














Prediction of tissue-of-origin of early stage cancers using serum miRNomes

Juntaro Matsuzaki, MD, PhD,^{1,2} Ken Kato, MD, PhD,³ Kenta Oono, PhD,⁴ Naoto Tsuchiya, PhD,⁵ Kazuki Sudo , MD, PhD,⁶ Akihiko Shimomura , MD, PhD,⁶ Kenji Tamura, MD, PhD,⁶ Sho Shiino, MD, PhD,⁷ Takayuki Kinoshita, MD, PhD,⁸ Hiroyuki Daiko , MD, PhD,⁹ Takeyuki Wada, MD, PhD,¹⁰ Hitoshi Katai, MD, PhD,¹⁰ Hiroki Ochiai, MD, PhD,¹¹ Yukihide Kanemitsu , MD,¹¹ Hiroyuki Takamaru, MD, PhD,¹² Seiichiro Abe , MD, PhD,¹² Yutaka Saito, MD, PhD,¹² Narikazu Boku , MD, PhD,³ Shunsuke Kondo , MD, PhD,¹³ Hideki Ueno, MD, PhD,¹³ Takuji Okusaka, MD, PhD,¹³ Kazuaki Shimada, MD, PhD,¹⁴ Yuichiro Ohe , MD, PhD,¹⁵ Keisuke Asakura, MD, PhD,¹⁶ Yukihiko Yoshida , MD, PhD,¹⁶ Shun-Ichi Watanabe, MD, PhD,¹⁶ Naofumi Asano, MD, PhD,¹⁷ Akira Kawai, MD, PhD,¹⁷ Makoto Ohno , MD, PhD,¹⁸ Yoshitaka Narita , MD, PhD,¹⁸ Mitsuya Ishikawa, MD, PhD,¹⁹ Tomoyasu Kato , MD, PhD,¹⁹ Hiroyuki Fujimoto, MD, PhD,²⁰ Shumpei Niida, PhD,²¹ Hiromi Sakamoto, PhD,²² Satoko Takizawa, PhD,^{1,23} Takuya Akiba, PhD,⁴ Daisuke Okanohara, PhD,⁴ Kouya Shiraiishi, PhD,²⁴ Takashi Kohno , PhD,²⁴ Fumitaka Takeshita, PhD,²⁵ Hitoshi Nakagama, MD, PhD,²⁶ Nobuyuki Ota, PhD,²⁷ Takahiro Ochiya, PhD,^{1,28,*} and the Project Team for Development and Diagnostic Technology for Detection of miRNA in Body Fluids[†]

¹Division of Molecular and Cellular Medicine, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

²Division of Pharmacotherapeutics, Keio University Faculty of Pharmacy, Minato-ku, Tokyo, Japan

³Department of Head and Neck, Esophageal Medical Oncology and Department of Gastrointestinal Medical Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

⁴Preferred Networks, Inc, Chiyoda-ku, Tokyo, Japan

⁵Laboratory of Molecular Carcinogenesis, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

⁶Department of Breast and Medical Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

⁷Department of Breast Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

⁸Department of Breast Surgery, National Hospital Organization Tokyo Medical Center, Meguro-ku, Tokyo, Japan

⁹Department of Esophageal Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁰Department of Gastric Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹¹Department of Colorectal Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹²Endoscopy Division, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹³Department of Hepatobiliary and Pancreatic Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁴Department of Hepatobiliary and Pancreatic Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁵Department of Thoracic Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁶Department of Thoracic Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁷Department of Musculoskeletal Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁸Department of Neurosurgery and Neuro-Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

¹⁹Department of Gynecology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

²⁰Department of Urology, National Cancer Center Hospital, Chuo-ku, Tokyo, Japan

²¹Research Institute, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan

²²Department of Biobank and Tissue Resources, Fundamental Innovative Oncology Core, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

²³Toray Industries, Inc, Kamakura, Kanagawa, Japan

²⁴Division of Genome Biology, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

²⁵Department of Translational Oncology, Fundamental Innovative Oncology Core, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

²⁶National Cancer Center, Chuo-ku, Tokyo, Japan

²⁷Preferred Medicine, Inc, Burlingame, CA, USA

²⁸Department of Molecular and Cellular Medicine, Tokyo Medical University, Shinjuku-ku, Tokyo, Japan

[†]The participating institutes and responsible members in the Project Team for Development and Diagnostic Technology for Detection of miRNA in Body Fluids can be found in the [Supplementary text](#) (available online).

