Title: Comprehensive machine-learning survival framework develop a consensus model in large scale multi-center cohorts for pancreatic cancer

Running Title: AIDPS guides individualized-treatment of PACA

Libo Wang[1,2,3†], Zaoqu Liu[4†]*, Ruopeng Liang[1,2,3†], Weijie Wang[1,2,3], Rongtao Zhu[1,2,3], Jian Li[1,2,3], Zhe Xing[5], Siyuan Weng[4], Xinwei Han[4]*, Yuling Sun[1,2,3]*

[1]Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan Province, China;

[2]Institute of Hepatobiliary and Pancreatic Diseases, Zhengzhou University, Zhengzhou 450052, Henan Province, China;

[3]Zhengzhou Basic and Clinical Key Laboratory of Hepatopancreatobiliary Diseases, Zhengzhou 450052, Henan Province, China;

[4]Department of Interventional Radiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan Province, China;

[5]Department of Neurosurgery, The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan Province, China.

[†]These authors contributed equally to this work

***For correspondence:**

ylsun@zzu.edu.cn (YLS);

fcchanxw@zzu.edu.cn (XWH);

liuzaoqu@163.com (ZQL)

Abstract

Background: As the most aggressive tumor, the outcome of pancreatic cancer (PACA) has not improved observably over the last decade. Anatomy-based TNM staging does not exactly identify treatment-sensitive patients, and an ideal biomarker is urgently needed for precision medicine.

Methods: A total of 1280 patients from 10 multi-center cohorts were enrolled. 10 machine-learning algorithms were transformed into 76 combinations, which were performed to construct an artificial intelligence-derived prognostic signature (AIDPS). The predictive performance, multi-omic alterations, immune landscape, and clinical significance of AIDPS were further explored.

Results: Based on 10 independent cohorts, we screened 32 consensus prognostic genes via univariate Cox regression. According to the criterion with the largest average C-index in the nine validation sets, we selected the optimal algorithm to construct the AIDPS. After incorporating several vital clinicopathological features and 86 published signatures, AIDPS exhibited robust and dramatically superior predictive capability. Moreover, in other prevalent digestive system tumors, the 9-gene AIDPS could still accurately stratify the prognosis. Of note, our AIDPS had important clinical implications for PACA, and patients with low AIDPS owned a dismal prognosis, relatively high frequency of mutations and copy number alterations, and denser immune cell infiltrates as well as were more sensitive to immunotherapy. Correspondingly, the high AIDPS group possessed dramatically prolonged survival, and panobinostat might be a potential agent for patients with high AIDPS.

Conclusions: The AIDPS could accurately predict the prognosis and immunotherapy efficacy of PACA, which might become an attractive tool to further guide the stratified management and individualized treatment.

Funding: This study was supported by the National Natural Science Foundation of China (No. 81870457, 82172944).

Keywords: Pancreatic cancer; Machine learning; Biomarker; Multi-omic; Immunotherapy

Introduction

As the most aggressive tumor, pancreatic cancer (PACA) has a 5-year survival rate of only 11% and ranks fourth among tumor-related deaths in the United States (***Siegel et al., 2022***). Due to its insidious clinical manifestations and lack of available early detection and screening tools, 80-85% of PACA patients have progressed or metastasized at the time of detection, and losing the opportunity for surgical resection (***Mizrahi et al., 2020***). Over the past decade, immunotherapy, especially immune checkpoint inhibitors (ICIs) has made encouraging progress in most solid tumors (***Billan et al., 2020***). Unfortunately, ICIs have yielded disappointing clinical results in PACA because of the complex composition and highly suppressive immune microenvironment (***Bear et al., 2020***). In terms of molecularly targeted drugs, PARP inhibitors once shed light on *BRCA*-mutated PACA patients. However, a recent study confirmed that although the PARP inhibitor Olaparib extended progression-free survival of patients (3.8 months vs 7.4 months), the overall survival (OS) was not significantly improved (***Golan et al., 2019***). Reassuringly, the result of a phase Ib multi-center study showed that the CD40 monoclonal antibody APX005M in combination with chemotherapy achieved a 58% response rate in advanced PACA (***O'Hara et al., 2021***). Thus, in the era of precision medicine, it is very urgent to explore novel individualized management and combination therapy strategies to markedly improve the prognosis of PACA patients.

In clinical practice, the decision-making, therapeutic management, and follow-up still rely on the traditional anatomy-based TNM staging system (***Katz et al., 2008***). Although this provides a relatively trustworthy reference for determining whether patients will undergo surgical resection, the high inter- and intra-tumoral heterogeneity of PACA results in a wide range of outcomes even among patients at the same stage (***Liu et al., 2022***). With the advancement of high throughput sequencing and evidence-based medicine, molecular biomarkers such as *BRCA1/2* mutations, *NTRK* fusion, DNA mismatch repair deficiency(dMMR), and microsatellite instability-high (MSI-H) have been gradually brought into clinical guideline (***Wattenberg et al., 2020; Doebele et al., 2020; Le et al., 2017***). However, given the relatively low incidence but extremely high mortality of PACA, coupled with the current lack of optimal biomarkers to guide treatment decisions, patients may be over- or under-treated resulting in heavy socioeconomic burden, serious toxic side effects, or rapid disease progression (***De Dosso et al., 2021***). In response to this problem, numerous multigene panels have been developed to address the wide heterogeneity of PACA and achieve relatively good performance in certain cohorts (***Wang et al., 2022; Tan et al., 2020; Yuan et al., 2021***). Considering that these prognostic models were based on the expression files of mRNAs, miRNAs, or lncRNAs in a specific pathway (e.g., immunity, metabolic reprogramming, m6A methylation), data utilization is insufficient (***Wang et al., 2022; Tan et al., 2020; Yuan et al., 2021***). In addition, due to uniqueness and inappropriateness of selected modelling methods, and the lack of strict validation in large multi-center cohorts, expression-based multigene signatures have great shortcomings thereby limiting their wide application in clinical settings (***Yokoyama et al., 2020***).

To develop an ideal biomarker, based on 32 consensus prognosis genes, we constructed and multi-center validated a 9-gene artificial intelligence-derived prognostic signature (AIDPS) via 76 machine-learning algorithm-combinations. In 10 independent cohorts, AIDPS exhibited robust performance in predicting OS, relapse-free survival (RFS), immunotherapy and drug treatment efficacy. After incorporating several vital clinicopathological features and 86 published signatures of PACA, our AIDPS also demonstrated stable and dramatically superior predictive capability. In addition, in other common digestive system tumors such as liver hepatocellular carcinoma (LIHC), stomach adenocarcinoma (STAD), colon adenocarcinoma (COAD) and rectum adenocarcinoma (READ), the AIDPS could still accurately stratify the prognosis. Overall, our study provides an important reference for achieving early diagnosis, prognostic evaluation, stratified management, individualized treatment, and follow-up of PACA in clinical practice.

Materials and methods

Data acquisition and preprocessing

Our workflow is outlined in ***Figure 1***. We collected datasets from The Cancer Genome Atlas (TCGA, http://portal.gdc.cancer.gov/), International Cancer Genome Consortium (ICGC, http://dcc.icgc.org/), ArrayExpress (https://www.ebi.ac.uk/arrayexpress/), and Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) public databases according to the following procedure: (1) more than 40 samples with survival information; (2) at least 15,000 clearly annotated genes; (3) patients with primary tumors and no other treatments were given before resection. Finally, we enrolled 1280 samples from 10 cohorts, TCGA-PAAD (n =176), ICGC-PACA-AU-Seq (PACA-AU-Seq, n =81), ICGC-PACA-AU-Array (PACA-AU-Array, n =267), ICGC-PACA-CA (PACA-CA-Seq, n =182), E-MTAB-6134 (n =288), GSE62452 (n =65), GSE28735 (n =42), GSE78229 (n =49), GSE79668 (n =51), and GSE85916 (n =79). The FPKM data in the TCGA-PAAD was downloaded from UCSC Xena database (https://xenabrowser.net/datapages/) and further converted into log2 (TPM+1) format. The RNA-Seq data of ICGC were downloaded from its portal and converted into log2 (TPM+1) format. The normalized exp-Array data from ICGC, ArrayExpress and GEO were generated directly from their portal. Detailed clinical and pathological information of these 10 cohorts is presented in ***Figure 1-source data 1***.

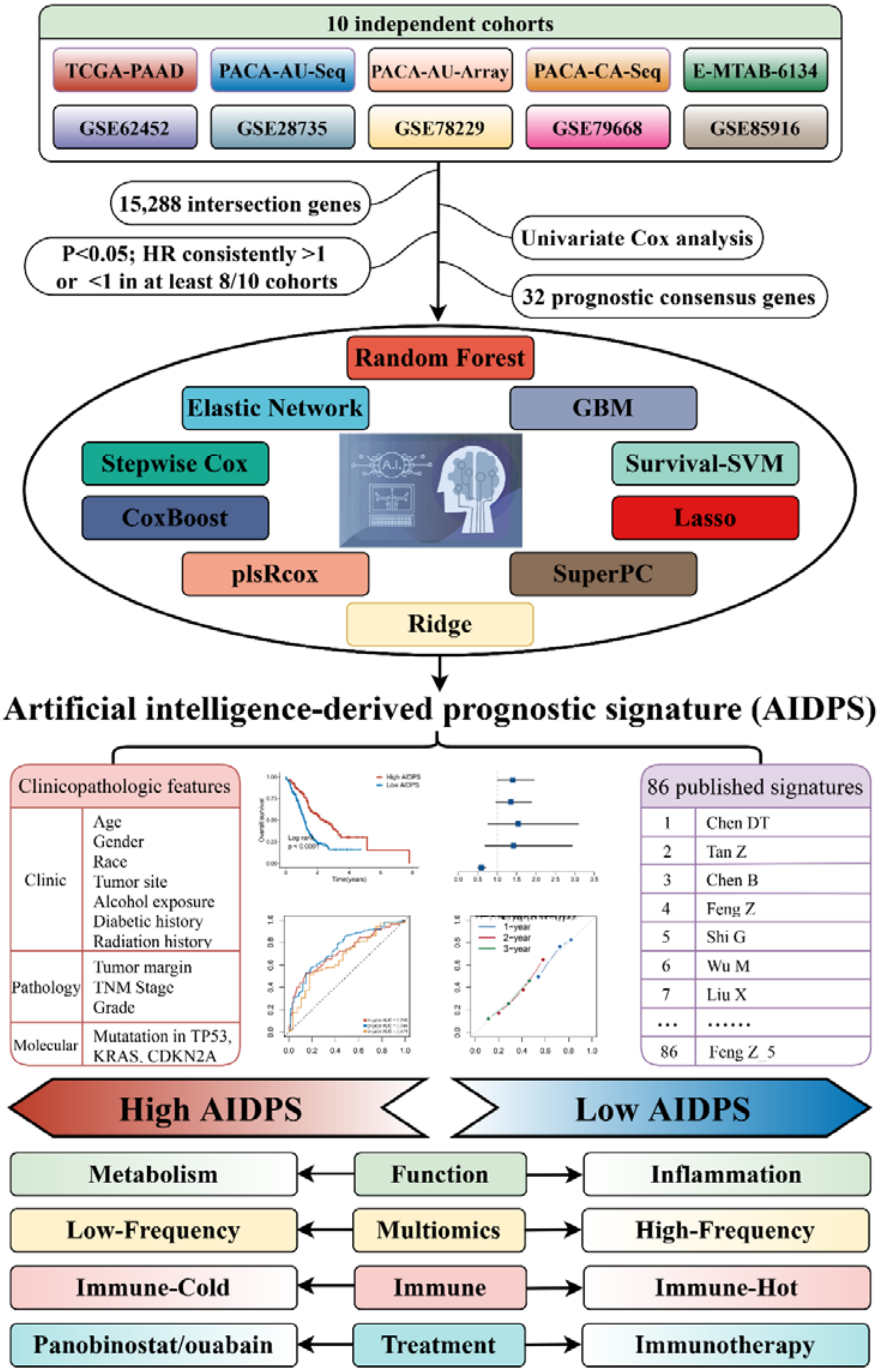


Figure 1. The workflow of our research.

Source data 1. Details of baseline information in 10 public datasets.

Univariate Cox regression analysis

Based on intersection genes, we performed univariate Cox regression analysis in 10 cohorts. We selected consensus prognosis genes (CPGs) for the next study according to the following criteria: $P < 0.05$ and all hazard ratios (HRs) consistently >1 or <1 in more than 8/10 cohorts.

Artificial intelligence-derived prognostic signature

To construct a consensus prognosis model for PACA, we performed our previous workflow (***Liu et al., 2022; Liu et al., 2022***). (1) First, we integrated 10 classical combinations algorithms: random forest (RSF), least absolute shrinkage and selection operator (LASSO), gradient boosting machine (GBM), survival support vector machine (Survival-SVM), supervised principal components (SuperPC), ridge regression, partial least squares regression for Cox (plsRcox), CoxBoost, Stepwise Cox, and elastic network (Enet). Among them, RSF, LASSO, CoxBoost and Stepwise Cox have the function of dimensionality reduction and variable screening, and we combined them with other algorithms into 76 machine-learning algorithm-combinations. (2) Next, we utilized the PACA-AU-Array with larger sample size in ICGC as the training set, and used these 76 combinations to construct signatures separately in the expression files with 32 CPGs. (3) Finally, in the nine validation sets, we calculated the AIDPS score for each cohort using the signature obtained in the training set. Based on the average C-index of the nine validation sets, we finally picked the best consensus prognosis model for PACA.

Validating the prognostic value of AIDPS in 11 datasets

Patients in 10 cohorts and Meta-Cohort were categorized into high and low AIDPS groups according to the median value. The prognostic value of AIDPS was evaluated by Kaplan-Meier curve and multivariate Cox regression analysis. The calibration curve and receiver operator characteristic (ROC) curve were plotted to assess the predictive accuracy of AIDPS.

Collection and calculation of PACA published signatures

With the attention paid to the stratified management and precise treatment of PACA, many prognostic signatures have been constructed, including m6A-related lncRNA signature, metabolic reprogramming-related signature, and SMAD4-driven immune signature, etc. (***Wang et al., 2022; Tan et al., 2020; Yuan et al., 2021***). To compare the predictive performance of AIDPS and these published signatures, we systematically searched PubMed for published prognostic model articles up to January 1, 2022. Afterwards, we calculated their AIDPS scores in the 11 cohorts based on the genes and coefficients provided by the article, and comprehensively evaluated their prognostic performance in PACA by univariate Cox analysis and C-index.

Evaluating of clinical significance of AIDPS

We compared the differences in several pivotal clinical traits such as age, gender, TNM stage, and grade between the high and low AIDPS groups. In addition, to explore the application value of AIDPS in other prevalent gastrointestinal tumors, we acquired the mRNA expression and survival data of LIHC, STAD, COAD and READ in the same way as TCGA-PAAD, and further performed Kaplan-Meier analysis.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was applied to identify specific functional pathways in the high and low AIDPS groups. After differential analysis using the *DESeq2* package, all genes were ranked in descending order according to log2FoldChange (log2FC). Next, we identified GO and KEGG enriched pathways by the *clusterProfiler* package, and further selected the top five pathways in normalized enrichment score (NES) for visualization.

Genomic alteration landscape

To investigate the genomic alteration landscape in the high and low AIDPS groups, we performed a comprehensive analysis of mutation and copy number alteration (CNA) data in the TCGA-PAAD. (1) After obtaining the raw mutation file using the *TCGAbiolinks* package, we calculated the tumor mutation burden (TMB) of each sample and visualized the top 15 genes through the *maftools* package; (2) As described in Lu X et al. (***Lu et al., 2021***), we applied the *deconstructSigs* package to extract the mutational signatures for each PACA patients, and selected mutational signature 1 (age-related), mutational signature 2 (*APOBEC* activity-related), mutational signature 3 (*BRCA1/2* mutations-related), and mutational signature 6 (dMMR-related) with higher frequency of occurrence in PACA for visualization (***Alexandrov et al., 2013***); (3) Recurrent amplified or deleted genome regions were decoded and localized through GISTIC 2.0 module in Firebrowse tool (http://firebrowse.org/). We finally selected regions with broad-level CNA frequency >20% and several genes located within chromosomes 8q24.21, 9p21.3, and 18q21.2 for display.

Estimation of methylation-driven events

Following the pipeline developed in previous studies (***Liu et al., 2021; Liu et al., 2021***), we identified methylation driver genes (MDGs) for TCGA-PAAD. Furthermore, we compared the differences in the methylation levels and mRNA expression levels of MDGs in the two groups, and further evaluated the effect of MDGs methylation levels on the prognosis by Kaplan-Meier survival curve.

Immune molecule expression and tumor microenvironment evaluation

The single sample gene set enrichment analysis (ssGSEA) was utilized to comprehensively infer the infiltration abundance of immune and stromal component in the high and low AIDPS groups (***Wang et al., 2022; Liu et al., 2021***). In addition, we recruited 27 immune checkpoint molecules (ICMs) from our previous study including *B7-CD28* family, *TNF* superfamily, etc., and measured their expression levels between the two groups (***Wang et al., 2022***) .

Response to immunotherapy

Tumour Immune Dysfunction and Exclusion (TIDE) web tool was used to predict responsiveness to ICIs between the high and low AIDPS groups, and lower TIDE scores suggested better immunotherapeutic efficacy (***Jiang et al., 2018***). Additionally, we applied the Subclass Mapping (*Submap*) to calculate the expression similarity between patients in the high and low AIDPS groups and patients who responded/non-responded to ICIs , and then speculated immunotherapy efficacy (***Hoshida et al., 2007***).

Development and validation of potential therapeutic agents

As shown in ***Figure 8D***, we developed potential therapeutic agents for high AIDPS group according to the protocol of Yang C et al (***Yang et al., 2021***). (1) First, we acquired drug sensitivity data for cancer cell lines (CCLs) from the Cancer Therapeutics Response Portal (CTRP) as well as PRISM repurposing datasets, and expression data of CCLs from the Cancer Cell Line Encyclopedia (CCLE) database. (2) The CTRP and PRISM datasets own AUC (area under the ROC curve) values, and lower AUC values suggest increased sensitivity to this compound. Moreover, as a first-line chemotherapeutic drug for PACA, we further selected gemcitabine to verify the scientific and rigor of this approach. (3) Based on Wilcox rank sum test, we performed a differential analysis of drug response between the high AIDPS (first 10%) and low AIDPS (last 10%) groups, and the threshold log2FC >0.2 was set to identify compounds with lower AUC values in the high AIDPS group. (4) Next, we applied

Spearman correlation to further screen compounds with AUC values that had negative correlation coefficients with AIDPS (setting threshold R <-0.4). (5) Finally, we identified potential drugs for patients in the high AIDPS group by the intersection of the compounds obtained from (3) and (4).

The Connectivity Map (CMap, https://clue.io/) is a publicly available web tool for discovering candidate compounds that may target AIDPS related pathways based on gene expression signature (***Subramanian et al., 2017; Malta et al., 2018***). Based on differential expression analysis, we identified potential compounds in PACA using CMap to further validate the results obtained from the CTRP and PRISM databases.

Statistical analysis

All data cleaning, analysis, and result visualization for this research was performed in R 4.1.2. Continuous variables were analyzed by Wilcox rank sum test or Student's t-test. Categorical variables were statistically compared using Chi-square test or Fisher's exact test. The *survival* package was used for univariate, multivariate Cox as well as Kaplan-Meier survival analysis. The *timeROC* package and *rms* package were utilized to plot ROC curve and calibration curve, respectively. The *MethylMix* package was applied to identify MDGs. P-value (two-sided) <0.05 was considered statistically significant.

Results

Integrated development of a pancreatic cancer consensus signature

Based on univariate Cox regression, we screened 32 CPGs from 15,288 intersection genes in 10 cohorts (***Figure 2B***). Next, these 32 CPGs were further incorporated into our integration program to develop an artificial intelligence-derived prognostic signature (AIDPS). In the PACA-AU-Array training set, we applied 76 algorithm-combinations via 10-fold cross-validation to construct prediction models and calculated the average C-index of each algorithm in the remaining nine validation sets. As shown in ***Figure 2A***, the combination of CoxBoost and Survival-SVM with the highest average C-index (0.675) was selected as the final model. We further calculated AIDPS scores of each sample in 10 cohorts according to the expression files of nine genes included in the AIDPS (***Figure 2-source data 1***).

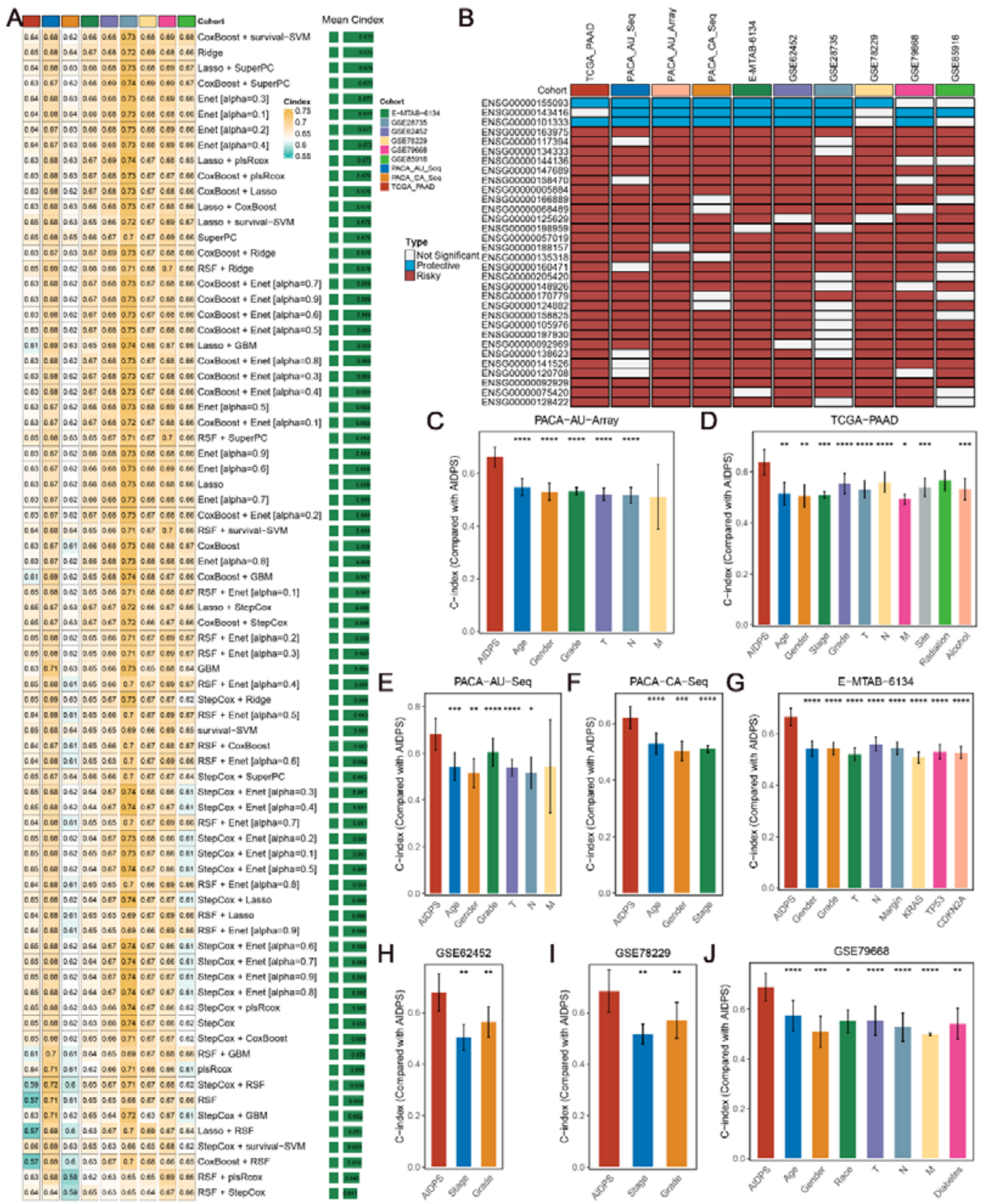


Figure 2. Construction and validation of the artificial intelligence-derived prognostic signature. (**A**) The C-indexes of 76 machine-learning algorithm-combinations in the nine validation cohorts. (**B**) Discovery of 32 consensus prognosis genes from 10 independent multi-center cohorts. (**C-J**) The predictive performance of AIDPS was compared with common clinical and molecular variables in the PACA-AU-Array (**C**),

TCGA-PAAD (**D**), PACA-AU-Seq (**E**), PACA-CA-Seq (**F**), E-MTAB-6134 (**G**), GSE62452 (**H**), GSE78229 (**I**), GSE79668 (**J**). Z-score test: *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

Source data 1. The 9 genes included in the AIDPS.

Consistent prognostic value of AIDPS

To evaluate the prognostic performance of AIDPS, we categorized PACA patients into high and low AIDPS groups according to the median value. The Kaplan-Meier curve for OS and RFS demonstrated the high AIDPS group possessed significantly longer survival in the PACA-AU-Array training set (P <0.0001 in OS and P =0.012 in RFS, ***Figure 3A and B***). After removing batch effects, the Meta-Cohort combining all 10 cohorts also exhibited the same trend (all P <0.05, ***Figure 3C and D***). In addition, we further enrolled several important clinical traits for multivariate Cox analysis, and the results indicated that AIDPS was an independent protective factor for OS and RFS in the PACA-AU-Array cohort (HR: 0.593 [0.504-0.697] for OS and 0.762 [0.611-0.949] for RFS, both P <0.05, ***Figure 3E and F***). Similar results were also found in the Meta-Cohort (HR: 0.603 [0.531-0.685] for OS and 0.667 [0.552-0.805] for RFS, both P <0.05, ***Figure 3G and H***).

In the nine external validation sets, Kaplan-Meier curves consistently showed a significantly prolonged OS in the high AIDPS group compared with the low AIDPS group (all P <0.05, ***Figure 3-figure supplement 1A-I***). Similarly, the comparison of RFS also demonstrated that patients in the high AIDPS group possessed dramatically lower recurrence rate than low AIDPS group in the TCGA-PAAD (n =69, P =0.029), PACA-CA-Seq (n =113, P =0.0023) and E-MTAB-6134 (n =288, P <0.0001) cohorts (***Figure 3-figure supplement 1J, L and M***). It is worth mentioning that only 28 samples in the PACA-AU-Seq cohort owned complete RFS information. Although Kaplan-Meier analysis showed a corresponding trend, the log-rank test did not reach statistical significance (P =0.063, ***Figure 3-figure supplement 1K***). After adjustment for available clinicopathological features, such as age, gender, TNM stage, grade, surgical margin, history of radiotherapy or alcohol consumption and *KRAS*, *TP53* or *CDKN2A* mutations, multivariate Cox analysis results still indicated that AIDPS was an independent prognostic factor for OS (all P <0.05, ***Figure 3-figure supplement 1N and 2A-F***). Consistently, the multivariate results of RFS also revealed that AIDPS remained statistically significant in the TCGA-PAAD, PACA-CA-Seq, and E-MTAB-6134 cohorts (all P <0.05, ***Figure 3-figure supplement 2G-I***). However, given the small sample size of PACA-AU-Seq cohort, the P value was not statistically significant (P =0.338, ***Figure 3-figure supplement 2J***).

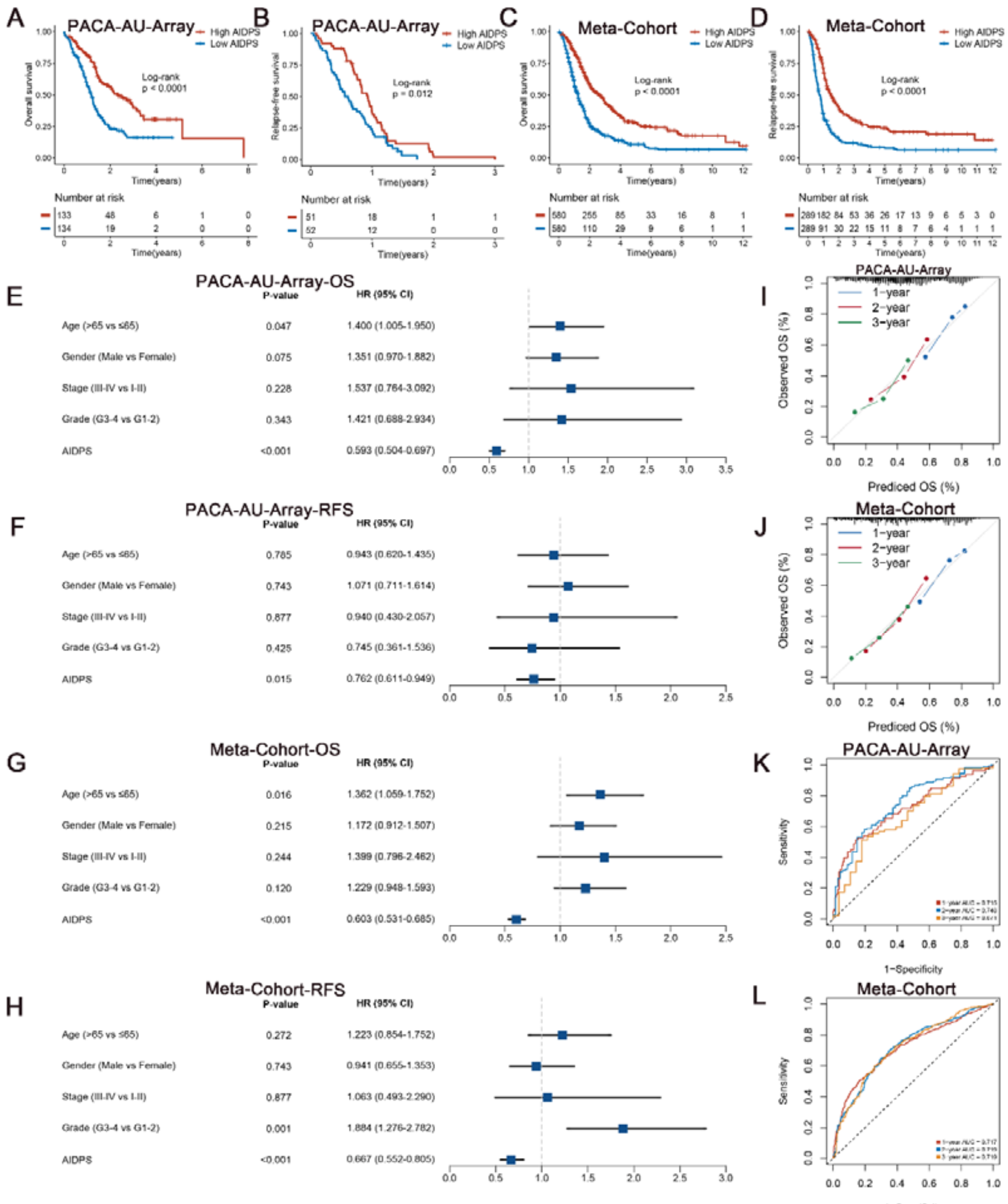


Figure 3. Survival analysis and predictive performance evaluation of AIDPS. (A and B) Kaplan-Meier survival analysis for OS (A) and RFS (B) between the high and low AIDPS groups in the PACA-AU-Array. (C and D) Kaplan-Meier survival analysis for OS (C) and RFS (D) between the high and low AIDPS groups in the Meta-Cohort. (E and F) Multivariate Cox regression analysis of OS (E) and RFS (F) in the

PACA-AU-Array. (**G and H**) Multivariate Cox regression analysis of OS (**G**) and RFS (**H**) in the Meta-Cohort. (**I and J**) Calibration curve for predicting 1-, 2-, and 3-years OS in the PACA-AU-Array (**I**), and Meta-Cohort (**J**). (**K and L**) Time-dependent ROC analysis for predicting 1-, 2-, and 3-years OS in the PACA-AU-Array (**K**), and Meta-Cohort (**L**). OS, overall survival; RFS, relapse-free survival; ROC, receiver operator characteristic.

