## **Supplementary Materials**

## **METHODS**

## **Nanoparticle tracking analysis (NTA)**

We evaluated the quality of extracellular vesicles isolated from 3 mL of urine using the Total Exosome Isolation Reagent (from urine) via the NTA system. The number and size of extracellular vesicles were analyzed using a nanoparticle tracking analysis (NTA) system (Malvern Panalytical, NanoSight NS300) according to the manufacturer's protocols. After appropriate dilutions in 1X PBS, video data were collected 5 times for a 60 second time period for each video. Camera level and detection threshold were set to 10 and 3, respectively. Data analysis was performed automatically by NanoSight NTA 3.4 software.

## **RT-qPCR**

Reverse transcription of miRNAs was conducted using the TaqMan Advanced miRNA cDNA Synthesis Kit (Applied Biosystems) according to the manufacturer's instructions. For each miRNA and sample, 5·0 μL of miRNA-enriched eluate was utilized in the initial reverse transcription step. Complementary DNA was pre-amplified as recommended by the kit. Real-time PCR was carried out using the FAST Advanced PCR mix (Applied Biosystems) on the QuantStudio 5 System (Applied Biosystems). Signal detection and analysis were performed with QuantStudio software (Applied Biosystems), and the threshold cycle (Ct) values were determined from the fixed threshold values for each probe as specified by the software. Ct values were not assigned to the measurements with poor amplification according to the amplification status provided by the software, and were removed from the analyses. Ct values of technical replicates ( $n=3$ ) were averaged after excluding outliers (defined as values deviating  $\pm 1$  or more from the median). miRNA expression levels were quantified using delta Ct values by subtracting from the average values of two reference miRNAs (hsa-miR-374a-5p and hsa-miR-30d-5p), and the log2 fold change values were calculated as the difference of the average values of delta Ct in the non-cancer group from that in the pancreatic cancer group. Student's t test was performed for the comparisons of the pancreatic cancer and non-cancer groups.

#### **Performance of prediction model with reduced feature sets**

To explore the feasibility of modeling with a smaller feature set, we performed a feature reduction approach with the recursive feature elimination (RFE) analysis with the linear support vector classifier (SVC) model as estimator. RFE employs a backward feature elimination approach. It initiates with the full set of features and iteratively constructs models, each time excluding the least important feature based on the model's performance metrics. This process continues until the desired number of features is reached, effectively ranking features by their predictive power and identifying the optimal subset for the target variable. We obtained performance curves for both the train and test sets by evaluating the model's area under the curve (AUC) of the receiver operating characteristic (ROC) curve against reduced feature sizes. In the train set, we conducted five-fold cross-validation repeated 100 times to obtain out-of-fold performance metrics. For the test set, we fitted the model using the entire train set and assessed its performance, allowing us to evaluate the generalization capabilities of reduced feature sets. To determine the minimum number of features while maintaining model performance, we conducted Student's t-tests comparing each point of the train set's performance curve with the full feature size of 183. Additionally, we performed Student's t-tests against reference AUC values of 0.90 and 0.95. All tests were upper one-sided, with  $p < 0.05$  considered statistically significant.

#### **RESULTS and DISCUSSIONS**

#### **Nanoparticle tracking analysis (NTA)**

The particle size (mode) ranged from 50 nm to 150 nm, corresponding to exosomes (Supplementary Fig. 1), and the particle concentration was approximately  $1-3\times10^{10}$  (Supplementary Fig. 1). These results indicate that our precipitation method for extraction of extracellular vesicles provides comparable particle size ranges with superior yields (10<sup>10</sup> particles / mL from only 3 mL of urine) compared to previous reports using ultracentrifugation, polyethylene glycol (PEG), and size exclusion chromatography (SEC) for exosome isolation.<sup>1-3</sup>

#### **Learning curve analysis of the prediction model**

To assess the potential risk of overfitting and evaluate our model's generalization capability, we conducted a learning curve analysis for our prediction model. As shown in Supplementary Fig. 5, the performance of the out-of-fold prediction on the validation set improved with an increase in the training set size, while the performance of the in-fold prediction on the training set remained almost constant. When the sample size was small, we observed a significant gap between the validation and training performances, indicating that the model was initially overfitted to the training set, and demonstrated poor generalizability to the out-of-fold data. As the training set size grew, we noted a gradual improvement in performance, suggesting that the model was acquiring greater generalizability with the addition of more training samples. The learning curve also indicated a gradual slowdown in performance improvement, likely approaching a saturation point with an AUC greater than  $0.95$ . Although it remains possible that the curve could show further growth with an even larger training set, this trend implies that the model effectively fitted the training set and achieved strong generalization, resulting in high performance in the out-of-fold prediction for the validation set. Given that the training set size in the current study was  $N = 315$ , which exceeds the saturation point, we conclude that our model was well-fitted to the train set and demonstrated robust generalization capabilities. These findings support the robustness of our prediction model and mitigate concerns about its generalizability.