*Correspondence to: Takahiro Ochiya, PhD, Department of Molecular and Cellular Medicine, Tokyo Medical University, 6-7-1 Nishishinjuku, Shinjuku-ku, Tokyo 160-0023, Japan (e-mail: tochiya@tokyo-med.ac.jp).

Abstract

Background: Noninvasive detection of early stage cancers with accurate prediction of tumor tissue-of-origin could improve patient prognosis. Because miRNA profiles differ between organs, circulating miRNomes represent a promising method for early detection of cancers, but this has not been shown conclusively.

Methods: A serum miRNA profile (miRNomes)-based classifier was evaluated for its ability to discriminate cancer types using advanced machine learning. The training set comprised 7931 serum samples from patients with 13 types of solid cancers and 5013

Received: August 05, 2022. Revised: October 02, 2022. Accepted: October 17, 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

noncancer samples. The validation set consisted of 1990 cancer and 1256 noncancer samples. The contribution of each miRNA to the cancer-type classification was evaluated, and those with a high contribution were identified.

Results: Cancer type was predicted with an accuracy of 0.88 (95% confidence interval [CI] = 0.87 to 0.90) in all stages and an accuracy of 0.90 (95% CI = 0.88 to 0.91) in resectable stages (stages 0-II). The F1 score for the discrimination of the 13 cancer types was 0.93. Optimal classification performance was achieved with at least 100 miRNAs that contributed the strongest to accurate prediction of cancer type. Assessment of tissue expression patterns of these miRNAs suggested that miRNAs secreted from the tumor environment could be used to establish cancer type-specific serum miRNomes.

Conclusions: This study demonstrates that large-scale serum miRNomics in combination with machine learning could lead to the development of a blood-based cancer classification system. Further investigations of the regulating mechanisms of the miRNAs that contributed strongly to accurate prediction of cancer type could pave the way for the clinical use of circulating miRNA diagnostics.

Improvements in nucleic acid sequencing technologies has led to an exponential increase in the demand for novel cancer diagnostics based on body fluids (1). Cancer detection systems with high specificity for the cancer site could help prevent unnecessary whole-body examinations. Although circulating tumor DNA (ctDNA) analysis has the potential to achieve this goal (2-4), the analysis of genomic abnormalities in ctDNA has a relatively low accuracy (0.61) for predicting cancer type (3), possibly because regions containing driver genes are commonly mutated in different cancer types. Analysis of ctDNA methylation status has a better performance for predicting cancer type than genomic abnormalities, with an accuracy of 0.93, although the sensitivity in early stage cancers is low (sensitivities = 0.17 in stage I, 0.40 in stage II) (5-7).

Extracellular microRNAs (ex-miRNAs) have been intensively studied worldwide as disease indicators. Ex-miRNAs are functional molecules that mediate cell-to-cell communication through packaging in extracellular vesicles (EVs); therefore, analyzing ex-miRNA profiles could provide helpful information on health and disease (8). The ability of serum miRNA profiling to discriminate between cancer and noncancer with high accuracy has been demonstrated extensively (9-11). However, although tissue miRNA expression profiles differ between cancer types (12-15), the idea that circulating miRNA profiles could serve as an accurate diagnostic tool for determining cancer type has not yet received the attention it deserves. To test this possibility, we launched a national project in Japan called Development and Diagnostic Technology for Detection of miRNA in Body Fluids (DDDmir-DB) and developed the world's largest cancer patient ex-miRNA database. In this study, we provide proof of concept that automatic extraction of key features of serum miRNA profiles (miRNomes) can discriminate between cancer types with high accuracy.