Figure supplement 1. Survival analysis of AIDPS in remaining nine validation cohorts.

Figure supplement 2. Survival analysis of AIDPS in remaining nine validation cohorts.

Figure supplement 3. Predictive performance evaluation of AIDPS.

Robust predictive performance of AIDPS

To measure the discrimination of AIDPS, we plotted calibration curves and ROC curves. The calibration curves of both the PACA-AU-Array training set and Meta-Cohort showed that AIDPS had good prediction performance (***Figure 3I and J***). The AUCs of 1-, 2-, and 3-year OS were 0.715, 0.748, and 0.671 in the PACA-AU-Array training set and 0.717, 0.719, and 0.719 in the Meta-Cohort (***Figure 3K and L***). Similarly excellent results were found in the nine external validation sets, with 0.705, 0.711, and 0.797 in the TCGA-PAAD, 0.749, 0.808, and 0.827 in the PACA-AU-Seq, 0.662, 0.683, and 0.691 in the PACA-CA-Seq, 0.773, 0.698, and 0.675 in the E-MTAB-6134, 0.676, 0.787 and 0.834 in the GSE62452, 0.734, 0.865, and 0.871 in the GSE28735, 0.669, 0.809, and 0.844 in the GSE78229, 0.791, 0.761, and 0.786 in the GSE79668; and 0.748, 0.766, and 0.811 in the GSE85916, respectively (***Figure 3-figure supplement 3A-I***). The results of AUCs greater than 0.65 in multiple independent cohorts indicated that our AIDPS could stably and robustly predict the prognosis of PACA patients.

In clinical settings, certain clinicopathological features such as surgical margin, stage, and grade are used for prognostic evaluation, clinical stratification management, and treatment decision-making (***Ferrone et al., 2005***). Therefore, we contrasted the predicted efficacy of AIDPS and these common clinical traits in the eight cohorts containing clinical information. The results of C-index indicated that AIDPS had significantly improved accuracy than these features, including age, gender, race, diabetic history, TNM stage, grade, primary location, history of radiotherapy or alcohol consumption, surgical margin, and *KRAS*, *TP53* or *CDKN2A* mutations (***Figure 2C-J***). Overall, our AIDPS might become an attractive tool to further service clinical practice.

Re-evaluation of previously 86 published signatures in PACA

The rapid development of high throughput sequencing has shed light on the stratified management and precise treatment of tumors. In recent years, numerous prognostic signatures of PACA have been constructed via machine-learning algorithms such as LASSO and Stepwise Cox based on large amounts of high-quality data (***Wang et al., 2022; Tan et al., 2020; Yuan et al., 2021***). Therefore, we additionally collected 86 published mRNA/lncRNA prognostic signatures to compare the predictive accuracy of AIDPS and these models (***Figure 4-source data 1***). Signatures containing miRNAs were excluded due to the lack of necessary miRNA expression information. The results of univariate Cox regression showed that only our AIDPS, the 36-gene signature of Haider S, and 20-gene signature of Demirkol CS had consistent statistical significance in all 10 independent cohorts and Meta-Cohort (***Figure 4A***).

We then compared the predictive power of AIDPS and these 86 signatures via C-index across the 10 independent cohorts and Meta-Cohort (***Figure 4B***). Our AIDPS exhibited distinctly superior accuracy than the other models in almost all cohorts (ranked first in 4 cohorts, ranked second in 3 cohorts, and ranked third in 2 cohorts), revealing the robustness of AIDPS. In addition, we noted that although the 36-gene signature of Haider S was first in three cohorts, it performed very poorly in the other cohorts, even lower than 0.6 in the TCGA-PAAD, GSE79668, GSE85916 and Meta-Cohort. While, the 20-gene signature of Demirkol CS was unsatisfactory across all cohorts (***Figure 4B***). Of note, various prognostic signatures owned higher C-index in the TCGA-PAAD training sets (e.g., Zhang C, Xu Q, Li Z, etc.), but performed poorly in other cohorts, which might be due to impaired generalization ability from overfitting (***Figure 4B***). In conclusion, the above results suggested that the 9-gene AIDPS could robustly predict the prognosis of PACA patients, and fewer genes might make it more valuable for clinical promotion.

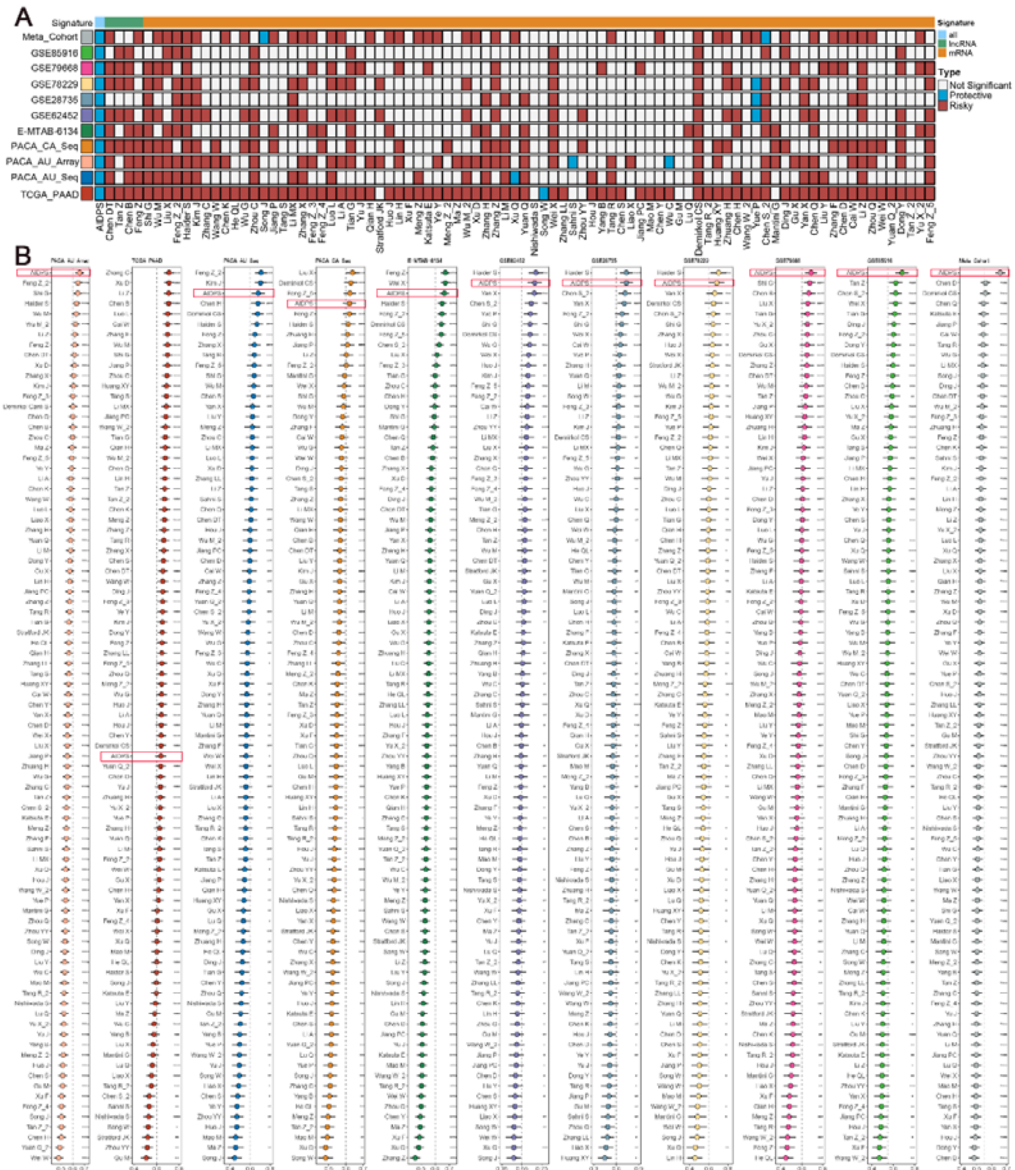


Figure 4. Comparisons between AIDPS and 86 expression-based signatures. (**A**) Univariate Cox regression analysis of AIDPS and 86 published signatures of PACA. (**B**) C-indexes of AIDPS and 86 published signatures in the PACA-AU-Array, TCGA-PAAD, PACA-AU-Seq, PACA-CA-Seq, E-MTAB-6134, GSE62452, GSE28735, GSE78229, GSE79668, GSE85916, and Meta-Cohort. Z-score test: *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

Source data 1. Details of 86 published mRNA/LncRNA signatures in PACA.

The clinical signature of AIDPS

We further compared several familiar clinical characteristics between the high and low AIDPS groups, and the results indicated the absence of statistically difference in age, gender, and TNM stage (***Figure 5A-C and Figure 5-figure supplement 1A-L***). However, patients with low AIDPS possessed more advanced grades, which might contribute to their worse prognosis (***Figure 5D and Figure 5-figure supplement 1M-P***).

In addition, given the excellent predictive power of AIDPS in PACA, we additionally tested its performance in several other common digestive system tumors. As shown in ***Figure 5E-H***, the Kaplan-Meier survival curves exhibited significantly dismal OS for patients in the low AIDPS group in four tumors, including LIHC (P =0.016), STAD (P =0.037), COAD (P =0.032), and READ (P =0.026). These results supported our hypothesis, suggesting that AIDPS constructed in PACA, as a biomarker, has broad prospects for generalization to other tumors.

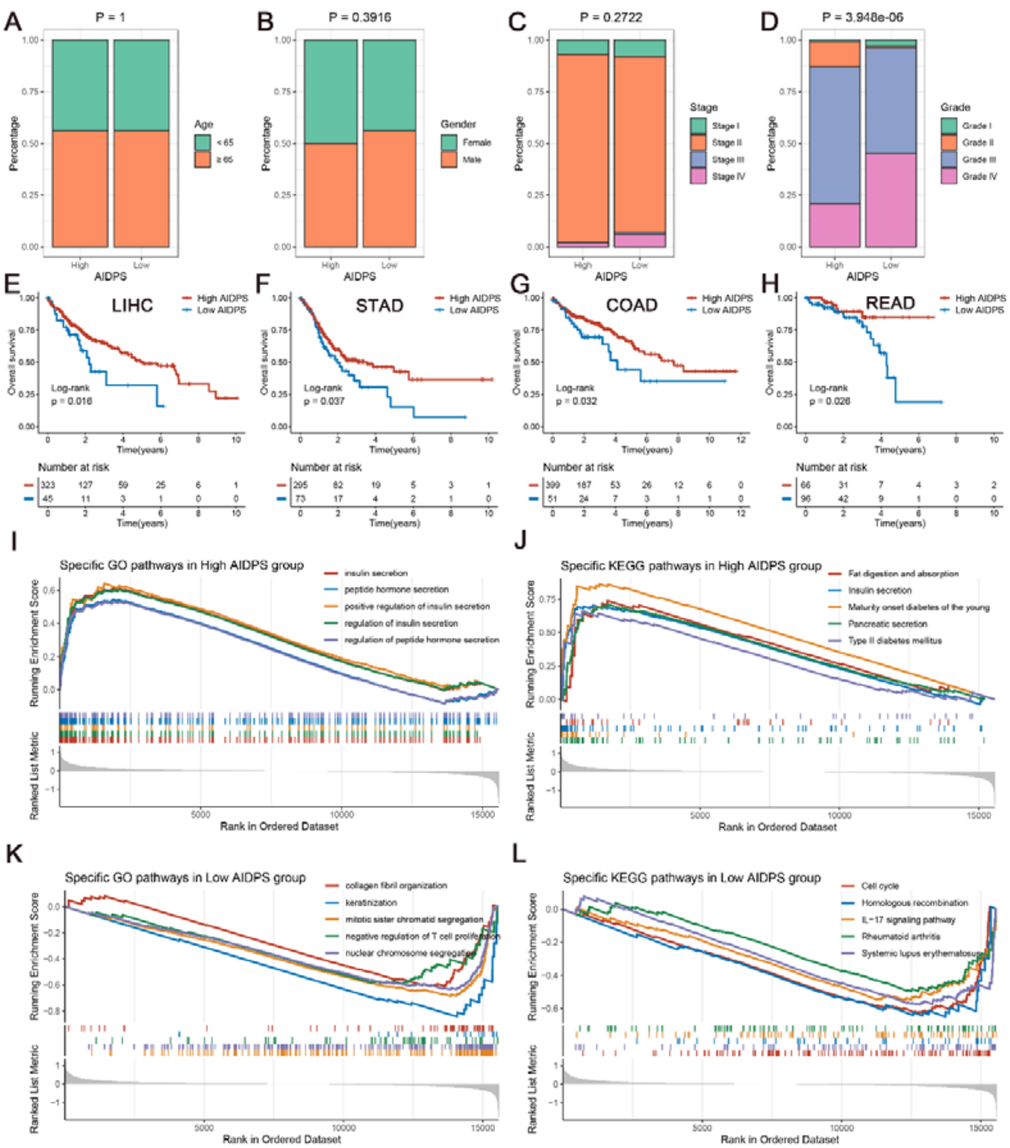


Figure 5. The clinical signature and functional characteristics of the high and low AIDPS groups. (**A-D**) Composition percentage of the two groups in clinical characteristics such as age (**A**), gender (**B**), stage (**C**), grade (**D**) in the PACA-AU-Array. (**E-H**) Kaplan-Meier survival analysis for OS in the TCGA-LIHC (**E**), TCGA-STAD (**F**), TCGA-COAD (**G**), TCGA-READ (**H**). (**I and J**) The top five GO enriched pathways (**I**) and KEGG enriched pathways (**J**) in the high AIDPS groups. (**K and L**) The top five GO enriched pathways (**K**) and KEGG enriched

pathways (**L**) in the low AIDPS groups. OS, overall survival; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Figure supplement 1. The clinical characteristics of the high and low AIDPS groups.

The underlying biological mechanisms of AIDPS

GSEA analysis was applied to elucidate the potential functional pathways of AIDPS. As illustrated in ***Figure 5I-J***, the high AIDPS group was remarkably enriched for digestive and metabolism-related pathways, such as insulin secretion and regulation, peptide hormone secretion and regulation, fat digestion and absorption, pancreatic secretion, maturity onset diabetes of the young, and type II diabetes mellitus. While, the low AIDPS group was predominantly correlated with the regulation of T cell proliferation, *IL-17* signaling pathway, and other immune-related pathways, as well as cell cycle, nuclear chromosome segregation, homologous recombination, and other proliferation-related biological processes, which partly explained its more advanced grades and worse prognosis (***Figure 5K and L***).

Genome alteration landscape of AIDPS

To investigate genomic heterogeneity between the high and low AIDPS groups, we performed a comprehensive analysis of mutations and CNA (***Figure 6A***). As shown in ***Figure 6C***, the low AIDPS group possessed observably higher TMB. Combining the 10 oncogenic signaling pathways in TCGA (***Sanchez-Vega et al., 2018***), we found that the classical tumor suppressor genes *TP53*, *CDKN2A*, and oncogene *KRAS* were more frequently mutated in the low AIDPS group than high AIDPS group, whereas the opposite was true for *SMAD4*, *TTN*, and *RNF43* (***Figure 6A and B***). Next, based on the popular mutational signatures in PACA, we discovered that mutational signature 3 (*BRCA1/2* mutations-related) was enriched in the high AIDPS group, while mutational signature 1 (age-related) was more dominant in low AIDPS group.

In addition, we further explored the CNA landscape of the two groups. Compared to the high AIDPS group, the low AIDPS group owned evidently higher amplification or deletion in the focal and chromosome arm levels, such as the amplification of 8q24.21, 19q13.2, and 8p11.22 as well as deletion of 9p21.3, 18q21.2, 6p22.2, and 22q13.31 (***Figure 6A and D***). This result was again corroborated in gene level by the obvious amplification of the oncogene *MYC* within 8q24.21, and the distinct deletion of the tumor suppressor genes *CDKN2A*, *CDKN2B* and *SMAD4* within 9p21.3 and 18q21.2 (***Figure 6A***). Overall, oncogenes amplification and tumor suppressor genes deletion in the low AIDPS group might contribute to their poor prognosis.

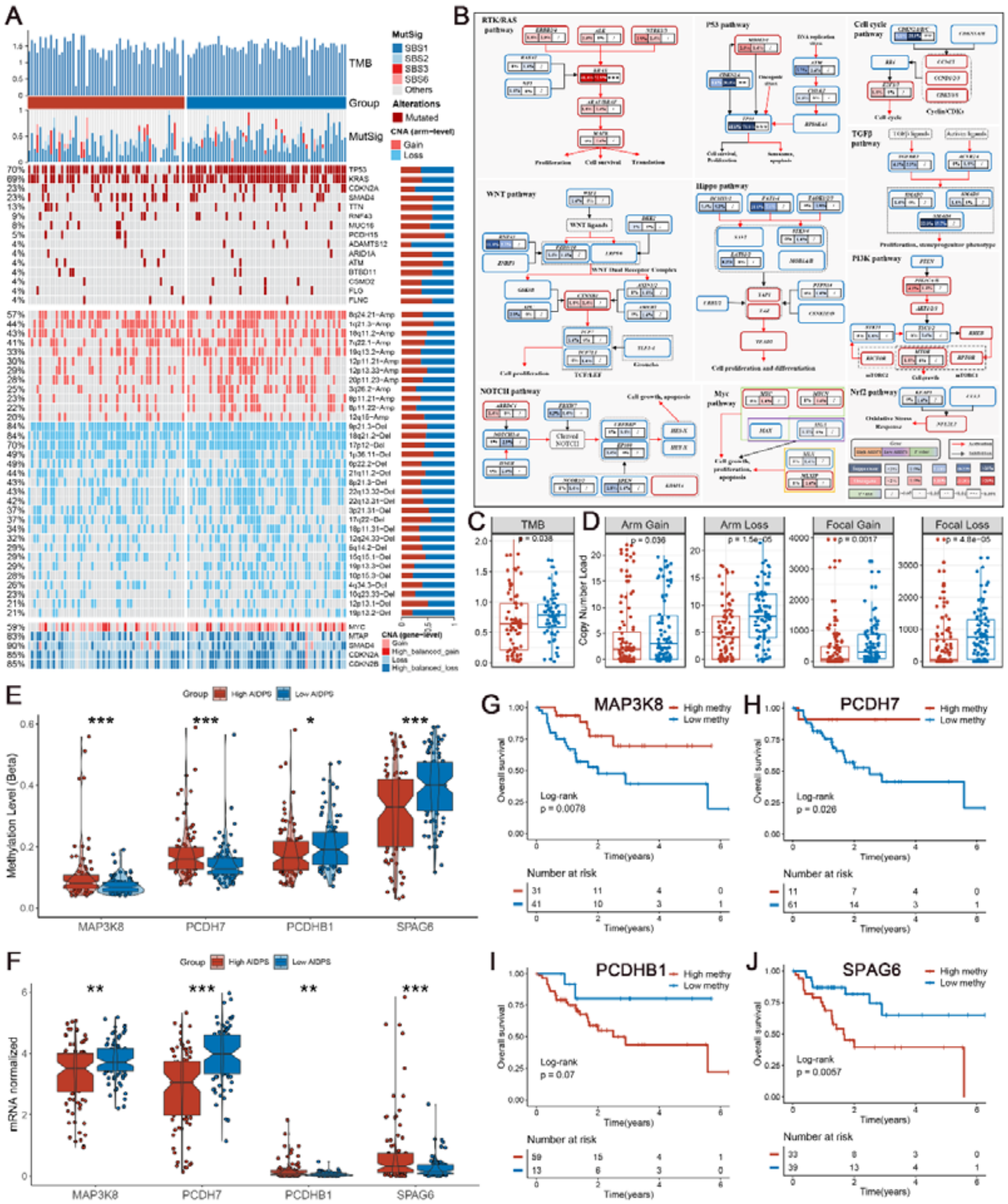


Figure 6. Multi-omics analysis based on mutation, copy number alteration (CNA) and methylation. (**A**) Genomic alteration landscape according to AIDPS. Tumour mutation burden (TMB), relative contribution of four mutational signatures, top 15 mutated genes and broad-level copy number alterations (>20%), and selected genes located within chromosomes 8q24.21, 9p21.3, and 18q21.2 are shown from the top to the

bottom panels. The proportion of the high and low AIDPS groups in each alteration is presented in the right bar charts. (**B**) Comprehensive comparison of mutation landscapes in 10 oncogenic signaling pathways across the high and low AIDPS groups. Genes are mutated at different frequencies (color intensity indicates the mutation frequency within the entire dataset) by oncogenic mutations (red) and tumor suppressor mutations (blue). Each gene box includes two percentages representing the mutation frequency in the high and low AIDPS groups, and another box representing the statistical p value. Genes are grouped by signaling pathways, with edges showing pairwise molecular interactions. (**C**) Comparison of the two groups in TMB. (**D**) Comparison of the two groups in arm and focal CNA burden. (**E and F**) Boxplot of DNA methylation level (**E**) and mRNA expression level (**F**) for methylation-driven genes in the high and low groups. (**G-J**) Kaplan-Meier survival analysis between the high and low methylation groups in the *MAP3K8* (**G**), *PCDH7* (**H**), *PCDHB1* (**I**), and *SPAG6* (**J**). *P <0.05, **P <0.01, ***P <0.001.

Methylation-driven events of AIDPS

Referring to our previous process (***Liu et al., 2021; Liu et al., 2021***), we screened four MDGs whose methylation levels were significantly inversely related to matched gene expression levels in PACA. Compared to the low AIDPS group, the high AIDPS group possessed higher *MAP3K8* and *PCDH7* methylation levels as well as significantly lower mRNA expression levels, while the opposite was true for *PCDHB1* and *SPAG6* (***Figure 6E and F***). Furthermore, the Kaplan-Meier analysis showed that higher methylation levels of *MAP3K8* and *PCDH7* and lower methylation level of *SPAG6* brought significantly prolonged OS for high AIDPS group (all P <0.05, ***Figure 6G, H and J***). *PCDHB1* also exhibited a similar trend with *SPAG6*, although statistical significance was not reached (P =0.07, ***Figure 6I***).

Immune landscape and molecular expression of AIDPS

The above GSEA revealed that several immune-related pathways were highly enriched in the low AIDPS group, and we consequently investigated the immune landscape and ICMs molecular expression between the two groups. According to ssGSEA, we found that the low AIDPS group exhibited a relatively higher infiltration abundance of immune cell types, including activated CD4+T cells, CD56dim natural killer cells, Central memory CD8+T cells, Gamma delta T cells, and Type 2 T helper cells (all P <0.05, ***Figure 7A and B***). In addition, among the 27 ICMs we included, the low AIDPS group had dramatically increased relative expression level, such as *CD274*, *CD276*, *PDCD1LG2*, *CD40*, *CD70*, *TNFRSF18*, *TNFRSF4*, *TNFRSF9*, and *NT5E* (***Figure 7C***). Together, these results consistently indicated that PACA patients with low AIDPS were more likely to respond to immunotherapy.

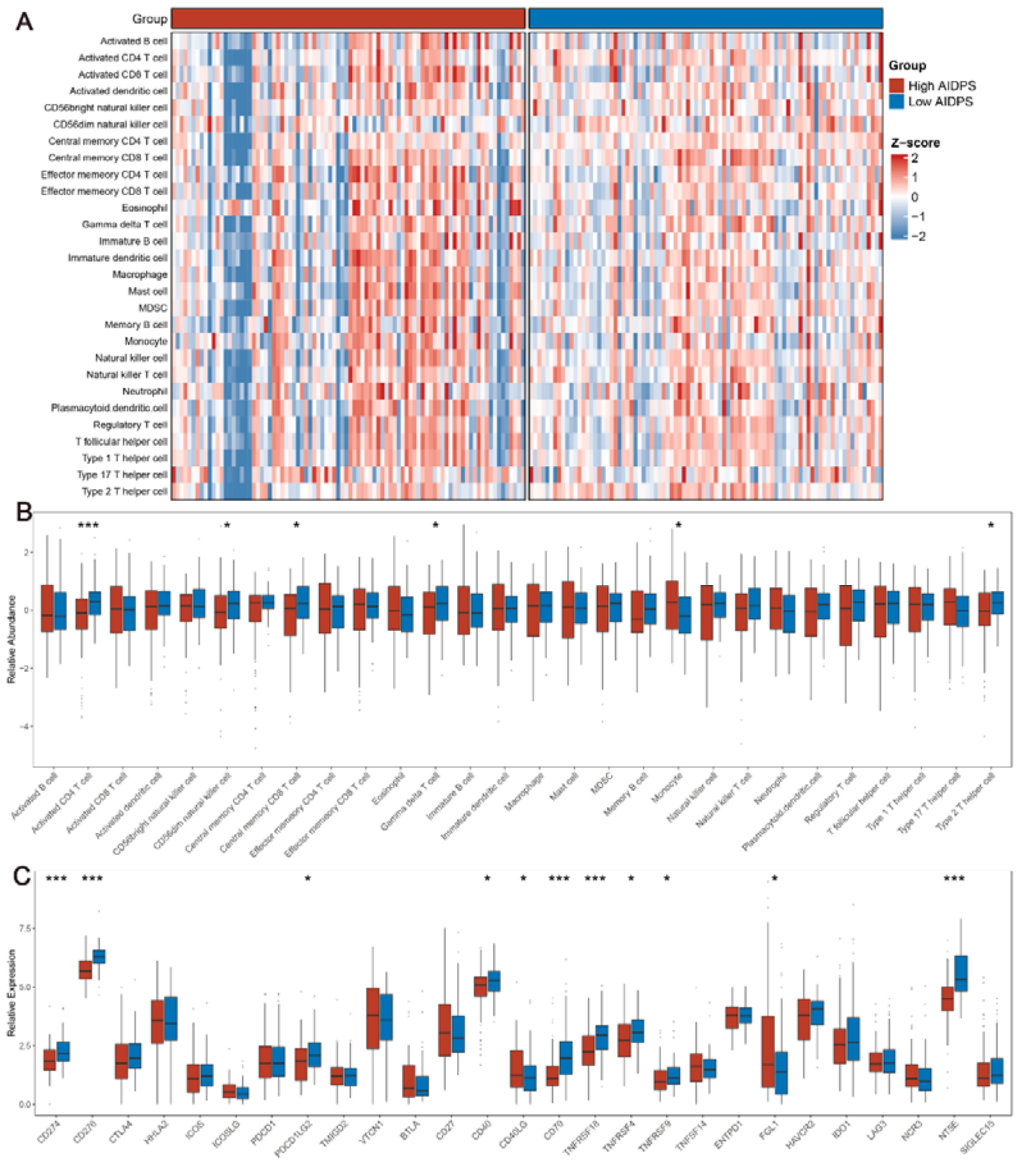


Figure 7. The immune landscape in the high and low AIDPS groups. (**A**) The heatmap of 28 immune cell types in the high and low AIDPS groups. (**B**) Boxplot of relative infiltrate abundance of 28 immune cells in patients with high and low AIDPS groups. (**C**) Boxplot of relative expression levels at 27 immune checkpoints profiles between the high and low AIDPS patients. *P <0.05, **P <0.01, ***P <0.001.

Predictive value of AIDPS for immunotherapy

Given that patients in the low AIDPS group possessed higher genomic variation frequency and TMB, combined with their relatively activated tumor microenvironment (TME) and increased ICMs expression, we speculated that PACA patients with low AIDPS were more sensitive to immunotherapy. Based on the TIDE web tool, the low AIDPS group resulted in significantly lower TIDE scores and higher immunotherapy response rates (***Figure 8A and B***). The results of the Submap also suggested that expression patterns of patients with low AIDPS was more similar to those of melanoma patients who responded to ICIs (***Figure 8C***). Overall, these results demonstrated that the low AIDPS group was more likely to benefit from immunotherapy.