#### **Evaluation of urinary bilirubin impact on miRNA assay performance**

We investigated urinary bilirubin using a simple test strip method (Aution Sticks 10EA, Arkray, Japan). Of the 99 and 54 pancreatic cancer cases in the train and test sets, 5 (Early stage: 2, Late stage: 2, Unknown: 1) and 2 (Late stage: 2) cases were bilirubin positive, respectively. All of the seven bilirubin-positive cases were positive in the urinary miRNA assay. In summary, a low incidence of bilirubin-positive cases was observed in the current pancreatic cancer cases, and those cases were successfully classified as pancreatic cancer positive in our urinary miRNA assay.

#### **Verification of DEMs train vs test**

To confirm the consistency of the results in the train and test sets, we performed differential expression analysis independently for the train and test sets. As shown in Supplementary Fig. 6, the log2 fold change values were highly correlated with each other (Pearson's  $r = 0.74$ ,  $p < 0.0001$ ). Of the 26 up-regulated and 19 down-regulated DEMs in the train set, 19 and 5 were consistently differentially expressed in the test set, respectively (Supplementary Table 1).

#### **Verification of DEMs by RT-qPCR**

We selected 30 pancreatic cancer and 30 non-cancer cases from the training set and targeted 15 miRNA probes for qPCR analysis using the TaqMan Assay. These probes were chosen based on prior confirmation of their detection among the 45 differentially expressed miRNAs (DEMs) identified through differential expression analysis (DEA) with small RNA-seq. The normalized Ct values showed a negative correlation (Pearson's  $r = -0.68$ ) with the normalized count from NGS, indicating that the overall miRNA expression levels were consistent (Supplementary Fig 7 (a)). DEA for the normalized Ct values revealed that 3 of the 8 up-regulated miRNAs identified in small RNA-seq exhibited a log2 fold change greater than  $0.5$  in qPCR (Supplementary Fig 7 (b) and Supplementary Table 5), indicating consistency between the small RNA sequencing and qPCR results.

#### **Performance of prediction model with reduced feature sets**

To explore the feasibility of modeling with a smaller feature set, we implemented a recursive feature elimination analysis. As shown in Supplementary Fig. 8 (a), reducing the feature set size from the full  $M = 183$  did not significantly affect model performance until it reached  $M = 96$  (Supplementary Fig. 8 (b)). Further reduction led to a gradual decline in performance, but the AUC was not significantly lower than 0.95 until  $M = 43$  (Supplementary Fig. 8 (c)). Below this threshold, the decline in performance accelerated, with AUC dropping below  $0.90$  at  $M = 10$ . Although the model showed reasonable performance with only 10 features, at least 43 features were necessary to maintain a high AUC of approximately 0·95. While the model performed comparably with 95 features in the training set, the full feature set demonstrated the most robust performance in the test set. These findings suggest that our modeling approach offers a robust prediction with the small RNA-seq assay.

## **Prediction performance in samples with storage conditions**.

To evaluate the potential confounding effect of urine storage conditions on our urinary miRNA assay, we compared specificity values for non-cancer samples stored at –80°C (Hokuto Hospital) and –20°C (Omiya City Clinic), as described in the Methods section. As shown in Supplementary Fig. 9, there were no statistically significant differences in specificity between the storage conditions of –80°C and –20°C, both in the training set (chi-squared test,  $p = 0.566$ ) and the test set (chi-squared test,  $p = 1.0$ ).

## **References**

- 1 Street JM, Koritzinsky EH, Glispie DM, Yuen PST. Urine Exosome Isolation and Characterization. In: Gautier J-C (ed). *Drug Safety Evaluation*. Springer New York: New York, NY, 2017, pp 413–423.
- 2 Gheinani AH, Vögeli M, Baumgartner U *et al.* Improved isolation strategies to increase the yield and purity of human urinary exosomes for biomarker discovery. *Sci Rep* 2018; **8**: 3945.
- 3 Cho S, Yang HC, Rhee WJ. Development and comparative analysis of human urine exosome isolation strategies. *Process Biochem* 2020; **88**: 197–203.

#### **Supplementary Tables**

**Supplementary Table 1. Differentially expressed miRNAs (DEMs) in urine.** (a) Upregulated miRNAs (b) Downregulated miRNAs.