Methods

Development of machine learning models for cancer type prediction

We developed an ensemble classifier, called the Hierarchical Ensemble Algorithm with Deep learning (HEAD) model, which combines 7 different learners. A 2-stage stacking technique was used to build the classifiers (see Figure 1, A) (16). The first stage consists of learners placed in parallel. A combination of the pre-processed dataset and the output of an unsupervised feature extractor was fed to the learners, each of which yielded prediction results. Random forest, logistic regression, extra tree classifier (17), support vector classifier, k-NN (nearest neighbor) classifier, gradient boosting decision tree (GBDT) (18), and multi-layer perceptron (MLP) (19) were used as the first stage weak

classifiers. In addition to learners, k-NN was used as an unsupervised feature extractor in the first stage. The diversity of constituent models is a key element of the success of ensemble models (20). The outputs of learners in the first stage were fed to a learner (GBDT) in the second stage.

Ethics statements

The study was approved by the National Cancer Center (NCC) Hospital institutional review board (2015-376, 2016-249), the Research Ethics Committee of Medical Corporation Shintokai of the Yokohama Minoru Clinic (6019-18-3772), and the Ethics and Conflict of Interest Committee of the National Center for Geriatrics and Gerontology (754). Written informed consent was obtained from each participant.

Results

Testing the diagnostic potential

We used HEAD to analyze a publicly available serum miRNA dataset (GSE59856) (21) consisting of 6 different digestive cancers, nontumor (NT) control samples, and benign disease in the biliary tract (BT) or the pancreas (Supplementary Figure 1, A, available online). Principal component analysis showed that sample categories could not be clearly segregated, suggesting that they share similar miRNA profiles and that it would be difficult to discriminate between them by conventional means (Figure 1, B). We randomly divided the data into a training set and validation set at a ratio of 4:1 and trained the HEAD and GBDT models using the training set (Supplementary Table 1, available online). In the validation set, HEAD achieved sensitivities of 0.60-1.00 for cancer discrimination, which was superior to that of GBDT alone (Figure 1, C). In line with previous results (9-11), NT was almost perfectly discriminated in cancer samples irrespective of the machine learning model used.

Concept verification

The performance of machine learning can be improved by training on large datasets. To evaluate the value of this approach, we analyzed the serum miRNomes of 16 190 serum samples preserved in the NCC Biobank using a microarray platform (3D-Gene v21, Toray Industries, Tokyo, Japan) and standard operating procedure (Supplementary Figures 1, B, and 2, A, available online). For data normalization, we used 3 internal control miRNAs (miR-149-3p, miR-2861, and miR-4463), which are suitable for 3D-Gene-derived serum miRNA datasets (10,22,23) (Supplementary Figure 2, B, available online). The following analyses were conducted after data normalization.

After feeding the whole training set (cancer, $n = 7931$; noncancer, $n = 5013$) into HEAD, we accurately distinguished 13 types