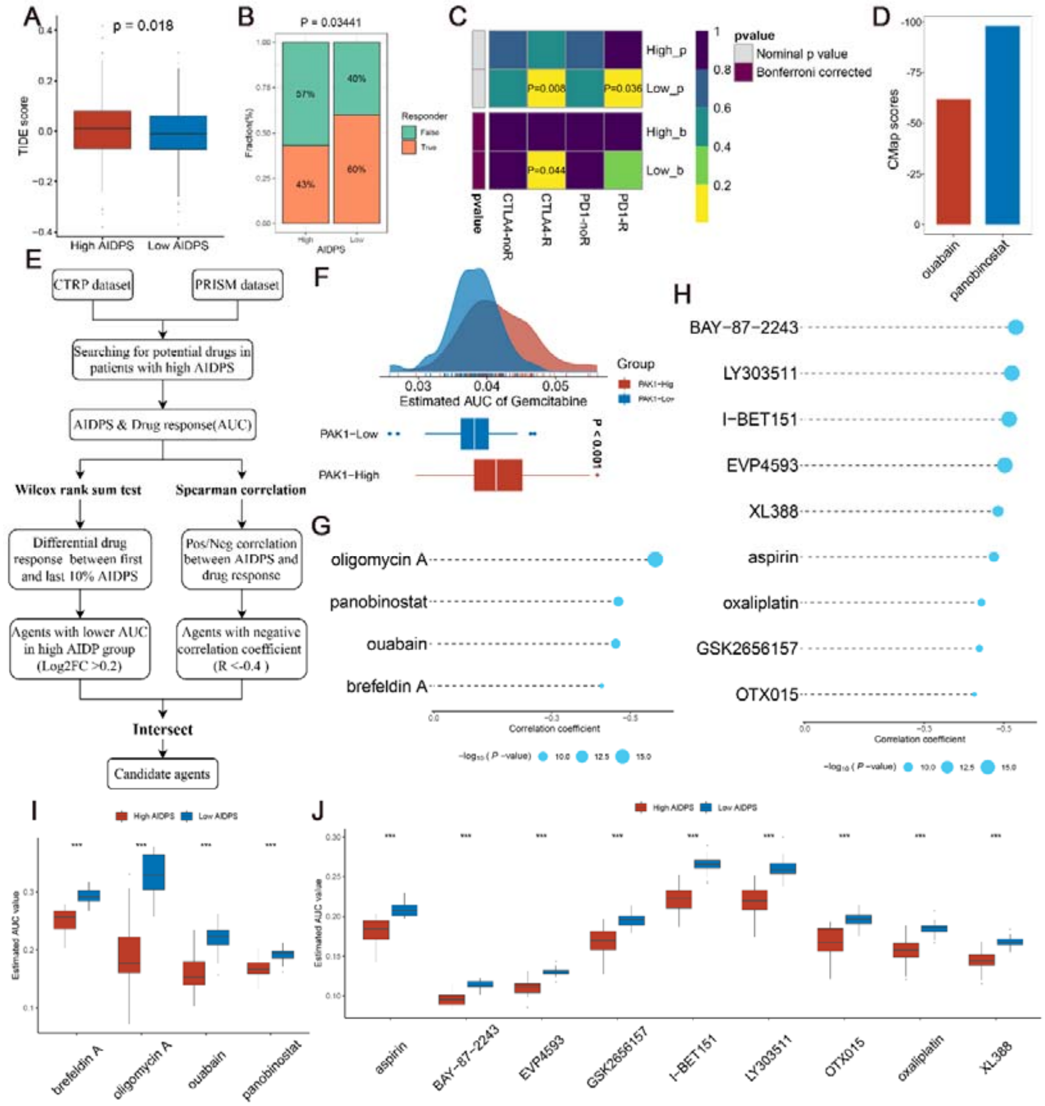


Figure 8. Evaluating therapeutic drug benefits. (**A**) Boxplot of TIDE score between the high and low AIDPS groups. (**B**) Percentage of immunotherapy responses at high and low AIDPS groups. (**C**) Submap analysis of the two groups and 47 pretreated patients with comprehensive immunotherapy annotations. For Submap analysis, a smaller p-value implied a more similarity of paired expression profiles. (**D**) Barplot of ouabain and panobinostat CMap scores in patients with high AIDPS. (**E**) Schematic outlining the strategy to develop potential therapeutic agents with higher drug

sensitivity in the high AIDPS group. (**F**) Comparison of estimated gemcitabine's sensitivity between high and low *PAK1* expression groups. (**G and H**) The results of Spearman's correlation analysis of CTRP-derived compounds (**G**) and PRISM-derived compounds (**H**). (**I and J**) The results of differential drug response analysis of CTRP-derived compounds (**I**) and PRISM-derived compounds (**J**), the lower values on the y-axis of boxplots imply greater drug sensitivity. TIDE, Tumour Immune Dysfunction and Exclusion; CMap, Connectivity Map; CTRP, Cancer Therapeutics Response Portal; PRISM, profiling relative inhibition simultaneously in mixtures. *P <0.05, **P <0.01, ***P <0.001.

Searching for potential therapeutic agents for high AIDPS group

As illustrated in ***Figure 8E***, we developed potential agents for PACA patients with high AIDPS using sensitivity data from CTRP (includes 481 compounds over 835 CCLs) and PRISM (includes 1448 compounds over 482 CCLs) datasets (***Yang et al., 2021***). To ensure the reliability of our protocol, gemcitabine, as a first-line treatment for PACA, was employed to investigate whether the estimated sensitivity and clinical practice were concordant. A laboratory study found that elevated *PAK1* activity was required for gemcitabine resistance in PACA, and that *PAK1* inhibition enhanced the efficacy of gemcitabine. Consistent with this study, our results revealed that patients with low *PAK1* expression possessed distinctly lower estimated AUC values, suggesting greater sensitivity to gemcitabine (***Figure 8F***). Afterwards, we applied this formula to identify potentially sensitive agents for high AIDPS group, and finally generated four CTRP-derived agents (brefeldin A, oligomycin A, ouabain, and panobinostat) and nine PRISM-derived agents (aspirin, BAY-87-2243, EVP4593, GSK2656157, I-BET151, LY303511, OTX015, oxaliplatin, and XL388). The estimated AUC values of these agents were not only statistically negatively correlated with AIDPS scores, but also significantly lower in the high AIDPS group (***Figure 8G-J***).

In addition, based on the differential expression profiles between PACA patients and normal controls, we further applied the CMap tool to identify candidate compounds for PACA. After taking the intersection with the results obtained by CTRP and PRISM, we ended up with two candidate compounds: ATPase inhibitor ouabain and Histone Deacetylase (HDAC) inhibitor panobinostat. Among them, panobinostat with a -98.11 CMap score was highly sensitive in PACA patients, suggesting that it could become a potential therapeutic agent for PACA patients in the high AIDPS group (***Figure 8D***).

Discussion

Over the past 20 years, the incidence of PACA has increased by 0.5% to 1.0% per year, but the 5-year survival rate only improved from 5.26% to 10%, without a significant breakthrough (***Park et al., 2021***). The lack of available biomarkers for screening, stratified management, and prognostic follow-up has been an urgent problem for clinicians and researchers, which may lead to over- or under-treatment. To bridge these gaps, we constructed and validated a 9-gene AIDPS via 76 machine-learning algorithm-combinations in 10 independent multi-center cohorts. Compared to several common clinicopathological features and 86 published signatures of PACA, our AIDPS demonstrated robust and superior predictive capacity. Moreover, we substantiated that the low AIDPS group had a dismal prognosis, more advanced grades, enrichment in immune and proliferation-related pathways, higher frequency mutations and CNAs, relatively activated TME and increased ICMs expression, and better immunotherapeutic efficacy. While the high AIDPS group owned remarkably prolonged OS and RFS, a significant enrichment of metabolism-related pathways and mutational signature 3 (*BRCA1/2* mutations-related), was more sensitive to the ATPase inhibitor ouabain and HDAC inhibitor panobinostat. Therefore, in clinical settings, our AIDPS may become a reliable platform to further serve individualized decision-making in PACA.

In the era of precision medicine, anatomy-based TNM staging is far from meeting the needs of clinicians for ideal biomarkers that could accurately evaluate the prognosis and predict the efficacy of PACA patients (***Katz et al., 2008; Liu et al., 2022***). Recently, numerous prognostic signatures of PACA have been constructed, but most of them are based on a specific biological pathway such as immunity, metabolic reprogramming, and m6A methylation (***Wang et al., 2022; Tan et al., 2020; Yuan et al., 2021***). This ignored information about other biological processes that played a crucial role in the oncogenesis and progression of PACA. In this study, based on 15,288 intersection genes in 10 independent cohorts, we further obtained 32 CPGs via univariate Cox regression analysis. In addition, in the existing studies, people mostly choose the modelling algorithms based on their knowledge limitations and preferences (***Liu et al., 2022; Liu et al., 2022***). To remedy this shortcoming, we collected 10 popular machine-learning algorithms that could be used to construct biomedical prognostic signatures. Among them, RSF, LASSO, CoxBoost and Stepwise Cox have the function of variable screening and data dimensionality reduction, so we further combined them into 76 algorithm combinations. Ultimately, the combination of CoxBoost and Survival-SVM with the largest average C-index (0.675) in the remaining nine external validation sets was identified as the final model.

Overfitting is one of the troublesome problems encountered by artificial intelligence and machine learning in biomedical model development, with several models fitting well in the training sets but poorly in other external validation cohorts (***Deo, 2015***). After minimizing redundant information by CoxBoost, we finally obtained a 9-gene signature termed AIDPS through Survival-SVM. The results of Kaplan-Meier analysis, univariate Cox regression, ROC curve and calibration curve all indicated that AIDPS had excellent predictive performance in the training set, nine external validation sets and Meta-Cohort. Moreover, compared with 86 published PACA signatures, AIDPS exhibited distinctly superior accuracy than the other models in almost all cohorts, revealing the robustness of AIDPS. It should be mentioned that, although the 36-gene signature of Haider S was superior to AIDPS in three cohorts, it performed very poorly in the TCGA-PAAD, GSE79668, GSE85916 and Meta-Cohort, with a C-index even lower than 0.6. While, the 21-gene signature of Zhang C, the

20-gene signature of Demirkol CS, the 14-gene signature of Wei X were better than AIDPS in a certain cohort, but they were unsatisfactory across almost all validation cohorts, which might due to very poor generalization ability from overfitting. In contrast, because contains fewer elements and thus AIDPS has a superior extrapolation possibility, and the findings of 10 independent multi-center cohorts also fully confirm this notion.

In addition, compared with several common clinical and molecular characteristics such as TNM stage, grade, *KRAS*, *TP53* or *CDKN2A* mutations, our AIDPS showed significantly improved accuracy. After stratifying PACA patients into the high and low AIDPS groups, we demonstrated that there was no outstanding difference in age, gender, and TNM stage, but the low AIDPS group had more advanced grades, which also contributed to its worse prognosis to some extent. Furthermore, given the excellent performance of AIDPS in PACA, we additionally measured its performance in four common digestive system tumors: LIHC, STAD, COAD, and READ, and found that AIDPS could accurately stratify patients. These findings indicated that AIDPS constructed in PACA, as a biomarker, has broad prospects for generalization to other tumors.

GSEA functional enrichment analysis was applied to elucidate the underlying biological mechanisms of AIDPS. The low AIDPS group was mostly enriched in immune and proliferation-related pathways (including regulation of T cell proliferation, *IL-17* signaling pathway, cell cycle, homologous recombination, etc.), which partly explained its more advanced grades and worse prognosis. In addition, many recent studies have reported the emerging promise of epigenetic alterations especially DNA methylation in the early diagnosis and prognostic follow-up of PACA (***Yokoyama et al., 2020; Grady et al., 2021***). Therefore, we identified four MDGs in PACA patients. Further investigations found that higher methylation levels of *MAP3K8* and *PCDH7*, and lower methylation levels of *PCDHB1* and *SPAG6* in the high AIDPS all corresponded to obviously prolonged OS, suggesting that methylation modification might play an indispensable role in its better prognosis.

Based on multi-omics data of TCGA-PAAD, we further investigated the genomic heterogeneity with regard to AIDPS. The results showed that the low AIDPS group owned higher TMB and superior mutation frequencies in the classical tumor suppressor genes *TP53*, *CDKN2A*, and oncogene *KRAS*. Numerous studies have revealed that *TP53*, *CDKN2A* and *KRAS* mutations promoted invasion, metastasis and immune escape in PACA patients, resulting in worse prognosis (***Hu et al., 2018; Hashimoto et al., 2019***). In addition, the results of CNA indicated that the low AIDPS group had evidently increased amplification of 8q24.21, 19q13.2, and 8p11.22 as well as deletion of 9p21.3, 18q21.2, 6p22.2, and 22q13.31 than those in the high AIDPS group. A recent study has shown that amplification of 19q13.2 and consequent overexpression of this loci was correlated with impaired survival in PACA (***Sandhu et al., 2016***). Morikawa A et al. also found that amplification of 8p11.22 was associated with chemotherapy resistance and shorter OS in ovarian clear cell carcinoma (***Morikawa et al., 2018***). Baker MJ et al. reviewed that people with 9p21.3 deletion were more susceptible to multiple types of tumors (***Baker et al., 2016***). Previous studies have demonstrated that deletion of 22q13.31 was an early genetic event in insulinoma independent of other genomic variants and related to more advanced stage and dismal prognosis (***Jonkers et al., 2006***). Overall, all this evidence consistently supported our conclusion that genome-driven events might contribute to worse prognosis of the low AIDPS group. On the other hand, relatively high frequency genomic variants and higher TMB also provide more neoantigens, which might point the way for immunotherapy in the low AIDPS group (***Sha et al., 2020***).

Next, we investigated the immune landscape between the high and low AIDPS groups. The results of ssGSEA exhibited that the low AIDPS group possessed superior abundance of immune cell types, including activated CD4+T cells, CD56dim natural killer cells, Central memory CD8+T cells. As everyone knows, these increased effector cells will enhance anti-tumor immunity and bring better immunotherapeutic effects for the low AIDPS group (***Borst et al., 2018; Wu et al., 2013***). Over the past decade, ICIs targeting ICMs have shed light on the treatment of solid tumors (***Billan et al., 2020***). As expected, our findings showed that the expression of ICMs such as *CD274*, *CD276*, *PDCD1LG2* were dramatically elevated in patients with low AIDPS, suggesting that they were more likely to benefit from ICIs treatment. TIDE and Submap are two widely recognized tools for predicting tumor patient sensitivity to ICIs based on expression profile (***Jiang et al., 2018; Hoshida et al., 2007***). Consistently, the results confirmed our previous conclusion that patients with low AIDPS group possessed a greater response rate to immunotherapy. Overall, these findings indicated that our AIDPS might provide a reference for early identification of immunotherapy-sensitive PACA patients receiving first-line immunotherapy.

Precision medicine requires clinicians to identify patients who are sensitive to various treatments as early as possible for further individualized treatment. Therefore, considering the higher sensitivity of the low AIDPS group to immunotherapy, we integrated CTRP, PRISM and CMap databases to develop specific drugs for patients in the high AIDPS group (***Yang et al., 2021; Subramanian et al., 2017; Malta et al., 2018***). Ultimately, an HDAC inhibitor, panobinostat attracted our attention as a potential drug for patients in the high AIDPS group. The latest studies report that panobinostat can synergistically enhance the antitumor effect of the selective Wee1 kinase inhibitor MK-1775 or CAR-T cell therapy in PACA (***Wang et al., 2015; Ali et al., 2021***). In the future, more clinical trials are required to confirm the broad prospects of panobinostat in PACA, especially in patients with high AIDPS.

This study differed from previous studies in the following aspects. (1) We systematically collected 10 large multi-center cohorts, and selected the algorithm with the largest average C-index in the 9 validation sets to construct our AIDPS. (2) Unlike current prognostic models for a certain pathway, our AIDPS was based on 15,288 intersection genes from 10 cohorts, which avoided the omission of other indispensable biological process in the initiation and progression of PACA. (3) In order to prevent the inappropriate modelling methods due to personal preference, we combined 10 recognized machine-learning algorithms into 76 combinations and selected the best model based on their accuracy. While we have tried to be as rigorous and comprehensive as possible in our research, some limitations should be noted. Firstly, although we collected 10 independent multi-center cohorts, further validation in prospective study was warranted. Secondly, in spite of the nine genes included in AIDPS have appeared in numerous prognostic signatures of PACA, which indicated their consistent prognostic value. The roles of them in PACA remain to be elucidated, more functional experimental validation is required in the future. Finally, further clinical trials are necessary to affirm the therapeutic effect of panobinostat in PACA patients with high AIDPS.

Conclusions

In conclusion, based on 32 consensus prognosis genes from 10 independent multi-center cohorts, we constructed and validated a consensus prognostic signature (termed AIDPS) via 76 machine-learning algorithm-combinations. After incorporating several vital clinicopathological features and 86 published signatures, AIDPS also exhibited robust and dramatically superior predictive capability. Of note, our AIDPS has important clinical implications for the clinical management and individualized treatment of PACA, and patients with low AIDPS are more sensitive to immunotherapy, while panobinostat may be a potential agent for patients with high AIDPS. In addition, in other prevalent digestive system tumors, the 9-gene AIDPS could still accurately stratify the prognosis, suggesting a strong possibility of extrapolation. Overall, our study provides an attractive tool for prognostic evaluation, risk stratification, and individualized treatment of PACA patients in clinical practice.

Additional information

Acknowledgements:

We sincerely thank the research group that contributed pancreatic cancer sequencing data as well as the staff who developed the R package.

Competing interests

The authors declare that they have no competing interests.

Funding

Funder	Grant reference number	Author
National Natural Science Foundation of China	81870457; 82172944	Yuling Sun

Author contributions

Libo Wang, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft preparation, Visualization; Zaoqu Liu, Resources, Conceptualization, Methodology, Software, Writing - review & editing, Project administration; Yuling Sun, Writing - review & editing, Supervision, Funding acquisition; Xinwei Han, Writing - review & editing, Supervision; Zhe Xing and Siyuan Weng, Resources, Data curation, Writing - review & editing; Ruopeng Liang, Weijie Wang, Rongtao Zhu, and Jian Li, Writing - review & editing.

Author ORCIDs

Libo Wang http://orcid.org/0000-0003-3745-9459

Zaoqu Liu https://orcid.org/0000-0002-0452-742X

Ruopeng Liang https://orcid.org/0000-0002-3789-1666

Xinwei Han https://orcid.org/0000-0003-4407-4864

Yuling Sun https://orcid.org/0000-0001-5289-4673

Ethics

This research has been conducted using publicly available datasets. De-identified data were used, and no ethical approval was required.

Additional files

Supplementary files

- Figure Supplements

Figure 3-figure supplement 1. Survival analysis of AIDPS in remaining nine validation cohorts.

Figure 3-figure supplement 2. Survival analysis of AIDPS in remaining nine validation cohorts.

Figure 3-figure supplement 3. Predictive performance evaluation of AIDPS.

Figure 5-figure supplement 1. The clinical characteristics of the high and low AIDPS groups.

- Source Data Files

Figure 1-source data 1. Details of baseline information in 10 public datasets.

Figure 2-source data 1. The 9 genes included in the AIDPS.

Figure 4-source data 1. Details of 86 published mRNA/LncRNA signatures in

PACA.

Data availability

All data generated during this study are included in the manuscript and supporting files. Other necessary data and supporting codes will be made available by contacting the corresponding author. Source data files have been provided for Figures 1, 2 and 4.

The following previously published datasets were used:

Author(s)	**Year**	**Dataset title**	**Dataset URL**	**Database and Identifier**
Goldman MJ, Craft B, Hastie M, et al.	2020	GDC TCGA Pancreatic Cancer (PAAD)	https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Pancreatic%20Cancer%20(PAAD)	The Cancer Genome Atlas, TCGA-PAAD
Zhang J, Bajari R, Andric D, et al.	2019	exp_seq. PACA-AU	https://dcc.icgc.org/releases/current/Projects/PACA-AU	International Cancer Genome Consortium, PACA-AU-Seq
Zhang J, Bajari R, Andric D, et al.	2019	exp_array. PACA-AU	https://dcc.icgc.org/releases/current/Projects/PACA-AU	International Cancer Genome Consortium, PACA-AU-Array
Zhang J, Bajari R, Andric D, et al.	2019	exp_seq. PACA-CA	https://dcc.icgc.org/releases/current/Projects/PACA-CA	International Cancer Genome Consortium, PACA-CA-Seq
Puleo F, Nicolle R, Blum Y, et al.	2018	The mRNA profiling by array for pancreatic ductal adenocarcinoma for clinical application	https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6134/	ArrayExpress, E-MTAB-6134
Yang S, He P, Wang J, et al.	2016	Microarray gene-expression profiles of 69 pancreatic tumors and 61 adjacent non-tumor tissue from patients with pancreatic ductal adenocarcinoma	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62452	Gene Expression Omnibus, GSE62452

Zhang G, He P, Tan H, et al.	2012	Microarray gene-expression profiles of 45 matching pairs of pancreatic tumor and adjacent non-tumor tissues from 45 patients with pancreatic ductal adenocarcinoma	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28735	Gene Expression Omnibus, GSE28735
Wang J, Yang S, He P, et al.	2017	Microarray gene-expression profiles of 50 pancreatic tumors tissue from patients with pancreatic ductal adenocarcinoma	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78229	Gene Expression Omnibus, GSE78229
Kirby MK, Ramaker RC, et al.	2016	RNA-sequencing of human pancreatic adenocarcinoma cancer tissues	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE79668	Gene Expression Omnibus, GSE79668
Puleo F, Maréchal R, Demetter P, et al.	2018	Patients with human resected pancreatic cancer	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85916	Gene Expression Omnibus, GSE85916

References

Siegel RL, Miller KD, Fuchs HE, Jemal A. 2022. Cancer statistics, 2022 *CA: A Cancer Journal For Clinicians* **72**. DOI: https://doi.org/10.3322/caac.21708, PMID: 35020204

Mizrahi JD, Surana R, Valle JW, Shroff RT. 2020. Pancreatic cancer *Lancet (London, England)* **395**:2008-2020. DOI: https://doi.org/10.1016/S0140-6736(20)30974-0, PMID: 32593337

Billan S, Kaidar-Person O, Gil Z. 2020. Treatment after progression in the era of immunotherapy *The Lancet Oncology* **21**:e463-e476. DOI: https://doi.org/10.1016/S1470-2045(20)30328-4, PMID: 33002442

Bear AS, Vonderheide RH, O'Hara MH. 2020. Challenges and Opportunities for Pancreatic Cancer Immunotherapy *Cancer cell* **38**:788-802. DOI: https://doi.org/10.1016/j.ccell.2020.08.004, PMID: 32946773

Golan T, Hammel P, Reni M, Van Cutsem E, Macarulla T, Hall MJ, Park J-O, Hochhauser D, Arnold D, Oh D-Y, Reinacher-Schick A, Tortora G, Algül H, O'Reilly EM, McGuinness D, Cui KY, Schlienger K, Locker GY, Kindler HL. 2019. Maintenance Olaparib for Germline -Mutated Metastatic Pancreatic Cancer *The New England journal of medicine* **381**:317-327. DOI: https://doi.org/10.1056/NEJMoa1903387, PMID: 31157963

O'Hara MH, O'Reilly EM, Varadhachary G, Wolff RA, Wainberg ZA, Ko AH, Fisher G, Rahma O, Lyman JP, Cabanski CR, Mick R, Gherardini PF, Kitch LJ, Xu J, Samuel T, Karakunnel J, Fairchild J, Bucktrout S, LaVallee TM, Selinsky C, et al. 2021. CD40 agonistic monoclonal antibody APX005M (sotigalimab) and chemotherapy, with or without nivolumab, for the treatment of metastatic pancreatic adenocarcinoma: an open-label, multicentre, phase 1b study *The Lancet Oncology* **22**:118-131. DOI: https://doi.org/10.1016/S1470-2045(20)30532-5, PMID: 33387490

Katz MHG, Hwang R, Fleming JB, Evans DB. 2008. Tumor-node-metastasis staging of pancreatic adenocarcinoma *CA: A Cancer Journal For Clinicians* **58**:111-125. DOI: https://doi.org/10.3322/CA.2007.0012, PMID: 18272835

Liu X, Wang W, Liu X, Zhang Z, Yu L, Li R, Guo D, Cai W, Quan X, Wu H, Dai M, Liang Z. 2022. Multi-omics analysis of intra-tumoural and inter-tumoural heterogeneity in pancreatic ductal adenocarcinoma *Clinical and Translational Medicine* **12**:e670. DOI: https://doi.org/10.1002/ctm2.670, PMID: 35061935

Wattenberg MM, Asch D, Yu S, O'Dwyer PJ, Domchek SM, Nathanson KL, Rosen MA, Beatty GL, Siegelman ES, Reiss KA. 2020. Platinum response characteristics of patients with pancreatic ductal adenocarcinoma and a germline BRCA1, BRCA2 or PALB2 mutation *British Journal of Cancer* **122**:333-339. DOI: https://doi.org/10.1038/s41416-019-0582-7, PMID: 31787751

Doebele RC, Drilon A, Paz-Ares L, Siena S, Shaw AT, Farago AF, Blakely CM, Seto T, Cho BC, Tosi D, Besse B, Chawla SP, Bazhenova L, Krauss JC, Chae YK, Barve M, Garrido-Laguna I, Liu SV, Conkling P, John T, et al. 2020. Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three phase 1-2 trials *The Lancet Oncology* **21**:271-282. DOI: https://doi.org/10.1016/S1470-2045(19)30691-6, PMID: 31838007

Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, Laheru D, Donehower R, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Greten TF, et al. 2017. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade *Science (New York, NY)* **357**:409-413. DOI: https://doi.org/10.1126/science.aan6733, PMID: 28596308

De Dosso S, Siebenhüner AR, Winder T, Meisel A, Fritsch R, Astaras C, Szturz P, Borner M. 2021. Treatment landscape of metastatic pancreatic cancer *Cancer Treatment Reviews* **96**:102180. DOI: https://doi.org/10.1016/j.ctrv.2021.102180, PMID: 33812339

Wang L, Liu Z, Zhu R, Liang R, Wang W, Li J, Zhang Y, Guo C, Han X, Sun Y. 2022. Multi-omics landscape and clinical significance of a SMAD4-driven immune signature: Implications for risk stratification and frontline therapies in pancreatic cancer *Computational and Structural Biotechnology Journal* **20**:1154-1167. DOI: https://doi.org/10.1016/j.csbj.2022.02.031, PMID: 35317237

Tan Z, Lei Y, Xu J, Shi S, Hua J, Zhang B, Meng Q, Liu J, Zhang Y, Wei M, Yu X, Liang C. 2020. The value of a metabolic reprogramming-related gene signature for pancreatic adenocarcinoma prognosis prediction *Aging* **12**:24228-24241. DOI: https://doi.org/10.18632/aging.104134, PMID: 33226369

Yuan Q, Ren J, Li L, Li S, Xiang K, Shang D. 2021. Development and validation of a novel N6-methyladenosine (m6A)-related multi- long non-coding RNA (lncRNA) prognostic signature in pancreatic adenocarcinoma *Bioengineered* **12**:2432-2448. DOI: https://doi.org/10.1080/21655979.2021.1933868, PMID: 34233576

Yokoyama S, Hamada T, Higashi M, Matsuo K, Maemura K, Kurahara H, Horinouchi M, Hiraki T, Sugimoto T, Akahane T, Yonezawa S, Kornmann M, Batra SK, Hollingsworth MA, Tanimoto A. 2020. Predicted Prognosis of Patients with Pancreatic Cancer by Machine Learning *Clinical cancer research : an official journal of the American Association for Cancer Research* **26**:2411-2421. DOI: https://doi.org/10.1158/1078-0432.CCR-19-1247, PMID: 31992588

Liu Z, Guo C, Dang Q, Wang L, Liu L, Weng S, Xu H, Lu T, Sun Z, Han X. 2022. Integrative analysis from multi-center studies identities a consensus machine learning-derived lncRNA signature for stage II/III colorectal cancer *EBioMedicine* **75**:103750. DOI: https://doi.org/10.1016/j.ebiom.2021.103750, PMID: 34922323

Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, Wang L, Lu T, Zhang Y, Sun Z, Han X. 2022. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer *Nature Communications* **13**:816. DOI: https://doi.org/10.1038/s41467-022-28421-6, PMID: 35145098

Lu X, Meng J, Su L, Jiang L, Wang H, Zhu J, Huang M, Cheng W, Xu L, Ruan X, Yeh S, Liang C, Yan F. 2021. Multi-omics consensus ensemble refines the classification of muscle-invasive bladder cancer with stratified prognosis, tumour microenvironment and distinct sensitivity to frontline therapies *Clinical and Translational Medicine* **11**:e601. DOI: https://doi.org/10.1002/ctm2.601, PMID: 34936229

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, et al. 2013. Signatures of mutational processes in human cancer *Nature* **500**:415-421. DOI: https://doi.org/10.1038/nature12477, PMID: 23945592

Liu Z, Wang L, Liu L, Lu T, Jiao D, Sun Y, Han X. 2021. The Identification and Validation of Two Heterogenous Subtypes and a Risk Signature Based on Ferroptosis in Hepatocellular Carcinoma *Frontiers In Oncology* **11**:619242. DOI: https://doi.org/10.3389/fonc.2021.619242, PMID: 33738257

Liu Z, Liu L, Lu T, Wang L, Li Z, Jiao D, Han X. 2021. Hypoxia Molecular Characterization in Hepatocellular Carcinoma Identifies One Risk Signature and Two Nomograms for Clinical Management *Journal of Oncology* **2021**:6664386. DOI: https://doi.org/10.1155/2021/6664386, PMID: 33552157

Liu Z, Lu T, Li J, Wang L, Xu K, Dang Q, Liu L, Guo C, Jiao D, Sun Z, Han X. 2021. Clinical Significance and Inflammatory Landscape of aNovel Recurrence-Associated Immune Signature in Stage II/III Colorectal Cancer *Frontiers in immunology* **12**:702594. DOI: https://doi.org/10.3389/fimmu.2021.702594, PMID: 34394098

Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW, Liu XS. 2018. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response *Nature Medicine* **24**:1550-1558. DOI: https://doi.org/10.1038/s41591-018-0136-1, PMID: 30127393

Hoshida Y, Brunet J-P, Tamayo P, Golub TR, Mesirov JP. 2007. Subclass mapping: identifying common subtypes in independent disease data sets *PLoS One* **2**:e1195. DOI: PMID: 18030330

Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S. 2021. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: an in silico strategy towards precision oncology *Briefings In Bioinformatics* **22**. DOI: https://doi.org/10.1093/bib/bbaa164, PMID: 32789496

Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, et al. 2017. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles *Cell* **171**. DOI: https://doi.org/10.1016/j.cell.2017.10.049, PMID: 29195078

Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, Colaprico A, Czerwińska P, Mazurek S, Mishra L, Heyn H, Krasnitz A, Godwin AK, Lazar AJ, Stuart JM, Hoadley KA, et al. 2018. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation *Cell* **173**. DOI: https://doi.org/10.1016/j.cell.2018.03.034, PMID: 29625051

Ferrone CR, Kattan MW, Tomlinson JS, Thayer SP, Brennan MF, Warshaw AL. 2005. Validation of a postresection pancreatic adenocarcinoma nomogram for disease-specific survival *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology* **23**:7529-7535. DOI: PMID: 16234519

Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dimitriadoy S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D, Daian F, Gao Q, Bailey MH, Liang W-W, Foltz SM, Shmulevich I, Ding L, Heins Z, Ochoa A, et al. 2018. Oncogenic Signaling Pathways in The Cancer Genome Atlas *Cell* **173**. DOI: https://doi.org/10.1016/j.cell.2018.03.035, PMID: 29625050

Park W, Chawla A, O'Reilly EM. 2021. Pancreatic Cancer: A Review *JAMA* **326**:851-862. DOI: https://doi.org/10.1001/jama.2021.13027, PMID: 34547082

Deo RC. 2015. Machine Learning in Medicine *Circulation* **132**:1920-1930. DOI: https://doi.org/10.1161/CIRCULATIONAHA.115.001593, PMID: 26572668

Grady WM, Yu M, Markowitz SD. 2021. Epigenetic Alterations in the Gastrointestinal Tract: Current and Emerging Use for Biomarkers of Cancer *Gastroenterology* **160**:690-709. DOI: https://doi.org/10.1053/j.gastro.2020.09.058, PMID: 33279516

Hu C, Hart SN, Polley EC, Gnanaolivu R, Shimelis H, Lee KY, Lilyquist J, Na J, Moore R, Antwi SO, Bamlet WR, Chaffee KG, DiCarlo J, Wu Z, Samara R, Kasi PM, McWilliams RR, Petersen GM, Couch FJ. 2018. Association Between Inherited Germline Mutations in Cancer Predisposition Genes and Risk of Pancreatic Cancer *JAMA* **319**:2401-2409. DOI: https://doi.org/10.1001/jama.2018.6228, PMID: 29922827

Hashimoto S, Furukawa S, Hashimoto A, Tsutaho A, Fukao A, Sakamura Y, Parajuli G, Onodera Y, Otsuka Y, Handa H, Oikawa T, Hata S, Nishikawa Y, Mizukami Y, Kodama Y, Murakami M, Fujiwara T, Hirano S, Sabe H. 2019. ARF6 and AMAP1 are major targets of and mutations to promote invasion, PD-L1 dynamics, and immune evasion of pancreatic cancer *Proceedings of the National Academy of Sciences of the United States of America* **116**:17450-17459. DOI: https://doi.org/10.1073/pnas.1901765116, PMID: 31399545

Sandhu V, Wedge DC, Bowitz Lothe IM, Labori KJ, Dentro SC, Buanes T, Skrede ML, Dalsgaard AM, Munthe E, Myklebost O, Lingjærde OC, Børresen-Dale A-L, Ikdahl T, Van Loo P, Nord S, Kure EH. 2016. The Genomic Landscape of Pancreatic and Periampullary Adenocarcinoma *Cancer Research* **76**:5092-5102. DOI: https://doi.org/10.1158/0008-5472.CAN-16-0658, PMID: 27488532

Morikawa A, Hayashi T, Kobayashi M, Kato Y, Shirahige K, Itoh T, Urashima M, Okamoto A, Akiyama T. 2018. Somatic copy number alterations have prognostic impact in patients with ovarian clear cell carcinoma *Oncology Reports* **40**:309-318. DOI: https://doi.org/10.3892/or.2018.6419, PMID: 29749539

Baker MJ, Goldstein AM, Gordon PL, Harbaugh KS, Mackley HB, Glantz MJ, Drabick JJ. 2016. An interstitial deletion within 9p21.3 and extending beyond predisposes to melanoma, neural system tumours and possible haematological malignancies *Journal of Medical Genetics* **53**:721-727. DOI: https://doi.org/10.1136/jmedgenet-2015-103446, PMID: 26794401

Jonkers YMH, Claessen SMH, Feuth T, van Kessel AG, Ramaekers FCS, Veltman JA, Speel EJM. 2006. Novel candidate tumour suppressor gene loci on chromosomes 11q23-24 and 22q13 involved in human insulinoma tumourigenesis *The Journal of Pathology* **210**:450-458. DOI: PMID: 17068744

Sha D, Jin Z, Budczies J, Kluck K, Stenzinger A, Sinicrope FA. 2020. Tumor Mutational Burden as a Predictive Biomarker in Solid Tumors *Cancer discovery* **10**:1808-1825. DOI: https://doi.org/10.1158/2159-8290.CD-20-0522, PMID: 33139244

Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W. 2018. CD4 T cell help in cancer immunology and immunotherapy *Nature Reviews Immunology* **18**:635-647. DOI: https://doi.org/10.1038/s41577-018-0044-0, PMID: 30057419

Wu F, Zhang W, Shao H, Bo H, Shen H, Li J, Liu Y, Wang T, Ma W, Huang S. 2013. Human effector T cells derived from central memory cells rather than CD8(+)T cells modified by tumor-specific TCR gene transfer possess superior traits for adoptive immunotherapy *Cancer Letters* **339**:195-207. DOI: https://doi.org/10.1016/j.canlet.2013.06.009, PMID: 23791878

Wang G, Niu X, Zhang W, Caldwell JT, Edwards H, Chen W, Taub JW, Zhao L, Ge Y. 2015. Synergistic antitumor interactions between MK-1775 and panobinostat in preclinical models of pancreatic cancer *Cancer Letters* **356**:656-668. DOI: https://doi.org/10.1016/j.canlet.2014.10.015, PMID: 25458954

Ali AI, Wang M, von Scheidt B, Dominguez PM, Harrison AJ, Tantalo DGM, Kang J, Oliver AJ, Chan JD, Du X, Bai Y, Lee B, Johnstone RW, Darcy PK, Kershaw MH, Slaney CY. 2021. A Histone Deacetylase Inhibitor, Panobinostat, Enhances Chimeric Antigen Receptor T-cell Antitumor Effect Against Pancreatic Cancer *Clinical cancer research : an official journal of the American Association for Cancer Research* **27**:6222-6234.

