10] Srivastava SK, Arora S, Singh S, et al. MicroRNAs in pancreatic malignancy: progress and promises. Cancer Lett 2014;347:167–74.
[11] Papaconstantinou IG, Manta A, Gazouli M, et al. Expression of microRNAs in patients with pancreatic cancer and its prognostic significance. Pancreas 2013;42:67–71.
[12] Du Y, Zhao B, Liu Z, et al. Molecular subtyping of pancreatic cancer: translating genomics and transcriptomics into the clinic. J Cancer 2017;8:513–22.
[13] Collisson EA, Bailey P, Chang DK, et al. Molecular subtypes of pancreatic cancer. Nat Rev Gastroenterol Hepatol 2019;16:207–20.
[14] Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary Oncol 2015;19(1A):A68–77.
[15] Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113–20.
[16] Kosvyra A, Maramis C, Chouvarda I. A data-driven approach to build a predictive model of cancer patients' disease outcome by utilizing co-expression networks. Comput Biol Med 2020;125:103971.
[17] Simon R, Lam A, Li MC, et al. Analysis of gene expression data using BRB-ArrayTools. Cancer Inform 2007;3:11–7.
[18] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26:1572–3.
[19] Eisen MB, Spellman PT, Brown PO, et al. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.
[20] Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333–9.
[21] Nones K, Waddell N, Song S, et al. Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. Int J Cancer 2014;135:1110–8.
[22] Ramaswamy S, Tamayo P, Rifkin R, et al. Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci U S A 2001;98:15149–54.
[23] Oh SC, Park YY, Park ES, et al. Prognostic gene expression signature associated with two molecularly distinct subtypes of colorectal cancer. Gut 2012;61:1291–8.
[24] Lee JS, Chu IS, Heo J, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–76.
[25] Lee JS, Chu IS, Mikaelyan A, et al. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nat Genet 2004;36:1306–11.
[26] Lee JS, Heo J, Libbrecht L, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410–6.

[27] Ferrone CR, Pieretti-Vanmarcke R, Bloom JP, et al. Pancreatic ductal adenocarcinoma: long-term survival does not equal cure. Surgery 2012;152(3 suppl 1):S43–9.
[28] Hong SM, Park JY, Hruban RH, et al. Molecular signatures of pancreatic cancer. Arch Pathol Lab Med 2011;135:716–27.
[29] Chrystoja CC, Diamandis EP, Brand R, et al. Pancreatic cancer. Clin Chem 2013;59:41–6.
[30] Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: an evidence based appraisal. J Gastrointest Oncol 2012;3:105–19.
[31] Geng Y, Guan R, Hong W, et al. Identification of m6A-related genes and m6A RNA methylation regulators in pancreatic cancer and their association with survival. Ann Transl Med 2020;8:387.
[32] Medicherla S, Li L, Ma JY, et al. Antitumor activity of TGF-beta inhibitor is dependent on the microenvironment. Anticancer Res 2007;27:4149–57.
[33] Ito T, Kojima T, Yamaguchi H, et al. Transcriptional regulation of claudin-18 via specific protein kinase C signaling pathways and modification of DNA methylation in human pancreatic cancer cells. J Cell Biochem 2011;112:1761–72.