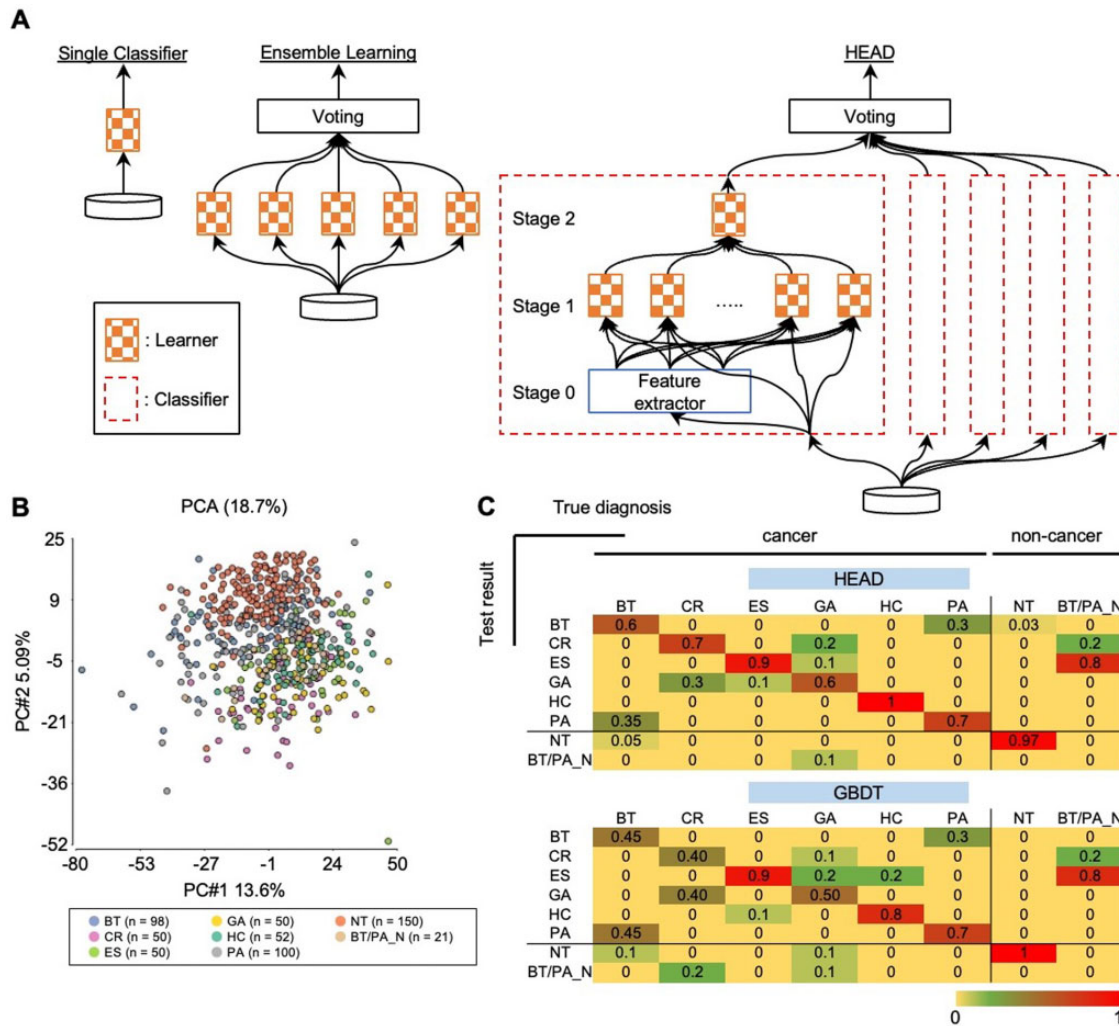


Figure 1. Cancer types can be classified by serum miRNA profiles using machine learning. **A**) Schematic view of the HEAD machine learning system. The system consists of multiple classifiers with the same architecture. Red narrow boxes with broken lines in the middle and right represent copies of the classifier on the left. Each classifier consists of 3 stages: unsupervised feature extraction in the first stage, various learners in the second stage, and a single classifier in the third stage. The output of previous stages is fed into the next stage. Learners in the figures can be of different types (eg, random forest, logistic regression, extra tree classifier, support vector classifier, k-NN, GBDT, and MLP). The results of prediction classifiers are aggregated using the voting method. For comparison, schematic views of a single classifier and an ensemble learning model are also shown. **B**) PCA plot showing that miRNA profiles in the GSE59856 dataset do not exhibit clear separation among cancer types. **C**) Two machine learning-based prediction models (HEAD and GBDT), developed using the training set, were applied to the validation set. The diagnostic sensitivities of 6 cancer types with HEAD and GBDT are shown. BT = biliary tract cancer; CR = colorectal cancer; ES = esophageal squamous cell carcinoma; GA = gastric cancer; GBDT = gradient boosting decision tree; HC = hepatocellular carcinoma; HEAD = Hierarchical Ensemble Algorithm with Deep learning; k-NN = k-nearest neighbors; miRNA = microRNA; MLP = multilayer perceptron; N = benign disease; NT = nontumor; PA = pancreatic cancer; PCA = principal component analysis.

of solid cancer samples in the validation set (cancer, $n=1990$; noncancer, $n=1256$) (Table 1 and Figure 2, A). In the validation set, the overall accuracy for HEAD was 0.91 (95% confidence interval [CI] = 0.91 to 0.92), which was statistically significantly higher than that of the ensemble 4-layer MLP (accuracy = 0.88, 95% CI = 0.88 to 0.89) and the other models (Supplementary Figure 3, A, Supplementary Table 2, available online). The accuracy of discrimination for the 13 cancer types was 0.88 (95% CI = 0.87 to 0.90) for HEAD, which was also the best among the evaluated models. The overall F1 score for HEAD was 0.89, whereas the F1 score for the discrimination of the 13 cancer types was 0.93. The superiority of the HEAD model was confirmed by five-fold cross-validation (Supplementary Figure 3, B, available online).