DOI: https://doi.org/10.1158/1078-0432.CCR-21-1141, PMID: 34475103

Figure Supplements

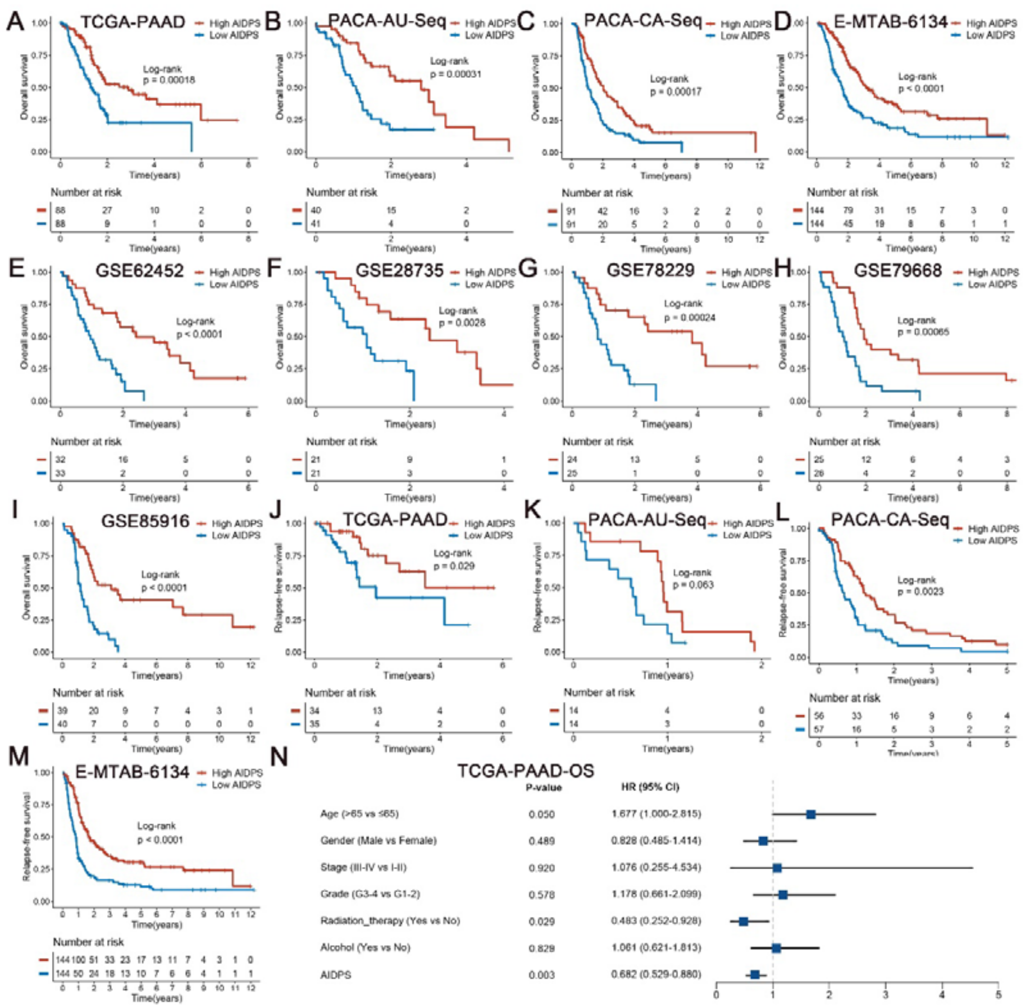


Figure 3-figure supplement 1. Survival analysis of AIDPS in remaining nine validation cohorts. (A-I) Kaplan-Meier survival analysis for OS between the high and low AIDPS groups in the TCGA-PAAD (A), PACA-AU-Seq (B), PACA-CA-Seq (C), E-MTAB-6134 (D), GSE62452 (E), GSE28735 (F), GSE78229 (G), GSE79668 (H), GSE85916 (I). (J-M) Kaplan-Meier survival analysis for RFS between the high and

low AIDPS groups in the TCGA-PAAD (J), PACA-AU-Seq (K), PACA-CA-Seq (L), E-MTAB-6134 (M). (N) Multivariate Cox regression analysis of OS in the TCGA-PAAD. OS, overall survival; RFS, relapse-free survival.

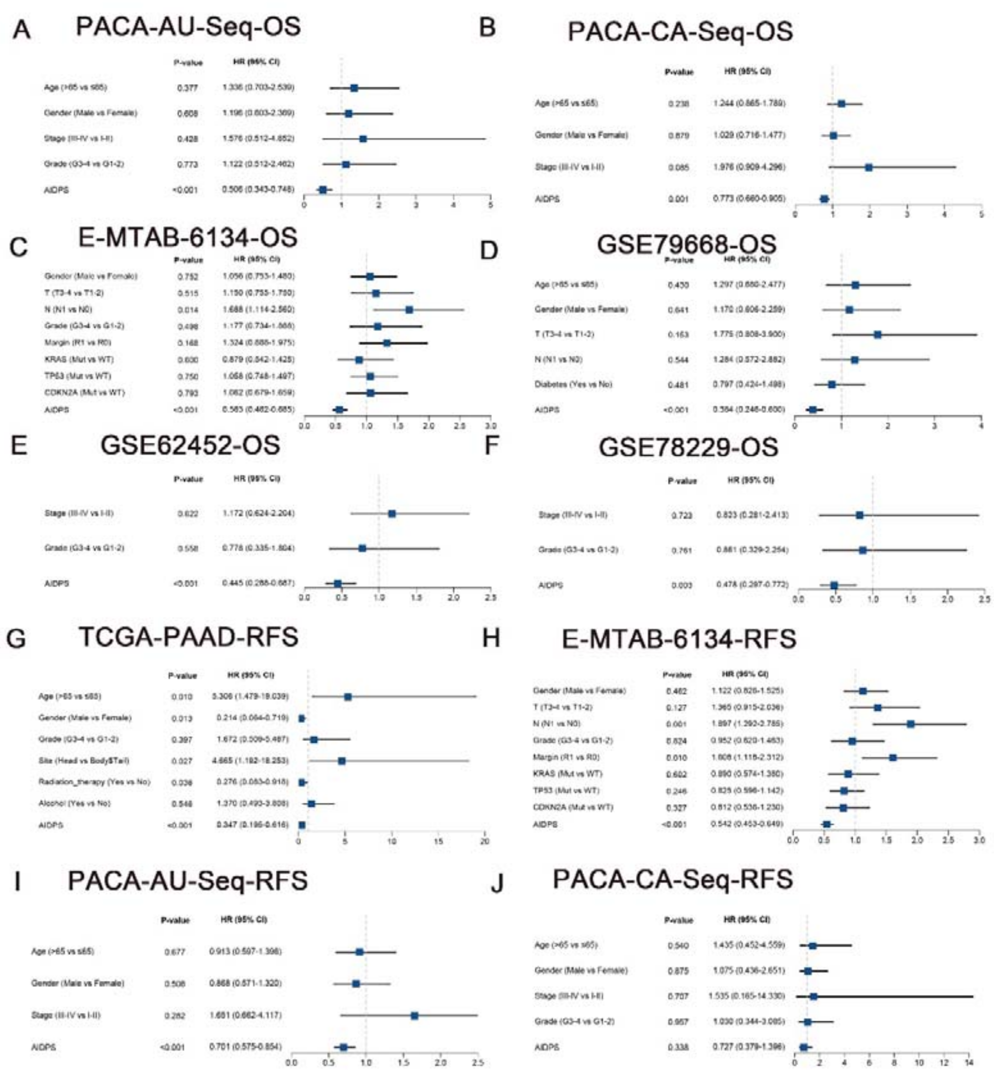


Figure 3-figure supplement 2. Survival analysis of AIDPS in remaining nine validation cohorts. (**A-F**) Multivariate Cox regression analysis of OS in the PACA-AU-Seq (**A**), PACA-CA-Seq (**B**), E-MTAB-6134 (**C**), GSE79668 (**D**), GSE62452 (**E**), GSE78229 (**F**). (**G-J**) Multivariate Cox regression analysis of RFS in the TCGA-PAAD (**G**), E-MTAB-6134 (**H**), PACA-AU-Seq (**I**), PACA-CA-Seq (**J**). OS, overall survival; RFS, relapse-free survival.

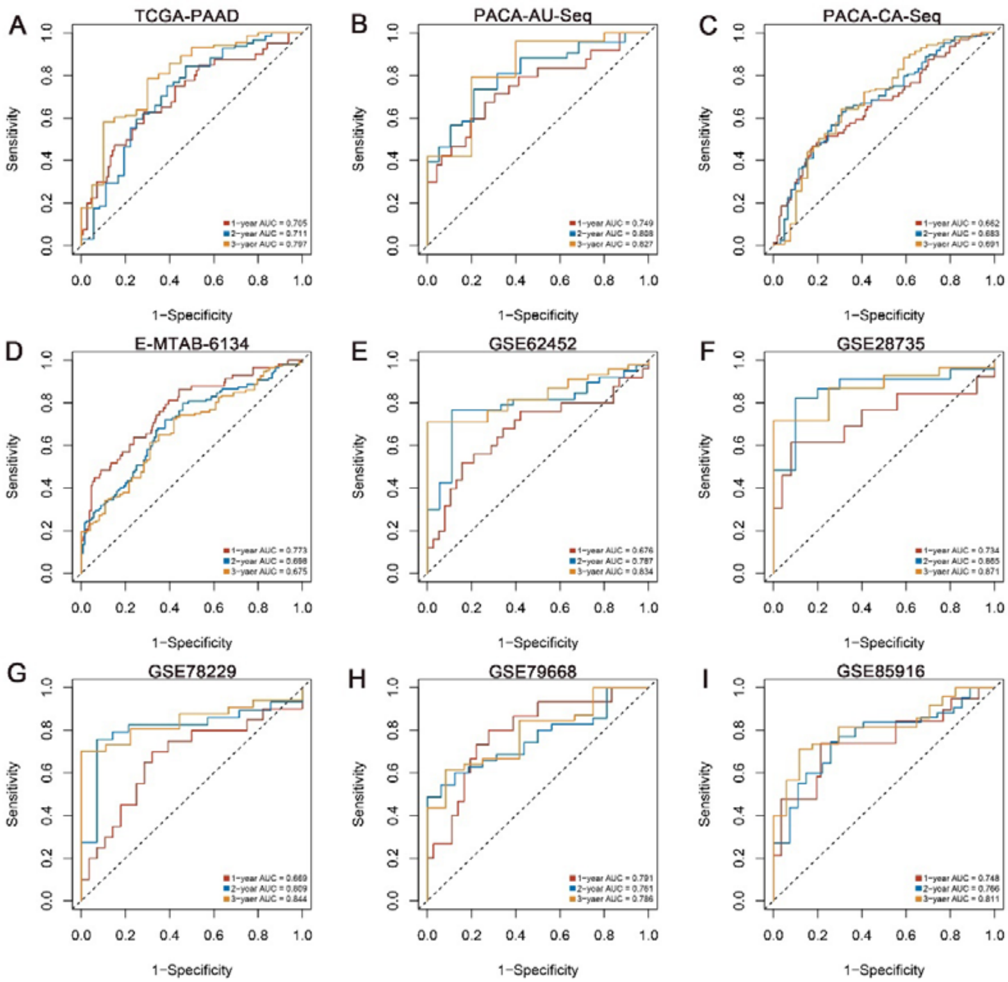


Figure 3-figure supplement 3. Predictive performance evaluation of AIDPS. Time-dependent ROC analysis for predicting 1-, 2-, and 3-years OS in the TCGA-PAAD (**A**), PACA-AU-Seq (**B**), PACA-CA-Seq (**C**), E-MTAB-6134 (**D**), GSE62452 (**E**), GSE28735 (**F**), GSE78229 (**G**), GSE79668 (**H**), GSE85916 (**I**). ROC, receiver operator characteristic; OS, overall survival.

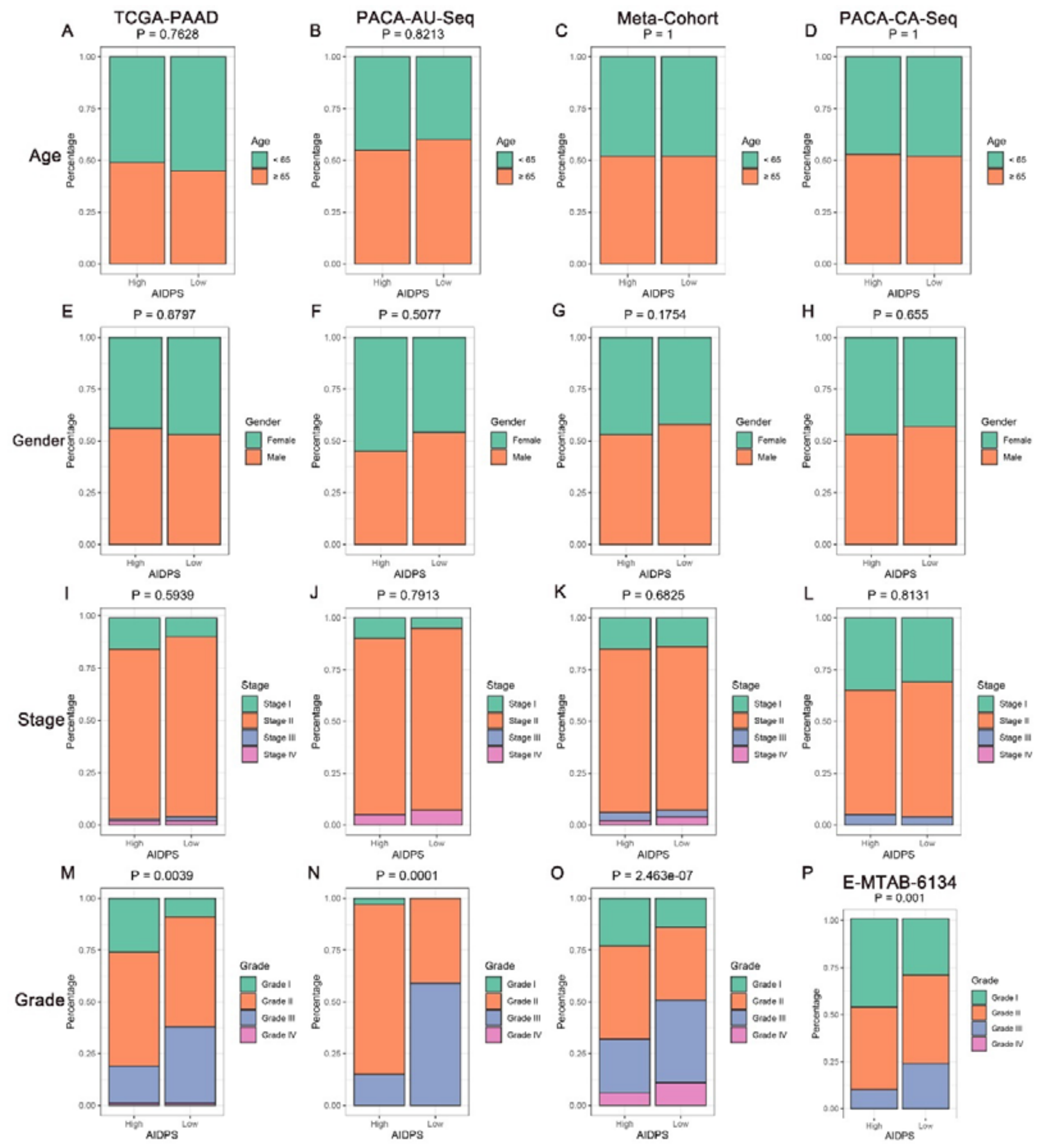


Figure 5-figure supplement 1. The clinical characteristics of the high and low AIDPS groups. (**A, E, I and M**) Composition percentage of the two groups in clinical characteristics such as age (**A**), gender (**E**), stage (**I**), grade (**M**) in the TCGA-PAAD cohort. (**B, F, J and N**) Composition percentage of the two groups in clinical characteristics such as age (**B**), gender (**F**), stage (**J**), grade (**N**) in the PACA-AU-Seq cohort. (**C, G, K and O**) Composition percentage of the two subtypes in clinical characteristics such as age (**C**), gender (**G**), stage (**K**), grade (**O**) in the Meta-Cohort

cohort. (**D, H and L**) Composition percentage of the two groups in clinical characteristics such as age (**D**), gender (**H**), stage (**L**) in the PACA-CA-Seq cohort. (**P**) Composition percentage of the two groups on grade in the E-MTAB-6134 cohort.

Source Data Files

Figure 1-source data 1. Details of baseline information in 10 public datasets

Accession	TCGA-PAAD	PACA-AU-Seq	PACA-AU-Array	PACA-CA-Seq	E-MTAB-6134	GSE62452	GSE28735	GSE78229	GSE79668	GSE85916
Number of Patients	176 (100%)	81 (100%)	267 (100%)	182 (100%)	288 (100%)	65 (100%)	42 (100%)	49 (100%)	51 (100%)	79 (100%)
Event										
Alive	84 (47.7%)	32 (39.5%)	106 (39.7%)	34 (18.7%)	107 (37.2%)	16 (24.6%)	13 (31.0%)	14 (28.6%)	6 (11.8%)	22 (27.8%)
Dead	92 (52.3%)	49 (60.5%)	161 (60.3%)	148 (81.3%)	181 (62.8%)	49 (75.4%)	29 (69.0%)	35 (71.4%)	45 (88.2%)	57 (72.2%)
Age										
<=65	93 (52.8%)	34 (42.0%)	116 (43.4%)	78 (42.9%)	--	--	--	--	25 (49.0%)	--
>65	83 (47.2%)	46 (56.8%)	150 (56.2%)	86 (47.3%)	--	--	--	--	26 (51.0%)	--
Not available	--	1 (1.2%)	1 (0.4%)	18 (9.9%)	--	--	--	--	--	--
Gender										
Female	80 (45.5%)	41 (50.6%)	125 (46.8%)	82 (45.1%)	122 (42.4%)	--	--	--	19 (37.3%)	--
Male	96 (54.5%)	40 (49.4%)	142 (53.2%)	99 (54.4%)	166 (57.6%)	--	--	--	32 (62.7%)	--
Not available	--	--	--	1 (0.5%)	--	--	--	--	--	--
Stage										
Ⅰ+Ⅱ	166(94.3%)	75 (92.6%)	251 (94.0%)	146 (80.2%)	--	49 (75.4%)	--	4 (8.2%)	--	--
Ⅲ+Ⅳ	7 (4.0%)	5 (6.2%)	122 (4.5%)	7 (3.8%)	--	16 (24.6%)	--	45 (91.8%)	--	--
Not available	3 (1.7%)	1 (1.2%)	4 (1.5%)	29 (15.9%)	--	--	--	--	--	--
T stage										
T1+T2	31 (17.6%)	9 (11.1%)	45 (16.8%)	--	51 (17.7%)	--	--	--	15 (29.4%)	--
T3+T4	143(81.2%)	70 (86.4%)	218 (81.6)	--	237 (82.3%)	--	--	--	36 (70.6%)	--
Not available	2 (1.1%)	2 (2.5%)	4 (1.5%)	--	--	--	--	--	--	--
N stage										
N0	49 (27.8%)	26 (32.1%)	63 (23.6%)	--	72(25.0%)	--	--	--	14 (27.5%)	--
N1+N2	122(69.3%)	53 (65.4%)	199 (74.5%)	--	216 (75.0%)	--	--	--	37 (72.5%)	--
Not available	5 (2.8%)	2 (2.5%)	5 (1.9%)	--	--	--	--	--	--	--
Grade										
G1-2	124(70.5%)	49 (60.5%)	22 (8.2%)	--	240 (83.3%)	34 (52.3%)	--	26 (53.1%)	--	--
G3-4	50 (28.4%)	29 (35.8%)	241 (90.3%)	--	48 (16.7%)	30 (46.2%)	--	22 (44.9%)	--	--
Not available	2 (1.1%)	3 (3.7%)	4 (1.5%)	--	--	1 (1.5%)	--	1 (2.0%)	--	--
Relapse status										
No	46 (26.1%)	40 (49.4%)	120 (44.9%)	85 (46.7%)	73 (25.3%)	--	--	--	--	--
Yes	23 (13.1%)	41 (50.6%)	147 (55.1%)	97 (53.3%)	215 (74.7%)	--	--	--	--	--
Not available	107(60.8%)	--	--	--		--	--	--	--	--
Radiation_therapy										

No	100(56.8%)	--	--	--	--	--	--	--	--	--
Yes	37 (21.0%)	--	--	--	--	--	--	--	--	--
Not available	39 (22.2%)	--	--	--	--	--	--	--	--	--
Alcohol										
No	64 (36.4%)	--	--	--	--	--	--	--	--	--
Yes	100(56.8%)	--	--	--	--	--	--	--	--	--
Not available	12 (6.8%)	--	--	--	--	--	--	--	--	--
Margin										
R0	--	--	--	--	235 (81.6%)	--	--	--	--	--
R1	--	--	--	--	49 (17.0%)	--	--	--	--	--
Not available	--	--	--	--	4 (1.4%)	--	--	--	--	--
Site										
Body&Tail	29 (16.5%)	--	--	--	--	--	--	--	--	--
Head	128(72.7%)	--	--	--	--	--	--	--	--	--
Not available	19 (10.8%)	--	--	--	--	--	--	--	--	--
Diabetes										
No	--	--	--	--	--	--	--	--	29 (56.9%)	--
Yes	--	--	--	--	--	--	--	--	22 (43.1%)	--
KRAS Status										
Wild	--	--	--	--	31 (10.8%)	--	--	--	--	--
Mutation	--	--	--	--	230 (79.9%)	--	--	--	--	--
Not available	--	--	--	--	27 (9.4%)	--	--	--	--	--
TP53 Status										
Wild	--	--	--	--	80 (27.8%)	--	--	--	--	--
Mutation	--	--	--	--	181 (62.8%)	--	--	--	--	--
Not available	--	--	--	--	27 (9.4%)	--	--	--	--	--
CDKN2A Status										
Wild	--	--	--	--	220 (76.4%)	--	--	--	--	--
Mutation	--	--	--	--	41 (14.2%)	--	--	--	--	--
Not available	--	--	--	--	27 (9.4%)	--	--	--	--	--

Figure 2-source data 1. The 9 genes included in the AIDPS

ENSEMBL	SYMBOL
ENSG00000143416	SELENBP1
ENSG00000057019	DCBLD2
ENSG00000124882	EREG
ENSG00000148926	ADM
ENSG00000170779	CDCA4
ENSG00000068489	PRR11
ENSG00000092929	UNC13D
ENSG00000101333	PLCB4
ENSG00000198959	TGM2

Figure 4-source data 1. Details of 86 published mRNA/LncRNA signatures in PACA.