## **(a) Upregulated miRNAs**





**Train Test**

## **(b) Downregulated miRNAs**



**Supplementary Table 2. Overrepresented pathways of DEMs in urine.** (a) 36 overrepresented pathways in KEGG (*p*-adjusted <0·05) (b) Top 100 overrepresented GO terms according to −log10 (*p*-adjusted) (*p*-adjusted <0·05).















hsa-let-7d-5p; hsa-miR-582-5p; hsa-miR-30c-1-3p; hsa-





5p; hsa-miR-1306-5p; hsa-miR-574-5p; hsa-miR-574-3p; hsa-miR-30c-1-3p; hsa-miR-296-3p; hsa-miR-30b-3p

















let-7d-5p; hsa-miR-744-5p; hsa-miR-574-5p; hsa-miR-

hsa-miR-4728-3p; hsa-miR-30c-1-3p; hsa-miR-30b-3p





miR-582-5p; hsa-miR-30c-1-3p; hsa-miR-30b-3p







**Supplementary Table 3. Coefficients of the multiple regression analysis with the generalized linear model with logit link function.**



BMI, Body Mass Index.

**Supplementary Table 4. Lists of overlap between urinary DEMs and organoid DEMs.** (a) Upregulated miRNAs in urine (b) Downregulated miRNAs in urine.

**(a) Upregulated miRNAs in urine**









# **Supplementary Table 5. Verification of 15 DEMs in qPCR.**



**Supplementary Figure 1. Nanoparticle tracking assay for urinary-EV.** (a) Particle size distribution of a representative EV sample (#2). The shaded region indicates the standard error of five repeated measurements. (b) Particle size estimated from the mode of the distribution, plotted for three biological replicates and two technical replicates. Error bars represent the standard error of five repeated measurements. (c) EV concentration estimated from the area of the distribution, plotted similarly to (b). EV, Extracellular vesicle.



**Supplementary Figure 2. Prediction score across demographic characteristic factors.** Prediction scores were plotted against values of (a) Sex, (b) Age, (c) BMI, (d) Alcohol, (e) Smoking history, and (f) Diagnosis modality. BMI, Body Mass Index.



**Supplementary Figure 3. Performance of the model developed for the age-match dataset.** (a) Box plot showing age was matched between groups in the age-match dataset. (b) ROC curve for SVC model with linear kernel developed in the age-matched. The mean value of five-folds is shown as the solid blue line, with standard deviations shown as the blue-shaded region. (c) Prediction scores obtained via cross-validation were plotted against age.





\*Identical patient

**Supplementary Figure 4. Organoid analysis.** (a) Schematic overview of the organoid analysis. (b) Donor characteristics of the organoid analysis. \*Identical patient.



**Supplementary Figure 5. Learning curve analysis of the prediction model.** The AUC values were plotted against varying train set sizes. Each point represents the mean AUC from 100 repetitions of five-fold cross-validation, with solid lines indicating mean values and regions representing standard deviations. Blue and Yellow lines denote the performance on the training and validation sets, respectively.



**Supplementary Figure 6. Comparison of DEMs between training and test sets.** Differential expression analyses were conducted independently for the training and test sets. Log2 fold change values were plotted on a scatter plot, with the horizontal axis representing the train set and the vertical axis representing the test set. Each point corresponds to a single miRNA. The solid line shows the regression line, with the shaded region representing the 95% confidence interval. Pearson's correlation coefficient (r) and its associated p-value were displayed in the upper left corner of the plot.



**Supplementary Figure 7. Verification of DEMs by the qPCR assay.** Pancreatic cancer (N=30) and non-cancer (N=30) cases in the train set were subjected to qPCR analysis, and the results were compared with the small RNA-seq results. (a) Normalized miRNA expression levels were compared for 15 target miRNAs, with the horizontal and vertical axes indicating the delta Ct values and the log2 transformed base mean values from DESeq2. The solid line shows the regression line, with the shaded region representing the 95% confidence interval. Pearson's correlation coefficient (r) was displayed in the upper right corner of the plot. (b) Log2 fold change values of 15 target miRNAs were plotted for both the small RNA-seq and qPCR. Error bars indicate the standard errors.



**Supplementary Figure 8. Feature reduction analysis of the prediction model.** (a) The AUC values were plotted against reducing the feature set size by employing the recursive feature elimination analysis. The blue line indicates the AUC values evaluated with out-of-fold prediction through 100 repetitions of five-fold cross-validation in the train set, with solid lines indicating mean values and error bars representing standard deviations. The yellow line indicates the AUC value of the out-of-fold prediction in the test set. (b) ROC curve of the model with the reduced feature size of 96. (c) ROC curve of the model with the reduced feature size of 43.



**Supplementary Figure 9. Performance comparison of storage conditions.** Specificity values were plotted against the storage temperature for the non-cancer cases in the train and test sets, respectively. Error bars indicate 95% confidence intervals.