Perfect discrimination between NT and the other samples in DDDmir-DB was observed consistently in the GSE59856 dataset.

Considering that differences in the preservation conditions between NCC Biobank samples and NT samples collected from other institutes could affect this perfect discrimination (Supplementary Figure 1, available online), we mainly focused on discriminating between NCC Biobank samples after excluding NT samples from the following analysis. The HEAD model can output the probability scores for each cancer type in all samples; the cancer type with the maximum probability score is treated as the final output. When we conducted the receiver operating characteristic curve analysis using the probability score after the exclusion of NT control samples, the area under the curve value was greater than 0.95 for all cancer types (Figure 2, B; Supplementary Figure 4, available online), and sensitivities at a threshold of 0.99 specificity were greater than 0.80 for almost all cancer types except for biliary tract cancer (BT) and hepatocellular carcinoma (HC) (Table 2). The high

Table 1. Age and sex distribution of the patients from whom the analyzed samples in the DDDmir-DB were obtained

Cancer type	Training set				Validation set			
	No.	Men, %	Women, %	Mean age (SD), y	No.	Men, %	Women, %	Mean age (SD), y
Cancer	7931	60.4	39.6	63.5 (12.0)	1990	61.2	38.8	64.3 (12.0)
Breast cancer	540	0.0	100.0	55.9 (12.2)	135	0.0	100.0	53.7 (12.3)
Bladder cancer	319	70.8	29.2	67.4 (10.7)	80	78.8	21.2	68.3 (11.5)
Biliary tract cancer	321	62.9	37.1	66.0 (9.2)	81	50.6	49.4	66.5 (9.8)
Colorectal cancer	1276	57.5	42.5	63.7 (11.2)	320	57.8	42.2	64.9 (11.6)
Esophageal squamous cell carcinoma	452	82.5	17.5	66.5 (8.6)	114	86.0	14.0	66.7 (8.1)
Gastric cancer	1134	69.8	30.2	65.0 (10.8)	284	74.3	25.7	65.9 (10.3)
Intraparenchymal brain tumor ^a	192	56.8	43.2	55.6 (16.9)	49	57.1	42.9	55.3 (16.9)
Hepatocellular cancer	278	77.0	23.0	67.2 (9.2)	70	80.0	20.0	68.7 (9.1)
Lung cancer	1359	57.6	42.4	65.3 (10.0)	340	56.8	43.2	65.3 (10.1)
Ovarian cancer	320	0.0	100.0	55.6 (12.4)	80	0.0	100.0	59.2 (12.5)
Pancreatic cancer	680	56.5	43.5	64.1 (9.9)	171	60.2	39.8	66.5 (9.9)
Prostate cancer	821	100.0	0.0	67.4 (7.5)	206	100.0	0.0	68.1 (8.0)
Sarcoma	239	62.8	37.2	46.4 (23.6)	60	56.7	43.3	48.8 (20.4)
Noncancer	5013	45.7	54.3	66.1 (16.2)	1256	44.8	55.2	65.3 (16.0)
Nontumor	4514	43.9	56.1	67.5 (15.4)	1129	43.1	56.9	66.6 (15.4)
Benign disease in the breast	24	0.0	100.0	52.8 (12.9)	7	0.0	100.0	47.1 (13.7)
Benign disease in the brain	19	26.3	73.7	60.6 (18.4)	5	43.1	56.9	72.0 (8.6)
Benign disease in the ovary	22	0.0	100.0	56.5 (10.6)	6	0.0	100.0	58.8 (10.3)
Benign disease in the prostate	184	100.0	0.0	65.4 (7.2)	46	100.0	0.0	64.7 (6.5)
Benign disease in the bone and soft tissue	250	49.2	50.8	44.3 (18.8)	63	42.9	57.1	45.4 (17.8)

^a Intraparenchymal brain tumor such as glioma. DDDmir-DB = Development and Diagnostic Technology for Detection of miRNA in Body Fluids.

diagnostic performance was not affected by patient age or sex (Figure 2, C and D).