Model	PMID	Type	Author	Year	Coef	Gene	ENSEMBL
Model-1	34930261	mRNA	Cai W	2021	0.912	CA12	ENSG00000074410
Model-1	34930261	mRNA	Cai W	2021	0.9	CDA	ENSG00000158825
Model-1	34930261	mRNA	Cai W	2021	-3.3881	DGKZ	ENSG00000149091
Model-1	34930261	mRNA	Cai W	2021	6.769	GMPS	ENSG00000163655
Model-1	34930261	mRNA	Cai W	2021	-6.0982	PI4KB	ENSG00000143393
Model-2	33465729	mRNA	Chen B	2021	-0.1661	AGT	ENSG00000135744
Model-2	33465729	mRNA	Chen B	2021	0.1638	CXCL9	ENSG00000138755
Model-2	33465729	mRNA	Chen B	2021	0.2654	EGF	ENSG00000138798
Model-2	33465729	mRNA	Chen B	2021	0.5676	MET	ENSG00000105976
Model-2	33465729	mRNA	Chen B	2021	0.1655	RARRES3	ENSG00000133321
Model-2	33465729	mRNA	Chen B	2021	0.1964	S100A14	ENSG00000189334
Model-2	33465729	mRNA	Chen B	2021	0.0953	SPP1	ENSG00000118785
Model-3	34691034	mRNA	Chen D	2021	-0.45	GDF11	ENSG00000135414
Model-3	34691034	mRNA	Chen D	2021	0.025	IL18	ENSG00000150782
Model-3	34691034	mRNA	Chen D	2021	0.18	PLAU	ENSG00000122861
Model-3	34691034	mRNA	Chen D	2021	0.03	S100A16	ENSG00000188643
Model-3	34691034	mRNA	Chen D	2021	-0.02	NR0B1	ENSG00000169297
Model-3	34691034	mRNA	Chen D	2021	0.17	SEMA3C	ENSG00000075223
Model-3	34691034	mRNA	Chen D	2021	0.01	PPP3CA	ENSG00000138814
Model-4	26247463	mRNA	Chen DT	2015	0.157	C6orf15	ENSG00000204542
Model-4	26247463	mRNA	Chen DT	2015	0.178	CAPN8	ENSG00000203697
Model-4	26247463	mRNA	Chen DT	2015	0.22	HIST1H3H	ENSG00000278828
Model-4	26247463	mRNA	Chen DT	2015	0.234	IGF2BP3	ENSG00000136231
Model-4	26247463	mRNA	Chen DT	2015	0.243	SCEL	ENSG00000136155
Model-4	26247463	mRNA	Chen DT	2015	0.261	KIF14	ENSG00000118193
Model-4	26247463	mRNA	Chen DT	2015	0.247	KRT6A	ENSG00000205420
Model-4	26247463	mRNA	Chen DT	2015	0.281	SLC2A1	ENSG00000117394
Model-4	26247463	mRNA	Chen DT	2015	0.247	PMAIP1	ENSG00000141682
Model-4	26247463	mRNA	Chen DT	2015	0.157	PPBP	ENSG00000163736
Model-4	26247463	mRNA	Chen DT	2015	0.17	UCA1	ENSG00000214049
Model-4	26247463	mRNA	Chen DT	2015	0.256	RTKN2	ENSG00000182010
Model-4	26247463	mRNA	Chen DT	2015	0.271	SLC45A3	ENSG00000158715
Model-4	26247463	mRNA	Chen DT	2015	0.271	SERPINB5	ENSG00000206075
Model-4	26247463	mRNA	Chen DT	2015	0.208	TMPRSS3	ENSG00000160183
Model-5	30710069	mRNA	Chen H	2019	0.65	MAP4K4	ENSG00000071054
Model-5	30710069	mRNA	Chen H	2019	0.129	IGF2BP3	ENSG00000136231
Model-5	30710069	mRNA	Chen H	2019	0.195	SULT1E1	ENSG00000109193
Model-6	31102348	mRNA	Chen K	2019	0.4099	HSBP1L1	ENSG00000226742
Model-6	31102348	mRNA	Chen K	2019	-0.3448	D--H10	ENSG00000197653
Model-6	31102348	mRNA	Chen K	2019	-0.3725	KIAA0513	ENSG00000135709
Model-6	31102348	mRNA	Chen K	2019	1.3175	MRPL3	ENSG00000114686

Model-7	34632851	mRNA	Chen Q	2021	0.2805	B3GNT3	ENSG00000179913
Model-7	34632851	mRNA	Chen Q	2021	0.3705	BCAT1	ENSG00000060982
Model-7	34632851	mRNA	Chen Q	2021	0.4273	KYNU	ENSG00000115919
Model-7	34632851	mRNA	Chen Q	2021	0.3774	LDHA	ENSG00000134333
Model-7	34632851	mRNA	Chen Q	2021	0.549	TYMS	ENSG00000176890
Model-8	34276760	mRNA	Chen S	2021	-0.3281	COL2A1	ENSG00000139219
Model-8	34276760	mRNA	Chen S	2021	-0.3961	CUX2	ENSG00000111249
Model-8	34276760	mRNA	Chen S	2021	0.2074	CXCL10	ENSG00000169245
Model-8	34276760	mRNA	Chen S	2021	-0.1998	TRPC7	ENSG00000069018
Model-9	34094925	mRNA	Chen S_2	2021	-0.329	COL11A1	ENSG00000060718
Model-9	34094925	mRNA	Chen S_2	2021	-1.358	COL12A1	ENSG00000111799
Model-9	34094925	mRNA	Chen S_2	2021	3.811	COL1A1	ENSG00000108821
Model-9	34094925	mRNA	Chen S_2	2021	-12.93	COL3A1	ENSG00000168542
Model-9	34094925	mRNA	Chen S_2	2021	3.86	COL5A2	ENSG00000204262
Model-9	34094925	mRNA	Chen S_2	2021	10.44	COL6A3	ENSG00000163359
Model-9	34094925	mRNA	Chen S_2	2021	-0.2232	ITGA2	ENSG00000164171
Model-9	34094925	mRNA	Chen S_2	2021	4.88	MMP14	ENSG00000157227
Model-9	34094925	mRNA	Chen S_2	2021	-4.655	THBS2	ENSG00000186340
Model-10	34090418	mRNA	Chen Y	2021	-0.5399	CBX8	ENSG00000141570
Model-10	34090418	mRNA	Chen Y	2021	-0.5193	CENPT	ENSG00000102901
Model-10	34090418	mRNA	Chen Y	2021	0.7465	DPY30	ENSG00000162961
Model-10	34090418	mRNA	Chen Y	2021	0.1046	PADI1	ENSG00000142623
Model-11	32310997	mRNA	Demirkol CS	2020	1	ARNTL2	ENSG00000029153
Model-11	32310997	mRNA	Demirkol CS	2020	-1	C2orf42	ENSG00000115998
Model-11	32310997	mRNA	Demirkol CS	2020	-1	CADPS2	ENSG00000081803
Model-11	32310997	mRNA	Demirkol CS	2020	-1	CBX7	ENSG00000100307
Model-11	32310997	mRNA	Demirkol CS	2020	1	EPS8	ENSG00000151491
Model-11	32310997	mRNA	Demirkol CS	2020	1	ERRFI1	ENSG00000116285
Model-11	32310997	mRNA	Demirkol CS	2020	-1	KANK1	ENSG00000107104
Model-11	32310997	mRNA	Demirkol CS	2020	1	GSK3B	ENSG00000082701
Model-11	32310997	mRNA	Demirkol CS	2020	-1	PITP--	ENSG00000174238
Model-11	32310997	mRNA	Demirkol CS	2020	1	LDHA	ENSG00000134333
Model-11	32310997	mRNA	Demirkol CS	2020	1	MAP4K4	ENSG00000071054
Model-11	32310997	mRNA	Demirkol CS	2020	-1	MIA3	ENSG00000154305
Model-11	32310997	mRNA	Demirkol CS	2020	1	TRIO	ENSG00000038382
Model-11	32310997	mRNA	Demirkol CS	2020	-1	NDUFB2	ENSG00000090266
Model-11	32310997	mRNA	Demirkol CS	2020	1	SLC20A1	ENSG00000144136
Model-11	32310997	mRNA	Demirkol CS	2020	-1	POLR3H	ENSG00000100413
Model-11	32310997	mRNA	Demirkol CS	2020	1	RAB7A	ENSG00000075785
Model-11	32310997	mRNA	Demirkol CS	2020	1	STX16	ENSG00000124222
Model-11	32310997	mRNA	Demirkol CS	2020	1	TFG	ENSG00000114354
Model-11	32310997	mRNA	Demirkol CS	2020	-1	ZNF557	ENSG00000130544
Model-12	33865399	mRNA	Ding J	2021	-0.799	ENO3	ENSG00000108515
Model-12	33865399	mRNA	Ding J	2021	1.026	LDHA	ENSG00000134333

Model-12	33865399	mRNA	Ding J	2021	-0.484	PGK1	ENSG00000102144
Model-12	33865399	mRNA	Ding J	2021	0.415	PGM1	ENSG00000079739
Model-13	34290570	mRNA	Dong Y	2021	0.32	ANO1	ENSG00000131620
Model-13	34290570	mRNA	Dong Y	2021	0.135	FAM83A	ENSG00000147689
Model-13	34290570	mRNA	Dong Y	2021	0.003	GPR87	ENSG00000138271
Model-13	34290570	mRNA	Dong Y	2021	2.00E-04	ITGB6	ENSG00000115221
Model-13	34290570	mRNA	Dong Y	2021	0.15	KLK10	ENSG00000129451
Model-13	34290570	mRNA	Dong Y	2021	0.03	SERPINE1	ENSG00000106366
Model-13	34290570	mRNA	Dong Y	2021	-0.18	SMIM32	ENSG00000271824
Model-14	34095229	mRNA	Feng Z	2021	0.2342	ASPH	ENSG00000198363
Model-14	34095229	mRNA	Feng Z	2021	-0.6197	BLOC1S3	ENSG00000189114
Model-14	34095229	mRNA	Feng Z	2021	0.162	FAM83A	ENSG00000147689
Model-14	34095229	mRNA	Feng Z	2021	0.2842	DDX10	ENSG00000178105
Model-14	34095229	mRNA	Feng Z	2021	-0.2568	PPM1H	ENSG00000111110
Model-14	34095229	mRNA	Feng Z	2021	-0.2905	NR0B2	ENSG00000131910
Model-14	34095229	mRNA	Feng Z	2021	-0.4056	SLAMF6	ENSG00000162739
Model-15	32958051	mRNA	Feng Z_2	2020	0.3111	DCBLD2	ENSG00000057019
Model-15	32958051	mRNA	Feng Z_2	2020	0.2939	GSDMD	ENSG00000104518
Model-15	32958051	mRNA	Feng Z_2	2020	0.5141	PLOD2	ENSG00000152952
Model-15	32958051	mRNA	Feng Z_2	2020	0.1919	PMAIP1	ENSG00000141682
Model-16	33748108	mRNA	Feng Z_3	2021	0.42	EREG	ENSG00000124882
Model-16	33748108	mRNA	Feng Z_3	2021	-0.61	MCM3AP	ENSG00000160294
Model-16	33748108	mRNA	Feng Z_3	2021	0.33	MCM7	ENSG00000166508
Model-16	33748108	mRNA	Feng Z_3	2021	-0.441	KCTD13	ENSG00000174943
Model-16	33748108	mRNA	Feng Z_3	2021	0.328	POLG2	ENSG00000256525
Model-16	33748108	mRNA	Feng Z_3	2021	0.263	TP73	ENSG00000078900
Model-16	33748108	mRNA	Feng Z_3	2021	-0.542	TERF2	ENSG00000132604
Model-17	33996821	mRNA	Feng Z_4	2021	0.3041	DLX2	ENSG00000115844
Model-17	33996821	mRNA	Feng Z_4	2021	0.2146	ITGB6	ENSG00000115221
Model-17	33996821	mRNA	Feng Z_4	2021	-0.2424	FGF9	ENSG00000102678
Model-17	33996821	mRNA	Feng Z_4	2021	-0.1568	LGR5	ENSG00000139292
Model-17	33996821	mRNA	Feng Z_4	2021	0.6384	MYC	ENSG00000136997
Model-17	33996821	mRNA	Feng Z_4	2021	-0.4059	IL6R	ENSG00000160712
Model-17	33996821	mRNA	Feng Z_4	2021	-0.4479	TNFSF12	ENSG00000239697
Model-17	33996821	mRNA	Feng Z_4	2021	-0.1231	S100A2	ENSG00000196754
Model-18	34631790	mRNA	Feng Z_5	2021	0.2247	CAV1	ENSG00000105974
Model-18	34631790	mRNA	Feng Z_5	2021	0.265	DDIT4	ENSG00000168209
Model-18	34631790	mRNA	Feng Z_5	2021	-0.24	SLC40A1	ENSG00000138449
Model-18	34631790	mRNA	Feng Z_5	2021	0.9346	SRXN1	ENSG00000271303
Model-18	34631790	mRNA	Feng Z_5	2021	0.1441	TFAP2C	ENSG00000087510
Model-19	33550277	mRNA	Gu M	2021	-0.884	CCNT1	ENSG00000129315
Model-19	33550277	mRNA	Gu M	2021	-2.446	HMOX2	ENSG00000103415
Model-19	33550277	mRNA	Gu M	2021	-1.934	ITGB3	ENSG00000259207
Model-19	33550277	mRNA	Gu M	2021	-3.871	SDS	ENSG00000135094

Model-20	33845841	mRNA	Gu X	2021	0.3784	FCGR2B	ENSG00000072694
Model-20	33845841	mRNA	Gu X	2021	0.4762	HLA-DRA	ENSG00000204287
Model-20	33845841	mRNA	Gu X	2021	-1.0217	IL10RA	ENSG00000110324
Model-21	25587357	mRNA	Haider S	2014	0.1231	ADAMTS14	ENSG00000138316
Model-21	25587357	mRNA	Haider S	2014	0.2345	ADM	ENSG00000148926
Model-21	25587357	mRNA	Haider S	2014	-0.2444	ARRB1	ENSG00000137486
Model-21	25587357	mRNA	Haider S	2014	-0.4377	B3GNT1	ENSG00000170340
Model-21	25587357	mRNA	Haider S	2014	0.0751	BLM	ENSG00000197299
Model-21	25587357	mRNA	Haider S	2014	-0.3389	CADPS	ENSG00000163618
Model-21	25587357	mRNA	Haider S	2014	0.3667	CDC45	ENSG00000093009
Model-21	25587357	mRNA	Haider S	2014	0.1296	CDK2AP1	ENSG00000111328
Model-21	25587357	mRNA	Haider S	2014	-0.0317	CIT	ENSG00000122966
Model-21	25587357	mRNA	Haider S	2014	0.1527	CKAP2L	ENSG00000169607
Model-21	25587357	mRNA	Haider S	2014	-0.1648	CNNM3	ENSG00000168763
Model-21	25587357	mRNA	Haider S	2014	-0.2899	EIF4E3	ENSG00000163412
Model-21	25587357	mRNA	Haider S	2014	0.184	KIF14	ENSG00000118193
Model-21	25587357	mRNA	Haider S	2014	0.1349	GRPEL2	ENSG00000164284
Model-21	25587357	mRNA	Haider S	2014	-0.3216	GTF2IRD2	ENSG00000196275
Model-21	25587357	mRNA	Haider S	2014	-0.1804	GTF2IRD2B	ENSG00000174428
Model-21	25587357	mRNA	Haider S	2014	-0.1242	ICOSLG	ENSG00000277117
Model-21	25587357	mRNA	Haider S	2014	0.2668	IGF2BP2	ENSG00000073792
Model-21	25587357	mRNA	Haider S	2014	0.4462	IL20RB	ENSG00000174564
Model-21	25587357	mRNA	Haider S	2014	0.1777	NPLOC4	ENSG00000182446
Model-21	25587357	mRNA	Haider S	2014	0.5888	ITGA5	ENSG00000161638
Model-21	25587357	mRNA	Haider S	2014	-0.0662	ITGBL1	ENSG00000198542
Model-21	25587357	mRNA	Haider S	2014	0.5459	KIF4A	ENSG00000090889
Model-21	25587357	mRNA	Haider S	2014	-0.366	NOSTRIN	ENSG00000163072
Model-21	25587357	mRNA	Haider S	2014	0.5742	SEMA3A	ENSG00000075213
Model-21	25587357	mRNA	Haider S	2014	-0.1145	SKA3	ENSG00000165480
Model-21	25587357	mRNA	Haider S	2014	0.2209	PHLDA1	ENSG00000139289
Model-21	25587357	mRNA	Haider S	2014	0.3898	SLC20A1	ENSG00000144136
Model-21	25587357	mRNA	Haider S	2014	0.3201	PXN	ENSG00000089159
Model-21	25587357	mRNA	Haider S	2014	-0.1181	QDPR	ENSG00000151552
Model-21	25587357	mRNA	Haider S	2014	0.3173	TMEM26	ENSG00000196932
Model-21	25587357	mRNA	Haider S	2014	0.1134	RFX8	ENSG00000196460
Model-21	25587357	mRNA	Haider S	2014	0.1698	RPL39L	ENSG00000163923
Model-21	25587357	mRNA	Haider S	2014	-0.2347	RPSAP58	ENSG00000225178
Model-21	25587357	mRNA	Haider S	2014	-0.1939	ZNF471	ENSG00000196263
Model-21	25587357	mRNA	Haider S	2014	0.3158	SSX3	ENSG00000165584
Model-22	32702921	mRNA	He QL	2020	0.02	BICC1	ENSG00000122870
Model-22	32702921	mRNA	He QL	2020	0.052	CYSLTR1	ENSG00000173198
Model-22	32702921	mRNA	He QL	2020	0.031	GBP5	ENSG00000154451
Model-22	32702921	mRNA	He QL	2020	0.078	P2RY6	ENSG00000171631
Model-22	32702921	mRNA	He QL	2020	-0.322	RAB39B	ENSG00000155961

Model-22	32702921	mRNA	He QL	2020	-0.449	VENTX	ENSG00000151650
Model-22	32702921	mRNA	He QL	2020	0.121	SLC7A14	ENSG00000013293
Model-23	32754191	mRNA	Hou J	2020	0.34	HNRNPC	ENSG00000092199
Model-23	32754191	mRNA	Hou J	2020	-0.11	METTL3	ENSG00000165819
Model-23	32754191	mRNA	Hou J	2020	0.28	IGF2BP2	ENSG00000073792
Model-23	32754191	mRNA	Hou J	2020	0.04	IGF2BP3	ENSG00000136231
Model-23	32754191	mRNA	Hou J	2020	0.28	KIAA1429	ENSG00000164944
Model-23	32754191	mRNA	Hou J	2020	-0.37	YTHDF1	ENSG00000149658
Model-24	34671679	mRNA	Huang XY	2021	-1.1241	BIRC5	ENSG00000089685
Model-24	34671679	mRNA	Huang XY	2021	-2.5941	BUB1	ENSG00000169679
Model-24	34671679	mRNA	Huang XY	2021	3.3769	CCNB2	ENSG00000157456
Model-24	34671679	mRNA	Huang XY	2021	3.7166	CDK1	ENSG00000170312
Model-24	34671679	mRNA	Huang XY	2021	5.0568	TPX2	ENSG00000088325
Model-24	34671679	mRNA	Huang XY	2021	-3.0209	UBE2C	ENSG00000175063
Model-24	34671679	mRNA	Huang XY	2021	-4.0606	ZWINT	ENSG00000122952
Model-25	34122496	mRNA	Huo J	2021	-0.0995	ABCA5	ENSG00000154265
Model-25	34122496	mRNA	Huo J	2021	0.0562	CAC--2D4	ENSG00000151062
Model-25	34122496	mRNA	Huo J	2021	0.004	GALNT10	ENSG00000164574
Model-25	34122496	mRNA	Huo J	2021	0.0064	DPYD	ENSG00000188641
Model-25	34122496	mRNA	Huo J	2021	0.0059	GART	ENSG00000159131
Model-25	34122496	mRNA	Huo J	2021	0.0016	GPD2	ENSG00000115159
Model-25	34122496	mRNA	Huo J	2021	0.0194	MTAP	ENSG00000099810
Model-25	34122496	mRNA	Huo J	2021	0.0083	INPP4B	ENSG00000109452
Model-25	34122496	mRNA	Huo J	2021	-0.0174	IP6K1	ENSG00000176095
Model-25	34122496	mRNA	Huo J	2021	0.0062	OAS2	ENSG00000111335
Model-25	34122496	mRNA	Huo J	2021	0.002	MTHFD1	ENSG00000100714
Model-25	34122496	mRNA	Huo J	2021	-0.0197	SLC25A27	ENSG00000153291
Model-25	34122496	mRNA	Huo J	2021	0.0065	STS	ENSG00000101846
Model-25	34122496	mRNA	Huo J	2021	0.0021	SULF2	ENSG00000196562
Model-25	34122496	mRNA	Huo J	2021	-0.0126	SLC2A8	ENSG00000136856
Model-26	33819918	mRNA	Jiang P	2021	-0.2256	ATG4D	ENSG00000130734
Model-26	33819918	mRNA	Jiang P	2021	0.2279	AURKA	ENSG00000087586
Model-26	33819918	mRNA	Jiang P	2021	-1.3943	BAP1	ENSG00000163930
Model-26	33819918	mRNA	Jiang P	2021	0.2577	CAPG	ENSG00000042493
Model-26	33819918	mRNA	Jiang P	2021	0.27	CAV1	ENSG00000105974
Model-26	33819918	mRNA	Jiang P	2021	0.1823	DDIT4	ENSG00000168209
Model-26	33819918	mRNA	Jiang P	2021	0.0971	PTGS2	ENSG00000073756
Model-26	33819918	mRNA	Jiang P	2021	0.4656	MAP3K5	ENSG00000197442
Model-26	33819918	mRNA	Jiang P	2021	0.1731	MT1G	ENSG00000125144
Model-26	33819918	mRNA	Jiang P	2021	0.0247	RRM2	ENSG00000171848
Model-26	33819918	mRNA	Jiang P	2021	-0.3538	SLC1A4	ENSG00000115902
Model-26	33819918	mRNA	Jiang P	2021	-0.3383	ZNF419	ENSG00000105136
Model-26	33819918	mRNA	Jiang P	2021	0.2476	STEAP3	ENSG00000115107
Model-26	33819918	mRNA	Jiang P	2021	-0.2033	TUBE1	ENSG00000074935

Model-27	34313248	mRNA	Jiang PC	2021	-0.5699	ATG4B	ENSG00000168397
Model-27	34313248	mRNA	Jiang PC	2021	0.191	BIRC5	ENSG00000089685
Model-27	34313248	mRNA	Jiang PC	2021	1.0582	BNIP1	ENSG00000113734
Model-27	34313248	mRNA	Jiang PC	2021	1.2169	CASP4	ENSG00000196954
Model-27	34313248	mRNA	Jiang PC	2021	-0.5859	EEF2	ENSG00000167658
Model-27	34313248	mRNA	Jiang PC	2021	0.9612	EEF2K	ENSG00000103319
Model-27	34313248	mRNA	Jiang PC	2021	0.4921	ITGA3	ENSG00000005884
Model-27	34313248	mRNA	Jiang PC	2021	-0.6779	RAB24	ENSG00000169228
Model-27	34313248	mRNA	Jiang PC	2021	0.6804	ULK1	ENSG00000177169
Model-27	34313248	mRNA	Jiang PC	2021	-0.7288	TSC1	ENSG00000165699
Model-28	34659888	mRNA	Katsuta E	2021	-0.242	HOXA4	ENSG00000197576
Model-28	34659888	mRNA	Katsuta E	2021	-0.101	DMRT3	ENSG00000064218
Model-28	34659888	mRNA	Katsuta E	2021	-0.094	ISL2	ENSG00000159556
Model-28	34659888	mRNA	Katsuta E	2021	-0.098	PHKG1	ENSG00000164776
Model-28	34659888	mRNA	Katsuta E	2021	-0.19	TRA2A	ENSG00000164548
Model-29	31703415	mRNA	Kim J	2019	0.5086	E2F7	ENSG00000165891
Model-29	31703415	mRNA	Kim J	2019	0.4543	IFI44	ENSG00000137965
Model-29	31703415	mRNA	Kim J	2019	0.4618	LAMA3	ENSG00000053747
Model-29	31703415	mRNA	Kim J	2019	-0.4652	LRIG1	ENSG00000144749
Model-29	31703415	mRNA	Kim J	2019	-0.4201	SLC12A2	ENSG00000064651
Model-30	33775699	mRNA	Li A	2021	-0.14	CHST2	ENSG00000175040
Model-30	33775699	mRNA	Li A	2021	0.1755	KIF20A	ENSG00000112984
Model-30	33775699	mRNA	Li A	2021	0.0214	MET	ENSG00000105976
Model-31	32522048	lncRNA	Li M	2020	-0.0066	AC009014.3	ENSG00000271824
Model-31	32522048	lncRNA	Li M	2020	0.0093	RP11-48O20.4	ENSG00000224259
Model-31	32522048	lncRNA	Li M	2020	0.0732	UCA1	ENSG00000214049
Model-32	33747931	mRNA	Li MX	2021	-0.082	ATP8B2	ENSG00000143515
Model-32	33747931	mRNA	Li MX	2021	-0.098	BIN1	ENSG00000136717
Model-32	33747931	mRNA	Li MX	2021	-0.112	ELOM1	ENSG00000155849
Model-32	33747931	mRNA	Li MX	2021	0.051	ERAP2	ENSG00000164308
Model-32	33747931	mRNA	Li MX	2021	-0.081	FAM118A	ENSG00000100376
Model-32	33747931	mRNA	Li MX	2021	0.017	RAPGEFL1	ENSG00000108352
Model-32	33747931	mRNA	Li MX	2021	0.394	KIF23	ENSG00000137807
Model-32	33747931	mRNA	Li MX	2021	0.005	LAPTM4A	ENSG00000068697
Model-32	33747931	mRNA	Li MX	2021	0.014	RGS16	ENSG00000143333
Model-33	34394706	mRNA	Li Z	2021	0.2339	ANLN	ENSG00000011426
Model-33	34394706	mRNA	Li Z	2021	0.1041	LY6D	ENSG00000167656
Model-33	34394706	mRNA	Li Z	2021	0.126	MYEOV	ENSG00000172927
Model-33	34394706	mRNA	Li Z	2021	-0.2197	SCN11A	ENSG00000168356
Model-33	34394706	mRNA	Li Z	2021	0.1388	ZNF488	ENSG00000265763
Model-34	28979141	mRNA	Liao X	2017	0.372	ARHGAP15	ENSG00000075884
Model-34	28979141	mRNA	Liao X	2017	-0.446	ARHGAP30	ENSG00000186517
Model-34	28979141	mRNA	Liao X	2017	0.267	CD247	ENSG00000198821
Model-34	28979141	mRNA	Liao X	2017	-0.376	CD96	ENSG00000153283

Model-34	28979141	mRNA	Liao X	2017	-0.284	FAM78A	ENSG00000126882
Model-34	28979141	mRNA	Liao X	2017	-0.211	HCLS1	ENSG00000180353
Model-34	28979141	mRNA	Liao X	2017	0.001	IL16	ENSG00000172349
Model-34	28979141	mRNA	Liao X	2017	0.137	GVINP1	ENSG00000254838
Model-34	28979141	mRNA	Liao X	2017	-0.443	SLA2	ENSG00000101082
Model-35	33901011	mRNA	Lin H	2021	0.476	ANO1	ENSG00000131620
Model-35	33901011	mRNA	Lin H	2021	0.2455	DSG3	ENSG00000134757
Model-35	33901011	mRNA	Lin H	2021	-0.598	GIMAP1	ENSG00000213203
Model-35	33901011	mRNA	Lin H	2021	0.2339	SPINK1	ENSG00000164266
Model-36	33789615	mRNA	Liu X	2021	0.1254	ERRFI1	ENSG00000116285
Model-36	33789615	mRNA	Liu X	2021	-0.1365	IL6R	ENSG00000160712
Model-36	33789615	mRNA	Liu X	2021	0.1544	SCEL	ENSG00000136155
Model-36	33789615	mRNA	Liu X	2021	-0.44	PPP1R10	ENSG00000204569
Model-36	33789615	mRNA	Liu X	2021	-0.4412	SSX2IP	ENSG00000117155
Model-36	33789615	mRNA	Liu X	2021	-0.3397	PTOV1-AS2	ENSG00000269352
Model-36	33789615	mRNA	Liu X	2021	-0.2231	TXNL4A	ENSG00000141759
Model-37	34902987	mRNA	Liu Y	2022	0.0137	EREG	ENSG00000124882
Model-37	34902987	mRNA	Liu Y	2022	-0.1893	PRLR	ENSG00000113494
Model-38	33761652	mRNA	Lu Q	2021	1.012	CA9	ENSG00000107159
Model-38	33761652	mRNA	Lu Q	2021	-1.5512	CD1D	ENSG00000158473
Model-38	33761652	mRNA	Lu Q	2021	1.1969	SPOCK2	ENSG00000107742
Model-39	33575083	mRNA	Luo L	2021	0.06	ANKRD22	ENSG00000152766
Model-39	33575083	mRNA	Luo L	2021	0.029	ANLN	ENSG00000011426
Model-39	33575083	mRNA	Luo L	2021	0.233	ARNTL2	ENSG00000029153
Model-39	33575083	mRNA	Luo L	2021	0.078	DSG3	ENSG00000134757
Model-39	33575083	mRNA	Luo L	2021	0.044	PTPRR	ENSG00000153233
Model-39	33575083	mRNA	Luo L	2021	0.055	S100A14	ENSG00000189334
Model-39	33575083	mRNA	Luo L	2021	-0.127	TSPAN7	ENSG00000156298
Model-40	34660806	mRNA	Ma Z	2021	0.549	ENPP2	ENSG00000136960
Model-40	34660806	mRNA	Ma Z	2021	-0.2933	CYP2S1	ENSG00000167600
Model-40	34660806	mRNA	Ma Z	2021	-0.5766	FTCD	ENSG00000160282
Model-40	34660806	mRNA	Ma Z	2021	-0.4561	GPX3	ENSG00000211445
Model-40	34660806	mRNA	Ma Z	2021	0.3473	XDH	ENSG00000158125
Model-40	34660806	mRNA	Ma Z	2021	0.1947	UGT1A10	ENSG00000242515
Model-41	32860207	mRNA	Mantini G	2020	-0.6349	KHSRP	ENSG00000088247
Model-41	32860207	mRNA	Mantini G	2020	0.5423	PYGL	ENSG00000100504
Model-41	32860207	mRNA	Mantini G	2020	-0.7133	SPTBN1	ENSG00000115306
Model-42	34335699	mRNA	Mao M	2021	-0.2278	FGF17	ENSG00000158815
Model-42	34335699	mRNA	Mao M	2021	0.0258	IL20RB	ENSG00000174564
Model-42	34335699	mRNA	Mao M	2021	0.0251	MET	ENSG00000105976
Model-42	34335699	mRNA	Mao M	2021	-0.3925	NPPA	ENSG00000175206
Model-42	34335699	mRNA	Mao M	2021	4.00E-04	OASL	ENSG00000135114
Model-42	34335699	mRNA	Mao M	2021	0.0016	PLAU	ENSG00000122861
Model-42	34335699	mRNA	Mao M	2021	-0.0081	SHC2	ENSG00000129946

Model-42	34335699	mRNA	Mao M	2021	-0.1301	WFIKKN1	ENSG00000127578
Model-43	32181755	mRNA	Meng Z	2020	-0.0158	ADH1B	ENSG00000196616
Model-43	32181755	mRNA	Meng Z	2020	0.0784	CA9	ENSG00000107159
Model-43	32181755	mRNA	Meng Z	2020	-0.7636	CDHR3	ENSG00000128536
Model-43	32181755	mRNA	Meng Z	2020	-0.7038	ICAM3	ENSG00000076662
Model-43	32181755	mRNA	Meng Z	2020	0.3381	CXCL9	ENSG00000138755
Model-43	32181755	mRNA	Meng Z	2020	-0.0661	GIMAP7	ENSG00000179144
Model-43	32181755	mRNA	Meng Z	2020	-0.2418	P2RY8	ENSG00000182162
Model-43	32181755	mRNA	Meng Z	2020	-0.1146	LDLRAD1	ENSG00000203985
Model-44	32490170	mRNA	Meng Z_2	2020	-0.0744	CDHR3	ENSG00000128536
Model-44	32490170	mRNA	Meng Z_2	2020	-0.0994	CELSR3	ENSG00000008300
Model-44	32490170	mRNA	Meng Z_2	2020	0.2942	EGF	ENSG00000138798
Model-44	32490170	mRNA	Meng Z_2	2020	0.2902	FGF10	ENSG00000070193
Model-44	32490170	mRNA	Meng Z_2	2020	-0.6095	GAD1	ENSG00000128683
Model-44	32490170	mRNA	Meng Z_2	2020	0.0881	MT1H	ENSG00000205358
Model-44	32490170	mRNA	Meng Z_2	2020	-0.4845	PAH	ENSG00000171759
Model-44	32490170	mRNA	Meng Z_2	2020	0.0536	PGC	ENSG00000096088
Model-44	32490170	mRNA	Meng Z_2	2020	0.0928	NMUR2	ENSG00000132911
Model-44	32490170	mRNA	Meng Z_2	2020	-0.2357	PGM5	ENSG00000154330
Model-44	32490170	mRNA	Meng Z_2	2020	-0.334	POPDC2	ENSG00000121577
Model-44	32490170	mRNA	Meng Z_2	2020	0.065	PPFIA3	ENSG00000177380
Model-44	32490170	mRNA	Meng Z_2	2020	-0.3769	TMEM145	ENSG00000167619
Model-44	32490170	mRNA	Meng Z_2	2020	0.2099	TNNT1	ENSG00000105048
Model-44	32490170	mRNA	Meng Z_2	2020	0.3893	ZPLD1	ENSG00000170044
Model-44	32490170	mRNA	Meng Z_2	2020	0.0593	SERPINA4	ENSG00000100665
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0077	AR	ENSG00000169083
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0199	cABL	ENSG00000097007
Model-45	32895958	mRNA	Nishiwada S	2021	0.0137	cJUN	ENSG00000177606
Model-45	32895958	mRNA	Nishiwada S	2021	0.1063	HDAC1	ENSG00000116478
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0371	IRF1	ENSG00000125347
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0663	PKC-beta	ENSG00000166501
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0202	PAK2	ENSG00000180370
Model-45	32895958	mRNA	Nishiwada S	2021	-0.058	RelA	ENSG00000173039
Model-45	32895958	mRNA	Nishiwada S	2021	-0.0102	STAT1	ENSG00000115415
Model-45	32895958	mRNA	Nishiwada S	2021	0.0659	SUMO1	ENSG00000116030
Model-46	34249934	mRNA	Qian H	2021	0.3119	BIRC5	ENSG00000089685
Model-46	34249934	mRNA	Qian H	2021	0.2271	CKLF	ENSG00000217555
Model-46	34249934	mRNA	Qian H	2021	0.2728	GBP2	ENSG00000162645
Model-46	34249934	mRNA	Qian H	2021	0.1752	CRABP2	ENSG00000143320
Model-46	34249934	mRNA	Qian H	2021	0.3574	CXCL11	ENSG00000169248
Model-46	34249934	mRNA	Qian H	2021	0.0155	DKK1	ENSG00000107984
Model-46	34249934	mRNA	Qian H	2021	0.1067	EREG	ENSG00000124882
Model-46	34249934	mRNA	Qian H	2021	0.3033	FAM3C	ENSG00000196937
Model-46	34249934	mRNA	Qian H	2021	0.0672	FGFRL1	ENSG00000127418

Model-46	34249934	mRNA	Qian H	2021	-0.8333	FIGNL2	ENSG00000261308
Model-46	34249934	mRNA	Qian H	2021	-0.0764	GDF9	ENSG00000164404
Model-46	34249934	mRNA	Qian H	2021	-0.4339	RFXAP	ENSG00000133111
Model-46	34249934	mRNA	Qian H	2021	-0.067	IL32	ENSG00000008517
Model-46	34249934	mRNA	Qian H	2021	-0.1976	S100A11	ENSG00000163191
Model-46	34249934	mRNA	Qian H	2021	-0.0123	SLC22A17	ENSG00000092096
Model-46	34249934	mRNA	Qian H	2021	-0.1279	PSMB8	ENSG00000204264
Model-46	34249934	mRNA	Qian H	2021	-0.4381	PSPN	ENSG00000125650
Model-46	34249934	mRNA	Qian H	2021	-0.3633	SDC4	ENSG00000124145
Model-47	32195182	mRNA	Sahni S	2020	0.856	AGR2	ENSG00000106541
Model-47	32195182	mRNA	Sahni S	2020	0.699	CADM1	ENSG00000182985
Model-47	32195182	mRNA	Sahni S	2020	0.759	JTB	ENSG00000143543
Model-47	32195182	mRNA	Sahni S	2020	0.908	TMED2	ENSG00000086598
Model-48	29340021	mRNA	Shi G	2017	-0.3775	ACSL5	ENSG00000197142
Model-48	29340021	mRNA	Shi G	2017	0.3075	AMIGO2	ENSG00000139211
Model-48	29340021	mRNA	Shi G	2017	0.4454	ARNTL2	ENSG00000029153
Model-48	29340021	mRNA	Shi G	2017	0.6208	ASPM	ENSG00000066279
Model-48	29340021	mRNA	Shi G	2017	0.4087	BIK	ENSG00000100290
Model-48	29340021	mRNA	Shi G	2017	-0.0156	CA4	ENSG00000167434
Model-48	29340021	mRNA	Shi G	2017	0.2423	COL17A1	ENSG00000065618
Model-48	29340021	mRNA	Shi G	2017	0.2441	DKK1	ENSG00000107984
Model-48	29340021	mRNA	Shi G	2017	0.7712	ERP27	ENSG00000139055
Model-48	29340021	mRNA	Shi G	2017	-0.1605	F11	ENSG00000088926
Model-48	29340021	mRNA	Shi G	2017	-0.0249	FAM3B	ENSG00000183844
Model-48	29340021	mRNA	Shi G	2017	-0.1537	MBOAT2	ENSG00000143797
Model-48	29340021	mRNA	Shi G	2017	0.2407	MT1M	ENSG00000205364
Model-48	29340021	mRNA	Shi G	2017	0.223	SERPINB5	ENSG00000206075
Model-48	29340021	mRNA	Shi G	2017	-0.4084	SLC4A4	ENSG00000080493
Model-48	29340021	mRNA	Shi G	2017	0.2769	SPOCK1	ENSG00000152377
Model-49	29239017	lncRNA	Song J	2018	-0.4645	C9orf139	ENSG00000180539
Model-49	29239017	lncRNA	Song J	2018	0.2446	CTC-327F10.4	ENSG00000251320
Model-49	29239017	lncRNA	Song J	2018	-0.7083	MIR600HG	ENSG00000236901
Model-49	29239017	lncRNA	Song J	2018	0.2256	RP5-965G21.4	ENSG00000274414
Model-49	29239017	lncRNA	Song J	2018	-0.3029	RP11-436K8.1	ENSG00000231252
Model-50	34249078	mRNA	Song W	2021	0.0361	ABCB6	ENSG00000115657
Model-50	34249078	mRNA	Song W	2021	0.0056	ALDH3B1	ENSG00000006534
Model-50	34249078	mRNA	Song W	2021	3.00E-04	B3GNT3	ENSG00000179913
Model-50	34249078	mRNA	Song W	2021	0.0155	CACNA1H	ENSG00000196557
Model-50	34249078	mRNA	Song W	2021	0.0364	CDK1	ENSG00000170312
Model-50	34249078	mRNA	Song W	2021	0.513	CHST12	ENSG00000136213
Model-50	34249078	mRNA	Song W	2021	0.0126	NT5E	ENSG00000135318
Model-50	34249078	mRNA	Song W	2021	0.0494	KIF20A	ENSG00000112984
Model-50	34249078	mRNA	Song W	2021	0.0015	GPR87	ENSG00000138271
Model-50	34249078	mRNA	Song W	2021	0.0246	MET	ENSG00000105976

Model-50	34249078	mRNA	Song W	2021	0.0214	PGM1	ENSG00000079739
Model-51	20644708	mRNA	Stratford JK	2010	-1.214	CDX2	ENSG00000165556
Model-51	20644708	mRNA	Stratford JK	2010	3.641	FTSJ3	ENSG00000108592
Model-51	20644708	mRNA	Stratford JK	2010	-2.342	MACF1	ENSG00000127603
Model-51	20644708	mRNA	Stratford JK	2010	-1.381	RASSF4	ENSG00000107551
Model-51	20644708	mRNA	Stratford JK	2010	2.293	STAT1	ENSG00000115415
Model-51	20644708	mRNA	Stratford JK	2010	1.9	STX2	ENSG00000111450
Model-52	33226369	mRNA	Tan Z	2020	-0.1765	CD36	ENSG00000135218
Model-52	33226369	mRNA	Tan Z	2020	-0.4314	ENO3	ENSG00000108515
Model-52	33226369	mRNA	Tan Z	2020	1.0167	MET	ENSG00000105976
Model-53	34367961	mRNA	Tan Z_2	2021	0.3261	CXCL9	ENSG00000138755
Model-53	34367961	mRNA	Tan Z_2	2021	-0.071	PYHIN1	ENSG00000163564
Model-53	34367961	mRNA	Tan Z_2	2021	-0.0366	NAPSB	ENSG00000131401
Model-53	34367961	mRNA	Tan Z_2	2021	-0.2465	ZNF831	ENSG00000124203
Model-54	33827590	mRNA	Tang R	2021	-0.0088	ARRB2	ENSG00000141480
Model-54	33827590	mRNA	Tang R	2021	-0.0073	CCDC69	ENSG00000198624
Model-54	33827590	mRNA	Tang R	2021	-0.0013	CD81	ENSG00000110651
Model-54	33827590	mRNA	Tang R	2021	-0.0423	HHEX	ENSG00000152804
Model-54	33827590	mRNA	Tang R	2021	0.0284	EIF2A	ENSG00000144895
Model-54	33827590	mRNA	Tang R	2021	-0.0104	FILIP1L	ENSG00000168386
Model-54	33827590	mRNA	Tang R	2021	0.0457	GBP1	ENSG00000117228
Model-54	33827590	mRNA	Tang R	2021	0.0017	HIST1H1C	ENSG00000187837
Model-54	33827590	mRNA	Tang R	2021	-0.0046	LCP1	ENSG00000136167
Model-54	33827590	mRNA	Tang R	2021	1.00E-04	REG3A	ENSG00000172016
Model-54	33827590	mRNA	Tang R	2021	0.011	MMP28	ENSG00000271447
Model-54	33827590	mRNA	Tang R	2021	-0.0094	SDC3	ENSG00000162512
Model-54	33827590	mRNA	Tang R	2021	-0.0054	PDX1	ENSG00000139515
Model-54	33827590	mRNA	Tang R	2021	0.0023	SPRR1B	ENSG00000169469
Model-54	33827590	mRNA	Tang R	2021	2.00E-04	PPP1R15A	ENSG00000087074
Model-54	33827590	mRNA	Tang R	2021	-0.001	TUBA1A	ENSG00000167552
Model-55	33062408	mRNA	Tang R_2	2020	0.0168	HNRNPC	ENSG00000092199
Model-55	33062408	mRNA	Tang R_2	2020	0.2623	RBM15	ENSG00000162775
Model-55	33062408	mRNA	Tang R_2	2020	0.0367	IGF2BP2	ENSG00000073792
Model-56	34511988	mRNA	Tang S	2024	-0.309	MZB1	ENSG00000170476
Model-56	34511988	mRNA	Tang S	2021	-0.502	ARID5A	ENSG00000196843
Model-56	34511988	mRNA	Tang S	2022	1.226	CLEC2B	ENSG00000110852
Model-56	34511988	mRNA	Tang S	2021	-0.757	RAPGEF1	ENSG00000107263
Model-56	34511988	mRNA	Tang S	2023	-0.419	MICAL1	ENSG00000135596
Model-57	31968179	mRNA	Tian G	2020	0.683	B3GNT3	ENSG00000179913
Model-57	31968179	mRNA	Tian G	2020	0.7	MET	ENSG00000105976
Model-57	31968179	mRNA	Tian G	2020	0.662	SPAG4	ENSG00000061656
Model-58	33718118	mRNA	Wang W	2021	-0.298	ESR2	ENSG00000140009
Model-58	33718118	mRNA	Wang W	2021	0.259	IDO1	ENSG00000131203
Model-58	33718118	mRNA	Wang W	2021	0.298	IL20RB	ENSG00000174564

Model-58	33718118	mRNA	Wang W	2021	0.172	PPP3CA	ENSG00000138814
Model-58	33718118	mRNA	Wang W	2021	0.184	PLAU	ENSG00000122861
Model-59	33546693	mRNA	Wang W_2	2021	0.1587	COL7A1	ENSG00000114270
Model-59	33546693	mRNA	Wang W_2	2021	-0.6539	DENND4B	ENSG00000198837
Model-59	33546693	mRNA	Wang W_2	2021	-0.1696	ITGA7	ENSG00000135424
Model-59	33546693	mRNA	Wang W_2	2021	0.7816	NCBP2	ENSG00000114503
Model-59	33546693	mRNA	Wang W_2	2021	-0.1696	LQK1	ENSG00000198468
Model-59	33546693	mRNA	Wang W_2	2021	-0.9365	RBM14	ENSG00000239306
Model-59	33546693	mRNA	Wang W_2	2021	-0.312	ZNF709	ENSG00000242852
Model-59	33546693	mRNA	Wang W_2	2021	0.5165	SP1	ENSG00000185591
Model-60	34898275	mRNA	Wei W	2021	-0.8874	ALOX15	ENSG00000161905
Model-60	34898275	mRNA	Wei W	2021	-0.2495	MAP1LC3A	ENSG00000101460
Model-60	34898275	mRNA	Wei W	2021	0.144	PROM2	ENSG00000155066
Model-60	34898275	mRNA	Wei W	2021	0.1044	SAT1	ENSG00000130066
Model-60	34898275	mRNA	Wei W	2021	-0.4772	SAT2	ENSG00000141504
Model-60	34898275	mRNA	Wei W	2021	0.0629	TFRC	ENSG00000072274
Model-60	34898275	mRNA	Wei W	2021	0.2362	SLC39A8	ENSG00000138821
Model-61	33731794	mRNA	Wei X	2021	0.0238	ANO1	ENSG00000131620
Model-61	33731794	mRNA	Wei X	2021	-0.0244	AP1M2	ENSG00000129354
Model-61	33731794	mRNA	Wei X	2021	0.3211	BVES	ENSG00000112276
Model-61	33731794	mRNA	Wei X	2021	-0.0105	C4orf19	ENSG00000154274
Model-61	33731794	mRNA	Wei X	2021	-0.1332	CAC--1D	ENSG00000157388
Model-61	33731794	mRNA	Wei X	2021	-0.275	CCDC148	ENSG00000153237
Model-61	33731794	mRNA	Wei X	2021	0.1158	INPP4B	ENSG00000109452
Model-61	33731794	mRNA	Wei X	2021	0.1943	NET1	ENSG00000173848
Model-61	33731794	mRNA	Wei X	2021	0.2316	INSIG2	ENSG00000125629
Model-61	33731794	mRNA	Wei X	2021	0.11	SCEL	ENSG00000136155
Model-61	33731794	mRNA	Wei X	2021	-0.0932	PARP1	ENSG00000143799
Model-61	33731794	mRNA	Wei X	2021	-0.125	POLD3	ENSG00000077514
Model-61	33731794	mRNA	Wei X	2021	-0.1514	SH3RF2	ENSG00000156463
Model-61	33731794	mRNA	Wei X	2021	0.0387	VGLL1	ENSG00000102243
Model-62	32943033	mRNA	Wu C	2020	0.022	AADAC	ENSG00000114771
Model-62	32943033	mRNA	Wu C	2020	-0.989	CHFR	ENSG00000072609
Model-62	32943033	mRNA	Wu C	2020	0.007	HIST1H1C	ENSG00000187837
Model-62	32943033	mRNA	Wu C	2020	-0.32	DEF8	ENSG00000140995
Model-62	32943033	mRNA	Wu C	2020	0.041	MET	ENSG00000105976
Model-63	32596278	mRNA	Wu G	2020	0.9452	CKLF	ENSG00000217555
Model-63	32596278	mRNA	Wu G	2020	0.2968	ERAP2	ENSG00000164308
Model-63	32596278	mRNA	Wu G	2020	0.3896	EREG	ENSG00000124882
Model-64	31612115	mRNA	Wu M	2019	0.0041	ANKRD22	ENSG00000152766
Model-64	31612115	mRNA	Wu M	2019	0.0066	COL17A1	ENSG00000065618
Model-64	31612115	mRNA	Wu M	2019	0.0991	ARNTL2	ENSG00000029153
Model-64	31612115	mRNA	Wu M	2019	0.1188	CEP55	ENSG00000138180
Model-64	31612115	mRNA	Wu M	2019	-0.0076	MCOLN3	ENSG00000055732

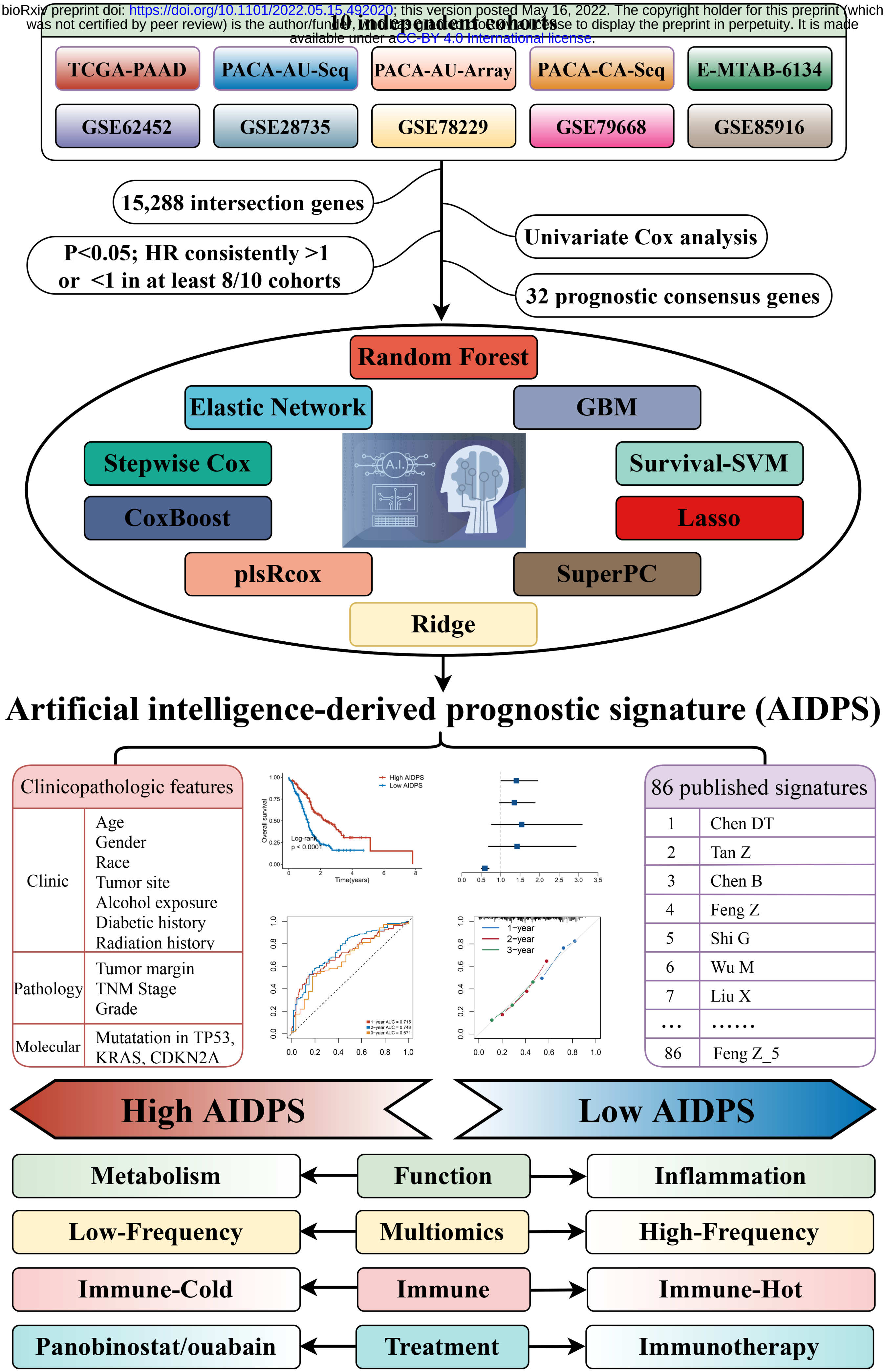
Model-64	31612115	mRNA	Wu M	2019	0.126	MET	ENSG00000105976
Model-64	31612115	mRNA	Wu M	2019	0.0479	ITGB6	ENSG00000115221
Model-64	31612115	mRNA	Wu M	2019	0.0276	KLK10	ENSG00000129451
Model-64	31612115	mRNA	Wu M	2019	-0.0397	SLC25A45	ENSG00000162241
Model-65	33033514	mRNA	Wu M_2	2020	0.0921	KIF14	ENSG00000118193
Model-65	33033514	mRNA	Wu M_2	2020	0.0097	GPR87	ENSG00000138271
Model-65	33033514	mRNA	Wu M_2	2020	0.0024	RACGAP1	ENSG00000161800
Model-65	33033514	mRNA	Wu M_2	2020	0.1136	TPX2	ENSG00000088325
Model-65	33033514	mRNA	Wu M_2	2020	0.0797	MMP28	ENSG00000271447
Model-65	33033514	mRNA	Wu M_2	2020	0.1156	RARRES3	ENSG00000133321
Model-65	33033514	mRNA	Wu M_2	2020	-0.0711	TSPAN7	ENSG00000156298
Model-66	33316381	mRNA	Xu D	2021	0.6265	AIM1	ENSG00000112297
Model-66	33316381	mRNA	Xu D	2021	0.6251	ARHGAP18	ENSG00000146376
Model-66	33316381	mRNA	Xu D	2021	0.5097	CAC--2D4	ENSG00000151062
Model-66	33316381	mRNA	Xu D	2021	-1.0684	MICAL1	ENSG00000135596
Model-66	33316381	mRNA	Xu D	2021	0.6057	DCBLD1	ENSG00000164465
Model-66	33316381	mRNA	Xu D	2021	0.1313	KLHDC7B	ENSG00000130487
Model-66	33316381	mRNA	Xu D	2021	-0.4401	KLHL32	ENSG00000186231
Model-66	33316381	mRNA	Xu D	2021	-0.4933	UNC13B	ENSG00000198722
Model-66	33316381	mRNA	Xu D	2021	-1.1333	TSPYL4	ENSG00000187189
Model-67	33176521	mRNA	Xu F	2021	-0.0286	ALKBH5	ENSG00000091542
Model-67	33176521	mRNA	Xu F	2021	-0.132	METTL14	ENSG00000145388
Model-67	33176521	mRNA	Xu F	2021	-0.087	METTL3	ENSG00000165819
Model-67	33176521	mRNA	Xu F	2021	0.233	KIAA1429	ENSG00000164944
Model-67	33176521	mRNA	Xu F	2021	-0.035	YTHDF1	ENSG00000149658
Model-68	34717651	mRNA	Xu Q	2021	-0.5483	ABCB4	ENSG00000005471
Model-68	34717651	mRNA	Xu Q	2021	-7.7451	GH1	ENSG00000259384
Model-68	34717651	mRNA	Xu Q	2021	-0.0924	FAM53B	ENSG00000189319
Model-68	34717651	mRNA	Xu Q	2021	-0.5285	INTU	ENSG00000164066
Model-68	34717651	mRNA	Xu Q	2021	-0.432	SPINK2	ENSG00000128040
Model-69	30643453	mRNA	Yan X	2018	0.4623	CDC6	ENSG00000094804
Model-69	30643453	mRNA	Yan X	2018	0.4106	IGF2BP2	ENSG00000073792
Model-69	30643453	mRNA	Yan X	2018	0.3707	KNTC1	ENSG00000184445
Model-69	30643453	mRNA	Yan X	2018	-0.4705	LYRM1	ENSG00000102897
Model-70	33968691	mRNA	Yang B	2021	2.8607	CHST11	ENSG00000171310
Model-70	33968691	mRNA	Yang B	2021	-0.5144	CRHBP	ENSG00000145708
Model-70	33968691	mRNA	Yang B	2021	0.7636	CXCL5	ENSG00000163735
Model-70	33968691	mRNA	Yang B	2021	-1.1445	GALNT16	ENSG00000100626
Model-70	33968691	mRNA	Yang B	2021	0.4442	MUC16	ENSG00000181143
Model-70	33968691	mRNA	Yang B	2021	-2.1059	NOD2	ENSG00000167207
Model-70	33968691	mRNA	Yang B	2021	1.0661	ZNF683	ENSG00000176083
Model-71	33386701	mRNA	Ye Y	2020	-0.0666	CA8	ENSG00000178538
Model-71	33386701	mRNA	Ye Y	2020	0.0413	CEP55	ENSG00000138180
Model-71	33386701	mRNA	Ye Y	2020	-0.2189	GNB3	ENSG00000111664

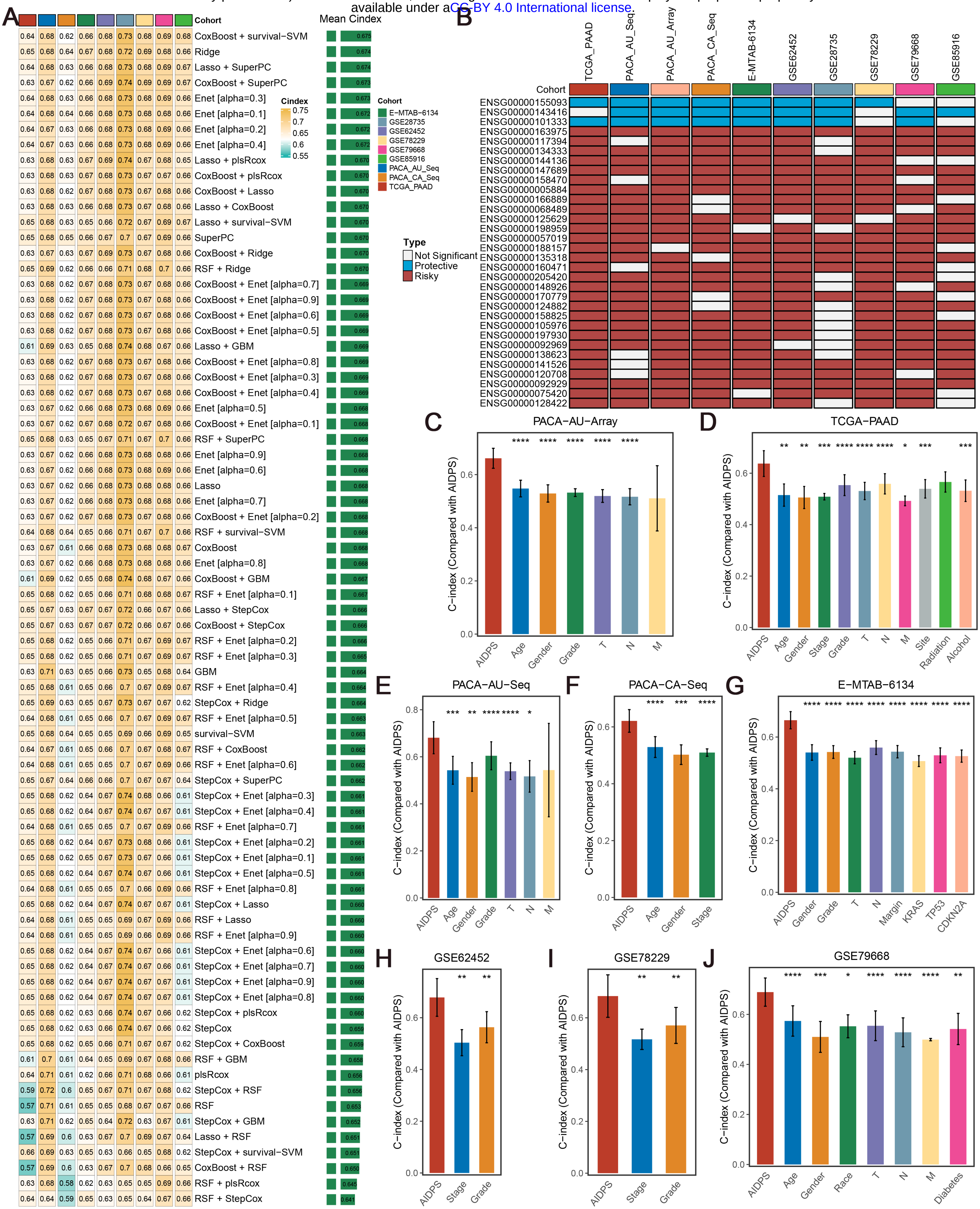
Model-71	33386701	mRNA	Ye Y	2020	-0.0339	SGSM2	ENSG00000141258
Model-72	34262410	mRNA	Yu J	2021	1.3721	APOL1	ENSG00000100342
Model-72	34262410	mRNA	Yu J	2021	0.769	PTK6	ENSG00000101213
Model-72	34262410	mRNA	Yu J	2021	-0.863	RAB24	ENSG00000169228
Model-72	34262410	mRNA	Yu J	2021	0.442	TP63	ENSG00000073282
Model-73	34820374	mRNA	Yu X_2	2021	0.34	CD44	ENSG00000026508
Model-73	34820374	mRNA	Yu X_2	2021	0.216	MT1G	ENSG00000125144
Model-73	34820374	mRNA	Yu X_2	2021	0.225	SAT1	ENSG00000130066
Model-73	34820374	mRNA	Yu X_2	2021	0.05	PTGS2	ENSG00000073756
Model-73	34820374	mRNA	Yu X_2	2021	0.186	TFRC	ENSG00000072274
Model-73	34820374	mRNA	Yu X_2	2021	0.207	STEAP3	ENSG00000115107
Model-74	34233576	lncRNA	Yuan Q	2021	0.3069	AC099850.3	ENSG00000265415
Model-74	34233576	lncRNA	Yuan Q	2021	-0.7104	AL513165.1	ENSG00000256966
Model-74	34233576	lncRNA	Yuan Q	2021	0.2433	AP005233.2	ENST00000561588
Model-74	34233576	lncRNA	Yuan Q	2021	-0.6473	PTOV1-AS2	ENSG00000269352
Model-74	34233576	lncRNA	Yuan Q	2021	0.1555	UCA1	ENSG00000214049
Model-75	34490033	mRNA	Yuan Q_2	2021	0.5671	SP1	ENSG00000185591
Model-75	34490033	mRNA	Yuan Q_2	2021	-1.147	SAFB	ENSG00000160633
Model-75	34490033	mRNA	Yuan Q_2	2021	0.4489	SERTAD3	ENSG00000167565
Model-76	32203889	mRNA	Yue P	2020	0.1063	CASP4	ENSG00000196954
Model-76	32203889	mRNA	Yue P	2020	0.0297	EIF4G1	ENSG00000114867
Model-76	32203889	mRNA	Yue P	2020	0.0416	CHMP2B	ENSG00000083937
Model-76	32203889	mRNA	Yue P	2020	-0.0498	GABARAP	ENSG00000170296
Model-76	32203889	mRNA	Yue P	2020	-0.1453	RAB24	ENSG00000169228
Model-76	32203889	mRNA	Yue P	2020	-0.3239	RPTOR	ENSG00000141564
Model-76	32203889	mRNA	Yue P	2020	0.0983	NCKAP1	ENSG00000061676
Model-76	32203889	mRNA	Yue P	2020	-0.2125	PELP1	ENSG00000141456
Model-76	32203889	mRNA	Yue P	2020	0.0078	TNFSF10	ENSG00000121858
Model-76	32203889	mRNA	Yue P	2020	-0.1622	WIPI2	ENSG00000157954
Model-77	33869254	mRNA	Zhang C	2021	-0.3606	AGT	ENSG00000135744
Model-77	33869254	mRNA	Zhang C	2021	-0.6498	CMTM6	ENSG00000091317
Model-77	33869254	mRNA	Zhang C	2021	0.3408	AREG	ENSG00000109321
Model-77	33869254	mRNA	Zhang C	2021	-0.8476	B2M	ENSG00000166710
Model-77	33869254	mRNA	Zhang C	2021	0.4845	CXCL9	ENSG00000138755
Model-77	33869254	mRNA	Zhang C	2021	0.5668	LTBP1	ENSG00000049323
Model-77	33869254	mRNA	Zhang C	2021	0.5127	HBEGF	ENSG00000113070
Model-77	33869254	mRNA	Zhang C	2021	0.8248	MET	ENSG00000105976
Model-77	33869254	mRNA	Zhang C	2021	0.1701	IL22RA1	ENSG00000142677
Model-77	33869254	mRNA	Zhang C	2021	0.4116	OAS1	ENSG00000089127
Model-77	33869254	mRNA	Zhang C	2021	0.127	PI3	ENSG00000124102
Model-77	33869254	mRNA	Zhang C	2021	0.1783	LCN2	ENSG00000148346
Model-77	33869254	mRNA	Zhang C	2021	-0.5991	NRP2	ENSG00000118257
Model-77	33869254	mRNA	Zhang C	2021	0.512	OASL	ENSG00000135114
Model-77	33869254	mRNA	Zhang C	2021	0.67	PAK3	ENSG00000077264

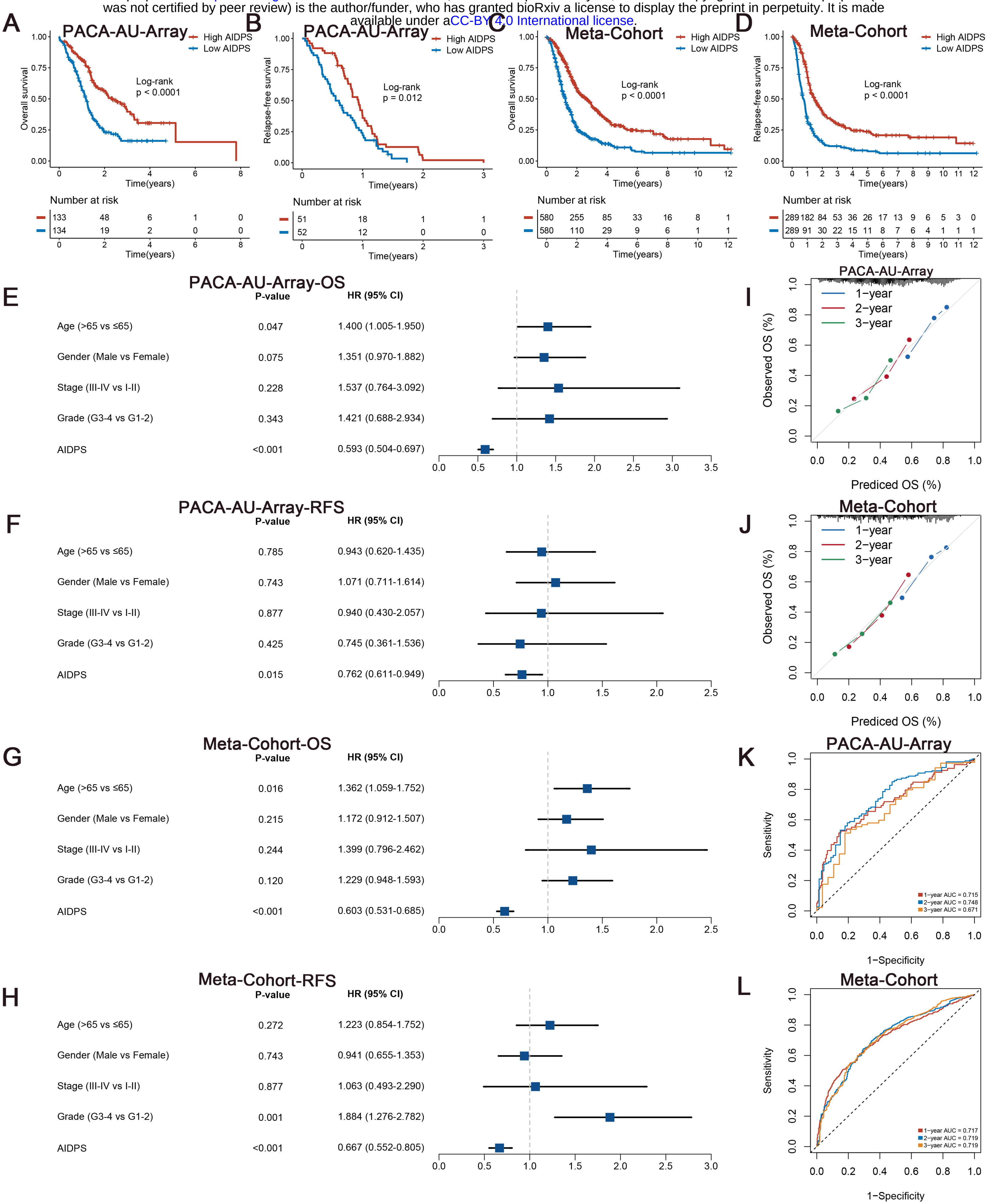
Model-77	33869254	mRNA	Zhang C	2021	-1.284	PLAUR	ENSG00000011422
Model-77	33869254	mRNA	Zhang C	2021	0.2425	PTGS2	ENSG00000073756
Model-77	33869254	mRNA	Zhang C	2021	0.3596	TMSB10	ENSG00000034510
Model-77	33869254	mRNA	Zhang C	2021	0.2132	SPP1	ENSG00000118785
Model-77	33869254	mRNA	Zhang C	2021	0.8734	S100A16	ENSG00000188643
Model-77	33869254	mRNA	Zhang C	2021	-0.6231	SDC4	ENSG00000124145
Model-78	34124046	mRNA	Zhang F	2021	-0.269	ARNT2	ENSG00000172379
Model-78	34124046	mRNA	Zhang F	2021	0.387	LINC01559	ENSG00000180861
Model-78	34124046	mRNA	Zhang F	2021	-0.395	SLC26A11	ENSG00000181045
Model-78	34124046	mRNA	Zhang F	2021	-0.374	TRIM67	ENSG00000119283
Model-78	34124046	mRNA	Zhang F	2021	0.314	UCA1	ENSG00000214049
Model-79	31031865	lncRNA	Zhang H	2019	0.028	ABHD11-AS1	ENSG00000225969
Model-79	31031865	lncRNA	Zhang H	2019	-0.453	AP000254.8	ENSG00000273271
Model-79	31031865	lncRNA	Zhang H	2019	0.198	CASC8	ENSG00000246228
Model-79	31031865	lncRNA	Zhang H	2019	-0.353	CTC-429P9.3	ENSG00000269044
Model-79	31031865	lncRNA	Zhang H	2019	-0.534	CTD-2186M15.3	ENSG00000272086
Model-79	31031865	lncRNA	Zhang H	2019	0.149	CYTOR	ENSG00000222041
Model-79	31031865	lncRNA	Zhang H	2019	0.108	LINC00941	ENSG00000235884
Model-79	31031865	lncRNA	Zhang H	2019	-0.335	LINC01089	ENSG00000212694
Model-79	31031865	lncRNA	Zhang H	2019	0.332	MIR4435-2HG	ENSG00000172965
Model-79	31031865	lncRNA	Zhang H	2019	-0.26	RP5-1085F17.3	ENSG00000260257
Model-79	31031865	lncRNA	Zhang H	2019	-0.197	RP5-890O3.9	ENSG00000240731
Model-79	31031865	lncRNA	Zhang H	2019	0.009	UCA1	ENSG00000214049
Model-80	33691542	mRNA	Zhang LL	2021	0.3543	ARNTL	ENSG00000133794
Model-80	33691542	mRNA	Zhang LL	2021	-0.2464	CRY1	ENSG00000008405
Model-80	33691542	mRNA	Zhang LL	2021	-0.4626	CRY2	ENSG00000121671
Model-80	33691542	mRNA	Zhang LL	2021	-0.2177	CSNK1D	ENSG00000141551
Model-80	33691542	mRNA	Zhang LL	2021	0.1566	CSNK1E	ENSG00000213923
Model-80	33691542	mRNA	Zhang LL	2021	-0.1878	CUL1	ENSG00000055130
Model-80	33691542	mRNA	Zhang LL	2021	-0.2242	DBP	ENSG00000105516
Model-80	33691542	mRNA	Zhang LL	2021	0.1836	NR1D1	ENSG00000126368
Model-80	33691542	mRNA	Zhang LL	2021	0.0974	RORA	ENSG00000069667
Model-80	33691542	mRNA	Zhang LL	2021	0.6332	RORB	ENSG00000198963
Model-81	33748143	mRNA	Zhang X	2021	0.0183	CDK1	ENSG00000170312
Model-81	33748143	mRNA	Zhang X	2021	0.079	CRABP2	ENSG00000143320
Model-81	33748143	mRNA	Zhang X	2021	0.3113	NUSAP1	ENSG00000137804
Model-81	33748143	mRNA	Zhang X	2021	0.2482	PERP	ENSG00000112378
Model-81	33748143	mRNA	Zhang X	2021	0.1529	TOP2A	ENSG00000131747
Model-82	33015155	mRNA	Zhang Z	2020	-0.0481	CHGA	ENSG00000100604
Model-82	33015155	mRNA	Zhang Z	2020	0.0402	COL17A1	ENSG00000065618
Model-82	33015155	mRNA	Zhang Z	2020	0.0021	LAMC2	ENSG00000058085
Model-82	33015155	mRNA	Zhang Z	2020	0.0697	ITGB6	ENSG00000115221
Model-82	33015155	mRNA	Zhang Z	2020	0.0063	S100P	ENSG00000163993
Model-83	31595147	mRNA	Zhou C	2019	0.012	ANKRD22	ENSG00000152766

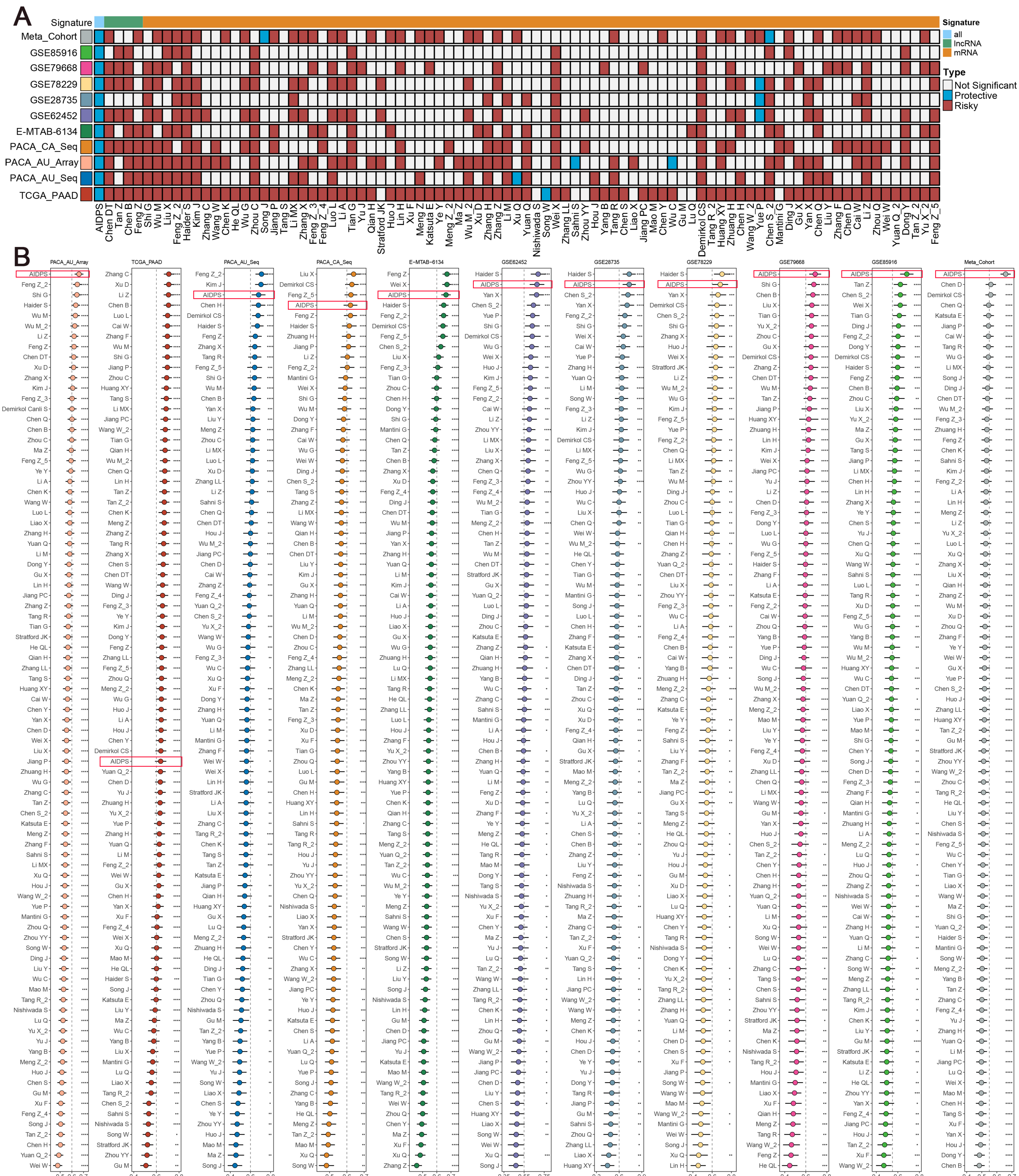
Model-83	31595147	mRNA	Zhou C	2019	0.332	MET	ENSG00000105976
Model-83	31595147	mRNA	Zhou C	2019	0.0034	IGF2BP3	ENSG00000136231
Model-83	31595147	mRNA	Zhou C	2019	0.0126	INPP4B	ENSG00000109452
Model-83	31595147	mRNA	Zhou C	2019	0.0677	KYNU	ENSG00000115919
Model-83	31595147	mRNA	Zhou C	2019	0.0413	TOP2A	ENSG00000131747
Model-84	34587439	mRNA	Zhou Q	2021	0.8544	APOL1	ENSG00000100342
Model-84	34587439	mRNA	Zhou Q	2021	-0.8465	ATG16L2	ENSG00000168010
Model-84	34587439	mRNA	Zhou Q	2021	2.0055	G--I3	ENSG00000065135
Model-84	34587439	mRNA	Zhou Q	2021	0.7124	PTK6	ENSG00000101213
Model-85	31706267	mRNA	Zhou YY	2019	-1.4	CEL	ENSG00000170835
Model-85	31706267	mRNA	Zhou YY	2019	1.321	CPA1	ENSG00000091704
Model-85	31706267	mRNA	Zhou YY	2019	0.454	POSTN	ENSG00000133110
Model-85	31706267	mRNA	Zhou YY	2019	1.011	PM20D1	ENSG00000162877
Model-86	33073486	mRNA	Zhuang H	2020	0.0036	ITGB1	ENSG00000150093
Model-86	33073486	mRNA	Zhuang H	2020	0.0013	ITGB4	ENSG00000132470
Model-86	33073486	mRNA	Zhuang H	2020	0.0032	ITGB5	ENSG00000082781
Model-86	33073486	mRNA	Zhuang H	2020	0.0055	ITGB6	ENSG00000115221

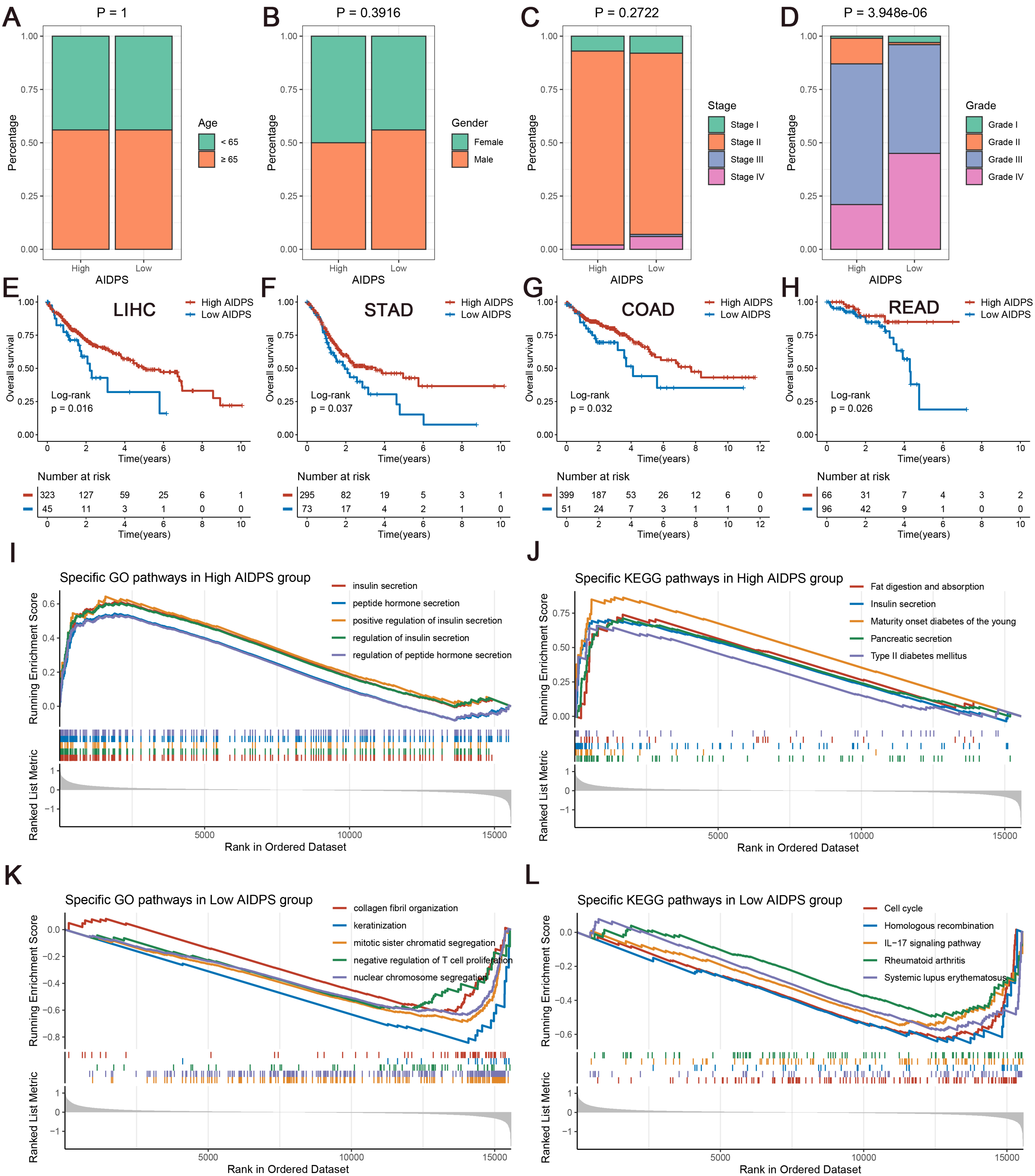
1008

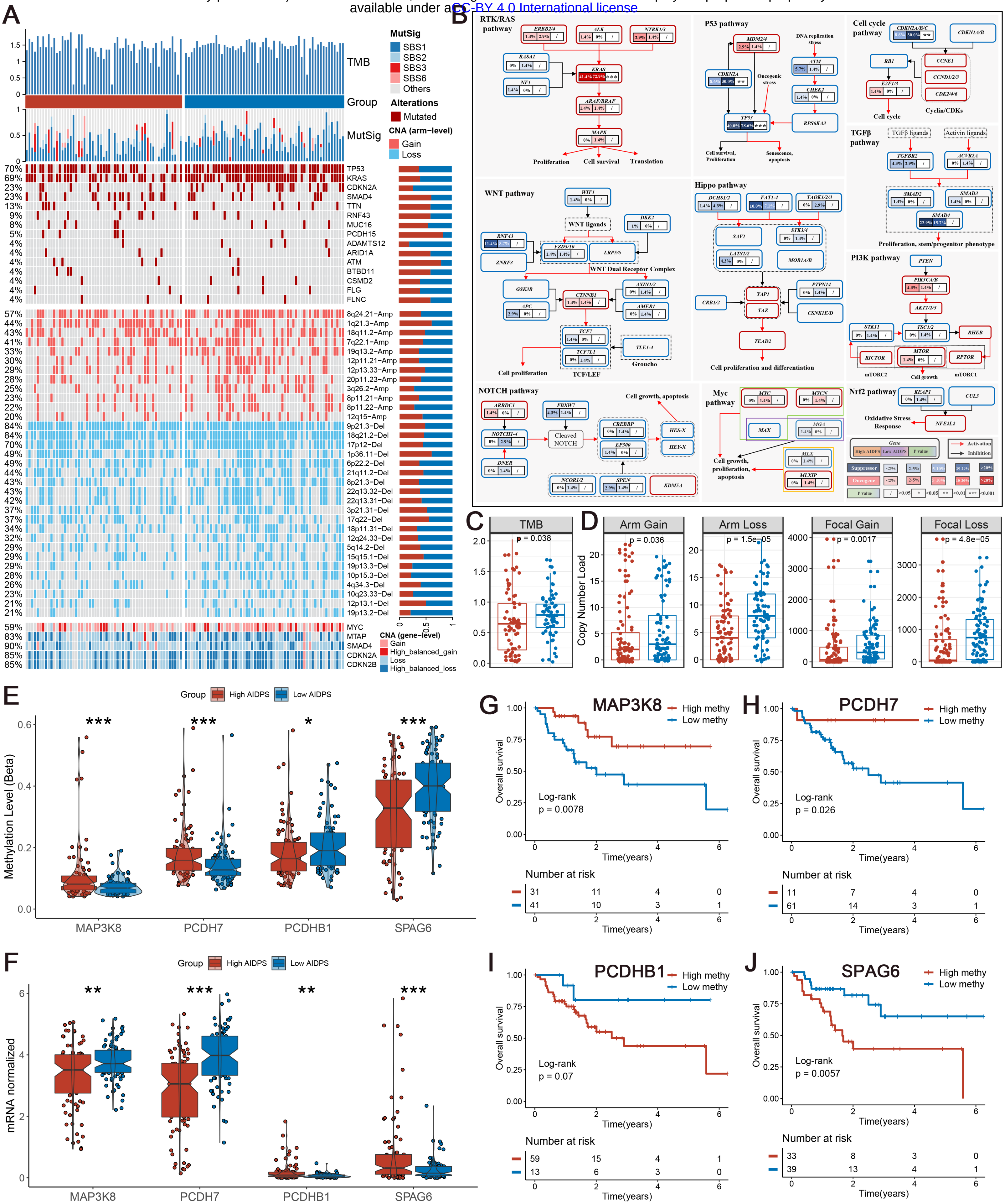

10 independent cohorts
TCGA-PAAD
PACA-AU-Seq
PACA-AU-Array
PACA-CA-Seq
E-MTAB-6134
GSE62452
GSE28735
GSE78229
GSE79668
GSE85916
15,288 intersection genes
Univariate Cox analysis
P<0.05; HR consistently >1 or <1 in at least 8/10 cohorts
32 prognostic consensus genes
Random Forest
Elastic Network
GBM
Stepwise Cox
Survival-SVM
CoxBoost
Lasso
plsRcox
SuperPC
Ridge
Artificial intelligence-derived prognostic signature (AIDPS)
Clinicopathologic features
Clinic
Age
Gender
Race
Tumor site
Alcohol exposure
Diabetic history
Radiation history
Pathology
Tumor margin
TNM Stage
Grade
Molecular
Mutatation in TP53, KRAS, CDKN2A
High AIDPS
Low AIDPS
Overall survival
Log-rank p < 0.0001
Time(years)
1-year
2-year
3-year
86 published signatures
1 Chen DT
2 Tan Z
3 Chen B
4 Feng Z
5 Shi G
6 Wu M
7 Liu X
… ……
86 Feng Z_5
High AIDPS
Low AIDPS
Metabolism
Function
Inflammation
Low-Frequency
Multiomics
High-Frequency
Immune-Cold
Immune
Immune-Hot
Panobinostat/ouabain
Treatment
Immunotherapy

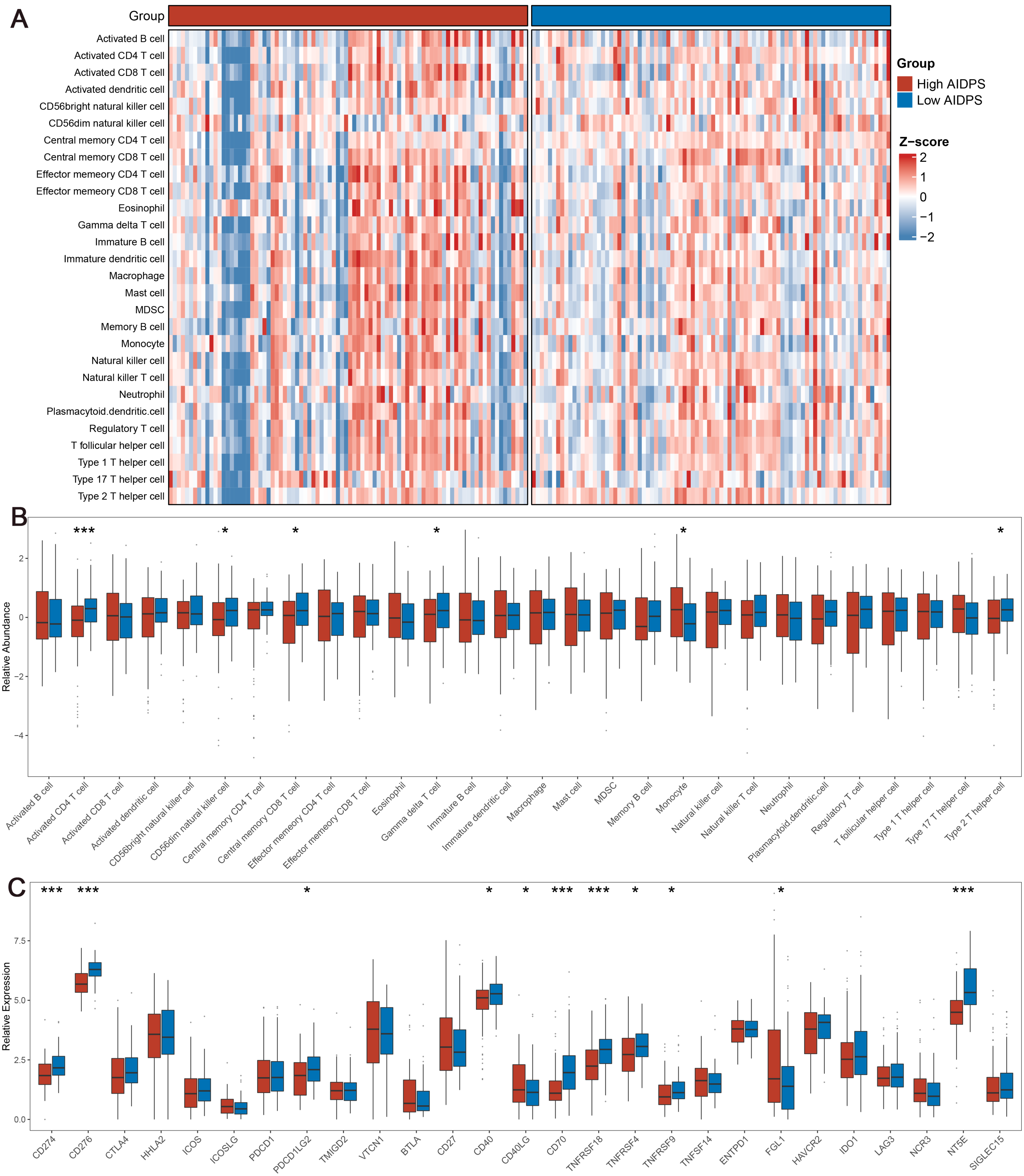

A
Cohort
Mean Cindex
Cindex
Cohort
E−MTAB−6134
GSE28735
GSE62452
GSE78229
GSE79668
GSE85916
PACA_AU_Seq
PACA_CA_Seq
TCGA_PAAD
CoxBoost + survival−SVM
Ridge
Lasso + SuperPC
CoxBoost + SuperPC
Enet [alpha=0.3]
Enet [alpha=0.1]
Enet [alpha=0.2]
Enet [alpha=0.4]
Lasso + plsRcox
CoxBoost + plsRcox
CoxBoost + Lasso
Lasso + CoxBoost
Lasso + survival−SVM
SuperPC
CoxBoost + Ridge
RSF + Ridge
CoxBoost + Enet [alpha=0.7]
CoxBoost + Enet [alpha=0.9]
CoxBoost + Enet [alpha=0.6]
CoxBoost + Enet [alpha=0.5]
Lasso + GBM
CoxBoost + Enet [alpha=0.8]
CoxBoost + Enet [alpha=0.3]
CoxBoost + Enet [alpha=0.4]
Enet [alpha=0.5]
CoxBoost + Enet [alpha=0.1]
RSF + SuperPC
Enet [alpha=0.9]
Enet [alpha=0.6]
Lasso
Enet [alpha=0.7]
CoxBoost + Enet [alpha=0.2]
RSF + survival−SVM
CoxBoost
Enet [alpha=0.8]
CoxBoost + GBM
RSF + Enet [alpha=0.1]
Lasso + StepCox
CoxBoost + StepCox
RSF + Enet [alpha=0.2]
RSF + Enet [alpha=0.3]
GBM
RSF + Enet [alpha=0.4]
StepCox + Ridge
RSF + Enet [alpha=0.5]
survival−SVM
RSF + CoxBoost
RSF + Enet [alpha=0.6]
StepCox + SuperPC
StepCox + Enet [alpha=0.3]
StepCox + Enet [alpha=0.4]
RSF + Enet [alpha=0.7]
StepCox + Enet [alpha=0.2]
StepCox + Enet [alpha=0.1]
StepCox + Enet [alpha=0.5]
RSF + Enet [alpha=0.8]
StepCox + Lasso
RSF + Lasso
RSF + Enet [alpha=0.9]
StepCox + Enet [alpha=0.6]
StepCox + Enet [alpha=0.7]
StepCox + Enet [alpha=0.9]
StepCox + Enet [alpha=0.8]
StepCox + plsRcox
StepCox
StepCox + CoxBoost
RSF + GBM
plsRcox
StepCox + RSF
RSF
StepCox + GBM
Lasso + RSF
StepCox + survival−SVM
CoxBoost + RSF
RSF + plsRcox
RSF + StepCox
B
TCGA_PAAD
PACA_AU_Seq
PACA_AU_Array
PACA_CA_Seq
E-MTAB-6134
GSE62452
GSE28735
GSE78229
GSE79668
GSE85916
Cohort
Type
Not Significant
Protective
Risky
ENSG00000155093
ENSG00000143416
ENSG00000101333
ENSG00000163975
ENSG00000117394
ENSG00000134333
ENSG00000144136
ENSG00000147689
ENSG00000158470
ENSG00000005884
ENSG00000166889
ENSG00000068489
ENSG00000125629
ENSG00000198959
ENSG00000057019
ENSG00000188157
ENSG00000135318
ENSG00000160471
ENSG00000205420
ENSG00000148926
ENSG00000170779
ENSG00000124882
ENSG00000158825
ENSG00000105976
ENSG00000197930
ENSG00000092969
ENSG00000138623
ENSG00000141526
ENSG00000120708
ENSG00000092929
ENSG00000075420
ENSG00000128422
C
PACA-AU-Array
C−index (Compared with AIDPS)
AIDPS
Age
Gender
Grade
T
N
M
D
TCGA−PAAD
AIDPS
Age
Gender
Stage
Grade
T
N
M
Site
Radiation
Alcohol
E
PACA-AU-Seq
AIDPS
Age
Gender
Grade
T
N
M
F
PACA-CA-Seq
AIDPS
Age
Gender
Stage
G
E−MTAB-6134
AIDPS
Gender
Grade
T
N
Margin
KRAS
TP53
CDKN2A
H
GSE62452
AIDPS
Stage
Grade
I
GSE78229
AIDPS
Stage
Grade
J
GSE79668
AIDPS
Age
Gender
Race
T
N
M
Diabetes

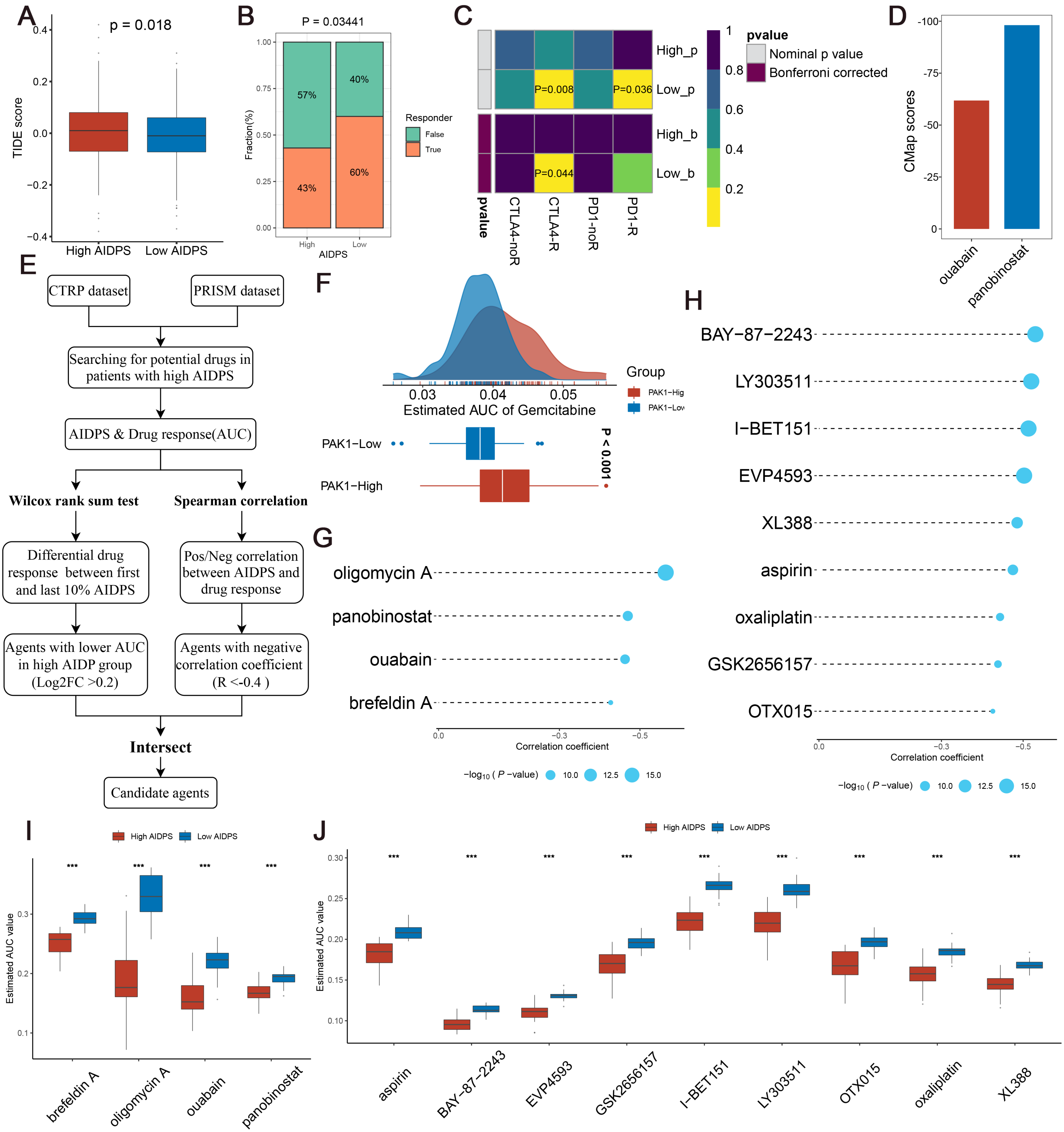

A
PACA-AU-Array
High AIDPS
Low AIDPS
Overall survival
Log-rank
p < 0.0001
Time(years)
Number at risk
133 48 6 1 0
134 19 2 0 0
B
PACA-AU-Array
Relapse-free survival
Log-rank
p = 0.012
51 18 1 1
52 12 0 0
C
Meta-Cohort
Overall survival
Log-rank
p < 0.0001
580 255 85 33 16 8 1
580 110 29 9 6 1 1
D
Meta-Cohort
Relapse-free survival
Log-rank
p < 0.0001
289 182 84 53 36 26 17 13 9 6 5 3 0
289 91 30 22 15 11 8 7 6 4 1 1 1
E
PACA-AU-Array-OS
P-value
HR (95% CI)
Age (>65 vs ≤65) 0.047 1.400 (1.005-1.950)
Gender (Male vs Female) 0.075 1.351 (0.970-1.882)
Stage (III-IV vs I-II) 0.228 1.537 (0.764-3.092)
Grade (G3-4 vs G1-2) 0.343 1.421 (0.688-2.934)
AIDPS <0.001 0.593 (0.504-0.697)
F
PACA-AU-Array-RFS
Age (>65 vs ≤65) 0.785 0.943 (0.620-1.435)
Gender (Male vs Female) 0.743 1.071 (0.711-1.614)
Stage (III-IV vs I-II) 0.877 0.940 (0.430-2.057)
Grade (G3-4 vs G1-2) 0.425 0.745 (0.361-1.536)
AIDPS 0.015 0.762 (0.611-0.949)
G
Meta-Cohort-OS
Age (>65 vs ≤65) 0.016 1.362 (1.059-1.752)
Gender (Male vs Female) 0.215 1.172 (0.912-1.507)
Stage (III-IV vs I-II) 0.244 1.399 (0.796-2.462)
Grade (G3-4 vs G1-2) 0.120 1.229 (0.948-1.593)
AIDPS <0.001 0.603 (0.531-0.685)
H
Meta-Cohort-RFS
Age (>65 vs ≤65) 0.272 1.223 (0.854-1.752)
Gender (Male vs Female) 0.743 0.941 (0.655-1.353)
Stage (III-IV vs I-II) 0.877 1.063 (0.493-2.290)
Grade (G3-4 vs G1-2) 0.001 1.884 (1.276-2.782)
AIDPS <0.001 0.667 (0.552-0.805)
I
PACA-AU-Array
1−year
2−year
3−year
Observed OS (%)
Prediced OS (%)
J
Meta-Cohort
K
PACA-AU-Array
Sensitivity
1−Specificity
1−year AUC = 0.715
2−year AUC = 0.748
3−yaer AUC = 0.671
L
Meta-Cohort
1−year AUC = 0.717
2−year AUC = 0.719
3−yaer AUC = 0.719


A
Signature
Meta_Cohort
GSE85916
GSE79668
GSE78229
GSE28735
GSE62452
E-MTAB-6134
PACA_CA_Seq
PACA_AU_Array
PACA_AU_Seq
TCGA_PAAD
Signature
all
lncRNA
mRNA
Type
Not Significant
Protective
Risky
B
PACA_AU_Array
TCGA_PAAD
PACA_AU_Seq
PACA_CA_Seq
E-MTAB-6134
GSE62452
GSE28735
GSE78229
GSE79668
GSE85916
Meta_Cohort
AIDPS


A
P = 1
Age
< 65
≥ 65
B
P = 0.3916
Gender
Female
Male
C
P = 0.2722
Stage
Stage I
Stage II
Stage III
Stage IV
D
P = 3.948e-06
Grade
Grade I
Grade II
Grade III
Grade IV
Percentage
AIDPS
High
Low
E
LIHC
Log-rank
p = 0.016
F
STAD
Log-rank
p = 0.037
G
COAD
Log-rank
p = 0.032
H
READ
Log-rank
p = 0.026
High AIDPS
Low AIDPS
Overall survival
Time(years)
Number at risk
I
Specific GO pathways in High AIDPS group
insulin secretion
peptide hormone secretion
positive regulation of insulin secretion
regulation of insulin secretion
regulation of peptide hormone secretion
J
Specific KEGG pathways in High AIDPS group
Fat digestion and absorption
Insulin secretion
Maturity onset diabetes of the young
Pancreatic secretion
Type II diabetes mellitus
K
Specific GO pathways in Low AIDPS group
collagen fibril organization
keratinization
mitotic sister chromatid segregation
negative regulation of T cell proliferation
nuclear chromosome segregation
L
Specific KEGG pathways in Low AIDPS group
Cell cycle
Homologous recombination
IL−17 signaling pathway
Rheumatoid arthritis
Systemic lupus erythematosus
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset


A
TMB
Group
MutSig
MutSig
SBS1
SBS2
SBS3
SBS6
Others
Alterations
Mutated
CNA (arm−level)
Gain
Loss
TP53
KRAS
CDKN2A
SMAD4
TTN
RNF43
MUC16
PCDH15
ADAMTS12
ARID1A
ATM
BTBD11
CSMD2
FLG
FLNC
8q24.21−Amp
1q21.3−Amp
18q11.2−Amp
7q22.1−Amp
19q13.2−Amp
12p11.21−Amp
12p13.33−Amp
20p11.23−Amp
3q26.2−Amp
8p11.21−Amp
8p11.22−Amp
12q15−Amp
9p21.3−Del
18q21.2−Del
17p12−Del
1p36.11−Del
6p22.2−Del
21q11.2−Del
8p21.3−Del
22q13.32−Del
22q13.31−Del
3p21.31−Del
17q22−Del
18p11.31−Del
12q24.33−Del
5q14.2−Del
15q15.1−Del
19p13.3−Del
10p15.3−Del
4q34.3−Del
10q23.33−Del
12p13.1−Del
19p13.2−Del
MYC
MTAP
SMAD4
CDKN2A
CDKN2B
CNA (gene−level)
Gain
High_balanced_gain
Loss
High_balanced_loss
B
RTK/RAS pathway
P53 pathway
Cell cycle pathway
WNT pathway
Hippo pathway
TGFβ pathway
PI3K pathway
NOTCH pathway
Myc pathway
Nrf2 pathway
Proliferation
Cell survival
Translation
Cell proliferation
Cell proliferation and differentiation
Proliferation, stem/progenitor phenotype
C
TMB
p = 0.038
D
Copy Number Load
Arm Gain
p = 0.036
Arm Loss
p = 1.5e−05
Focal Gain
p = 0.0017
Focal Loss
p = 4.8e−05
E
Group
High AIDPS
Low AIDPS
Methylation Level (Beta)
MAP3K8
PCDH7
PCDHB1
SPAG6
F
Group
High AIDPS
Low AIDPS
mRNA normalized
MAP3K8
PCDH7
PCDHB1
SPAG6
G
MAP3K8
High methy
Low methy
Overall survival
Log-rank
p = 0.0078
Time(years)
Number at risk
31 11 4 0
41 10 3 1
H
PCDH7
High methy
Low methy
Log-rank
p = 0.026
Number at risk
11 7 4 0
61 14 3 1
I
PCDHB1
High methy
Low methy
Log-rank
p = 0.07
Number at risk
59 15 4 1
13 6 3 0
J
SPAG6
High methy
Low methy
Log-rank
p = 0.0057
Number at risk
33 8 3 0
39 13 4 1


A
Group
High AIDPS
Low AIDPS
Z-score
Activated B cell
Activated CD4 T cell
Activated CD8 T cell
Activated dendritic cell
CD56bright natural killer cell
CD56dim natural killer cell
Central memory CD4 T cell
Central memory CD8 T cell
Effector memeory CD4 T cell
Effector memeory CD8 T cell
Eosinophil
Gamma delta T cell
Immature B cell
Immature dendritic cell
Macrophage
Mast cell
MDSC
Memory B cell
Monocyte
Natural killer cell
Natural killer T cell
Neutrophil
Plasmacytoid.dendritic.cell
Regulatory T cell
T follicular helper cell
Type 1 T helper cell
Type 17 T helper cell
Type 2 T helper cell
B
Relative Abundance
C
Relative Expression
CD274
CD276
CTLA4
HHLA2
ICOS
ICOSLG
PDCD1
PDCD1LG2
TMIGD2
VTCN1
BTLA
CD27
CD40
CD40LG
CD70
TNFRSF18
TNFRSF4
TNFRSF9
TNFSF14
ENTPD1
FGL1
HAVCR2
IDO1
LAG3
NCR3
NT5E
SIGLEC15


A
p = 0.018
TIDE score
High AIDPS
Low AIDPS
B
P = 0.03441
Fraction(%)
57%
43%
40%
60%
Responder
False
True
AIDPS
High
Low
C
High_p
Low_p
High_b
Low_b
P=0.008
P=0.036
P=0.044
pvalue
CTLA4-noR
CTLA4-R
PD1-noR
PD1-R
pvalue
Nominal p value
Bonferroni corrected
D
CMap scores
ouabain
panobinostat
E
CTRP dataset
PRISM dataset
Searching for potential drugs in patients with high AIDPS
AIDPS & Drug response(AUC)
Wilcox rank sum test
Spearman correlation
Differential drug response between first and last 10% AIDPS
Pos/Neg correlation between AIDPS and drug response
Agents with lower AUC in high AIDP group (Log2FC >0.2)
Agents with negative correlation coefficient (R <-0.4)
Intersect
Candidate agents
F
Estimated AUC of Gemcitabine
Group
PAK1-Low
PAK1-High
P < 0.001
G
oligomycin A
panobinostat
ouabain
brefeldin A
Correlation coefficient
H
BAY-87-2243
LY303511
I-BET151
EVP4593
XL388
aspirin
oxaliplatin
GSK2656157
OTX015
Correlation coefficient
I
High AIDPS
Low AIDPS
Estimated AUC value
brefeldin A
oligomycin A
ouabain
panobinostat
J
High AIDPS
Low AIDPS
Estimated AUC value
aspirin
BAY-87-2243
EVP4593
GSK2656157
I-BET151
LY303511
OTX015
oxaliplatin
XL388

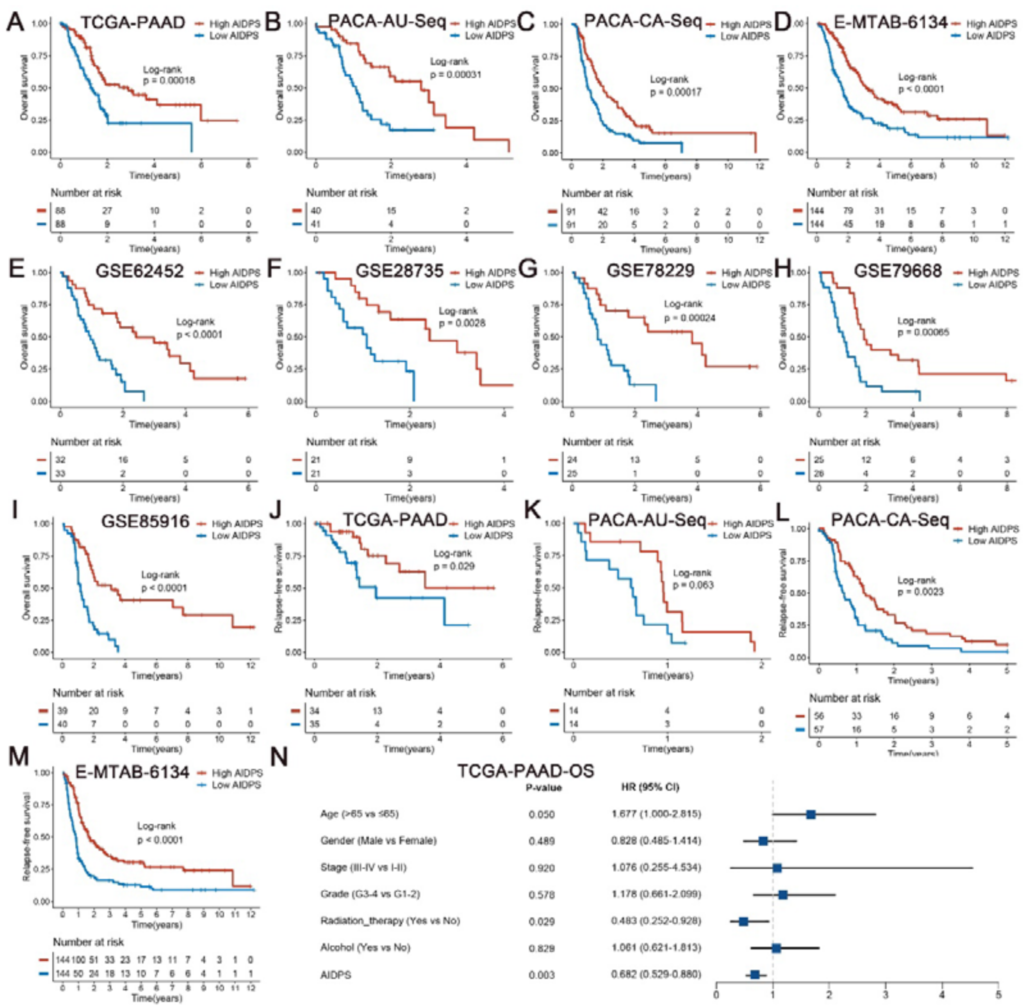


Figure 3-figure supplement 1. Survival analysis of AIDPS in remaining nine validation cohorts. (**A-I**) Kaplan-Meier survival analysis for OS between the high and low AIDPS groups in the TCGA-PAAD (**A**), PACA-AU-Seq (**B**), PACA-CA-Seq (**C**), E-MTAB-6134 (**D**), GSE62452 (**E**), GSE28735 (**F**), GSE78229 (**G**), GSE79668 (**H**), GSE85916 (**I**). (**J-M**) Kaplan-Meier survival analysis for RFS between the high and low AIDPS groups in the TCGA-PAAD (**J**), PACA-AU-Seq (**K**), PACA-CA-Seq (**L**), E-MTAB-6134 (**M**). (**N**) Multivariate Cox regression analysis of OS in the TCGA-PAAD. OS, overall survival; RFS, relapse-free survival.

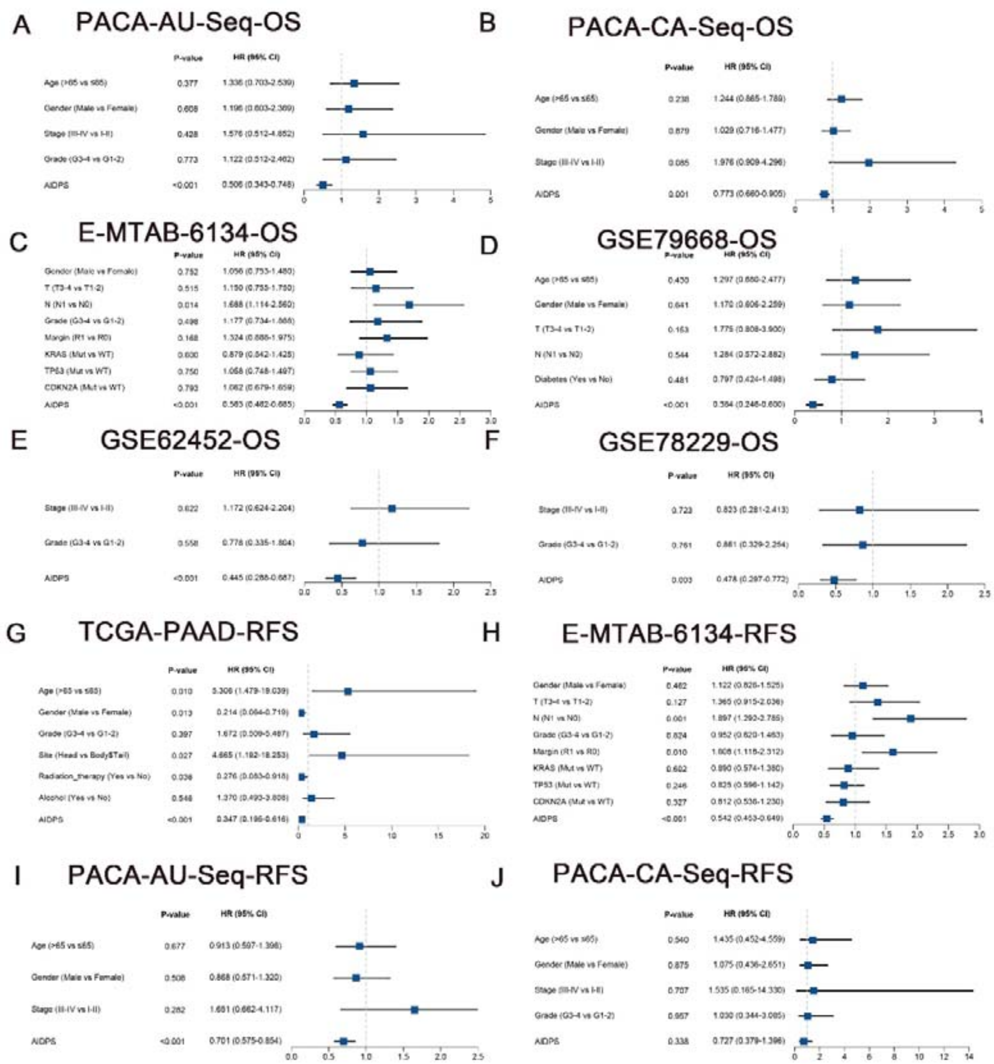


Figure 3-figure supplement 2. Survival analysis of AIDPS in remaining nine validation cohorts. (**A-F**) Multivariate Cox regression analysis of OS in the PACA-AU-Seq (**A**), PACA-CA-Seq (**B**), E-MTAB-6134 (**C**), GSE79668 (**D**), GSE62452 (**E**), GSE78229 (**F**). (**G-J**) Multivariate Cox regression analysis of RFS in the TCGA-PAAD (**G**), E-MTAB-6134 (**H**), PACA-AU-Seq (**I**), PACA-CA-Seq (**J**). OS, overall survival; RFS, relapse-free survival.

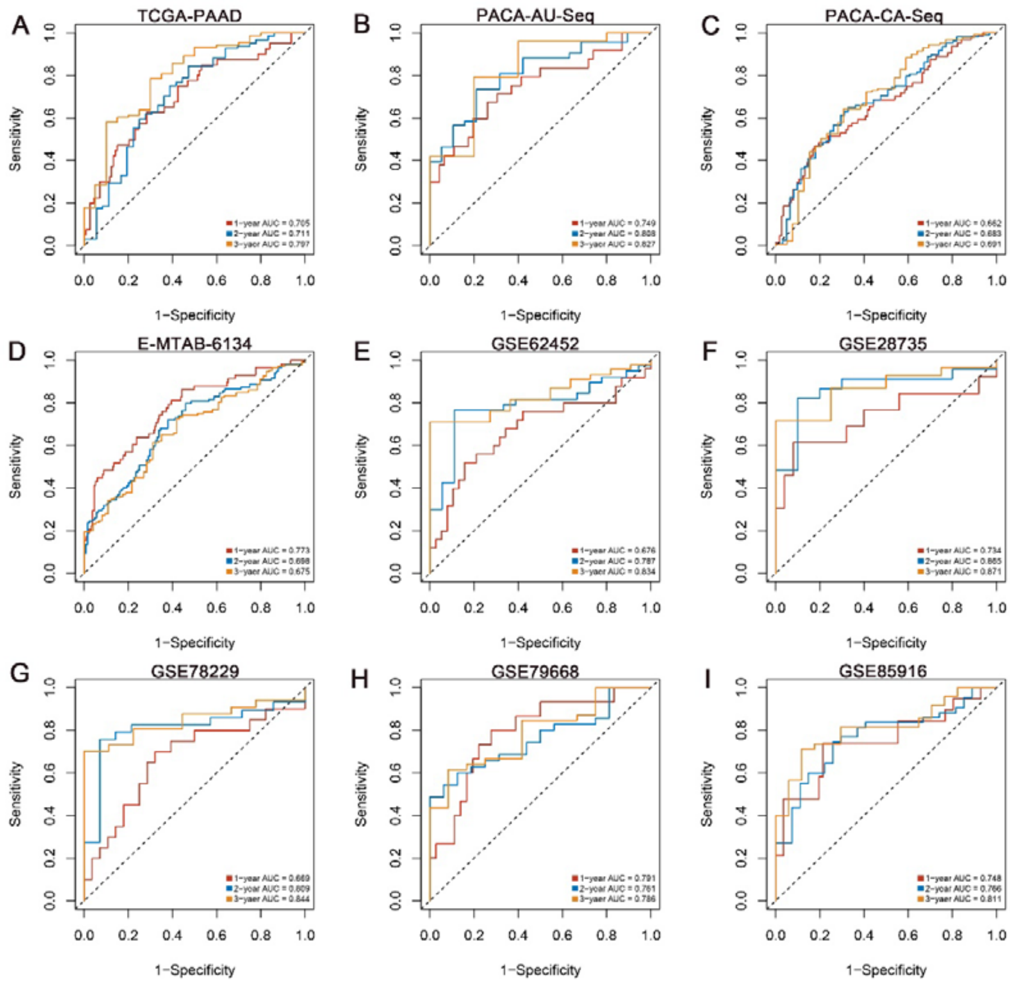


Figure 3-figure supplement 3. Predictive performance evaluation of AIDPS. Time-dependent ROC analysis for predicting 1-, 2-, and 3-years OS in the TCGA-PAAD (**A**), PACA-AU-Seq (**B**), PACA-CA-Seq (**C**), E-MTAB-6134 (**D**), GSE62452 (**E**), GSE28735 (**F**), GSE78229 (**G**), GSE79668 (**H**), GSE85916 (**I**). ROC, receiver operator characteristic; OS, overall survival.

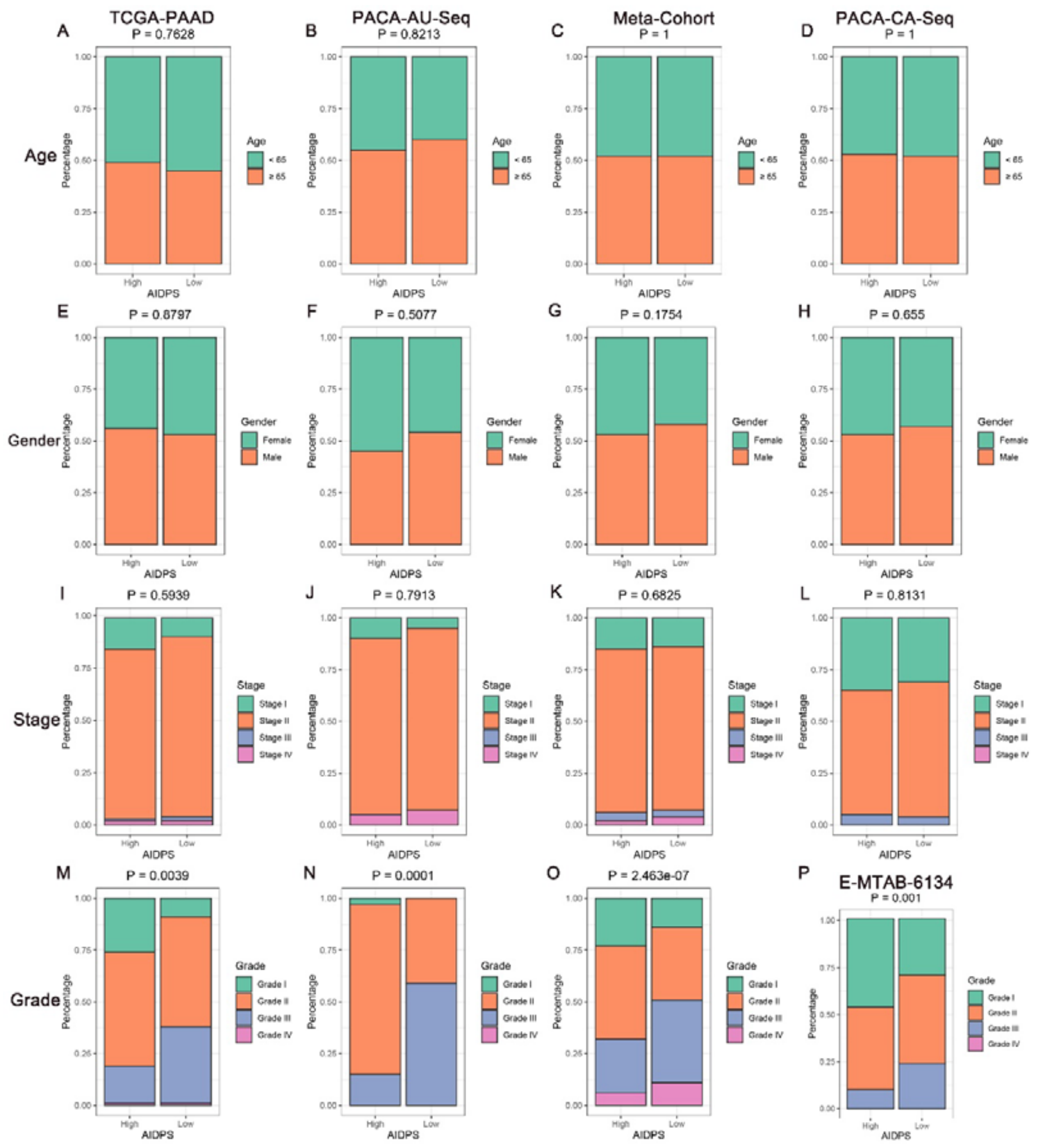


Figure 5-figure supplement 1. The clinical characteristics of the high and low AIDPS groups. (**A, E, I and M**) Composition percentage of the two groups in clinical characteristics such as age (**A**), gender (**E**), stage (**I**), grade (**M**) in the TCGA-PAAD cohort. (**B, F, J and N**) Composition percentage of the two groups in clinical characteristics such as age (**B**), gender (**F**), stage (**J**), grade (**N**) in the PACA-AU-Seq cohort. (**C, G, K and O**) Composition percentage of the two subtypes in clinical characteristics such as age (**C**), gender (**G**), stage (**K**), grade (**O**) in the Meta-Cohort cohort. (**D, H and L**) Composition percentage of the two groups in clinical characteristics such as age (**D**), gender (**H**), stage (**L**) in the PACA-CA-Seq cohort. (**P**) Composition percentage of the two groups on grade in the E-MTAB-6134 cohort.