We integrated the data of serum miRNomes obtained from patients with benign diseases (N) in 5 different organs (breast [BR], brain [GL], ovary [OV], prostate, and bone and soft tissue). Only a few cancer samples were misdiagnosed as noncancer. Although we could not obtain a large enough sample size to predict BR_N, GL_N, and OV_N as noncancer samples (Supplementary Figure 2, available online), disease location was almost perfectly predicted (Figure 2, A). This finding suggests that serum miRNomes are dysregulated in an organ-specific manner. Although additional collection and training of noncancer samples are necessary to improve the diagnostic performance, such excellent predictive performance regarding disease location will be helpful in various clinical settings, such as organ-specific cancer risk stratification and the management of cancers of unknown primary origin.

High discrimination accuracy was achieved regardless of disease stage using the HEAD model (Figure 2, E; Supplementary Figure 3, C, available online), indicating that circulating miRNA diagnosis could be useful even for the detection of early stage cancers. The accuracy of discrimination for the 13 cancer types was 0.90 (95% CI = 0.88 to 0.91), even for tumors in the resectable stages (stage 0, I, and II), whereas it was 0.86 (95% CI = 0.83 to 0.89) in stage III or IV. Most of the cancer types except BT and HC were diagnosed with a true prediction rate of more than 0.75, even when they were in the resectable stage (Figure 2, F). In addition, we compared the probability scores among disease stages in each cancer type (Figure 3). The probability scores were not associated with disease stage in most cancer types. In BT and pancreatic cancer, the probability scores increased with disease progression.

Transfer learning

In recent years, machine learning has been used for clinical diagnosis in various applications, but the integrated analysis of data collected by different protocols remains challenging. We were unable to adapt the HEAD model trained by the DDDmir-DB

alone for the prediction of cancer type in the GSE59856 dataset because of differences in sample preparation methods and microarray platforms. Thus, we used the domain adversarial neural network (DANN), a transfer learning framework, to integrate 2 datasets (24). The DANN is designed to extract the common features of each cancer type regardless of the differences in data sources (Figure 4, A). The data for DDDmir-DB and GSE59856 were fed to the DANN algorithm, and the predictive performance of GSE59856 was evaluated. Integrating 2 datasets statistically significantly improved the true prediction rate of GSE59856 for pancreatic cancer; the sensitivities for the other cancer types also improved compared with the model trained by GSE59856 alone (Figure 4, B). Thus, the accrual of blood miRNome data obtained even under different conditions can improve the accuracy of cancer type prediction.

miRNAs that contribute strongly to cancer classification

We sought to determine which serum miRNAs contribute to diagnostic performance by computing the contribution of each miRNA based on the number of associated nodes in the decision tree (Figure 5, A; Supplementary Table 3, available online). By calculating the true prediction rate for each cancer type with the 10, 30, 50, 100, 300, and 1000 most strongly contributing miRNAs, we found that diagnostic performance was optimal with 100 miRNAs (Figure 5, B). Next, we focused on the 179 miRNAs that contributed the most to cancer classification in the 2 datasets (contribution >0.01, blue and red dots in Figure 5, C). By calculating the average serum levels of the 179 miRNAs for each cancer type, the similarity of serum miRNomes among cancer types was investigated. Principal component analysis of the average serum miRNomes revealed that GL was segregated from the other cancer types (Figure 5, D). Unsupervised clustering analysis suggested that cancer types whose primary sites shared similar developmental processes produced similar serum miRNomes (Figure 5, E). In lung samples, pathological differences (small cell, squamous cell, or adenocarcinoma) and differences in driver

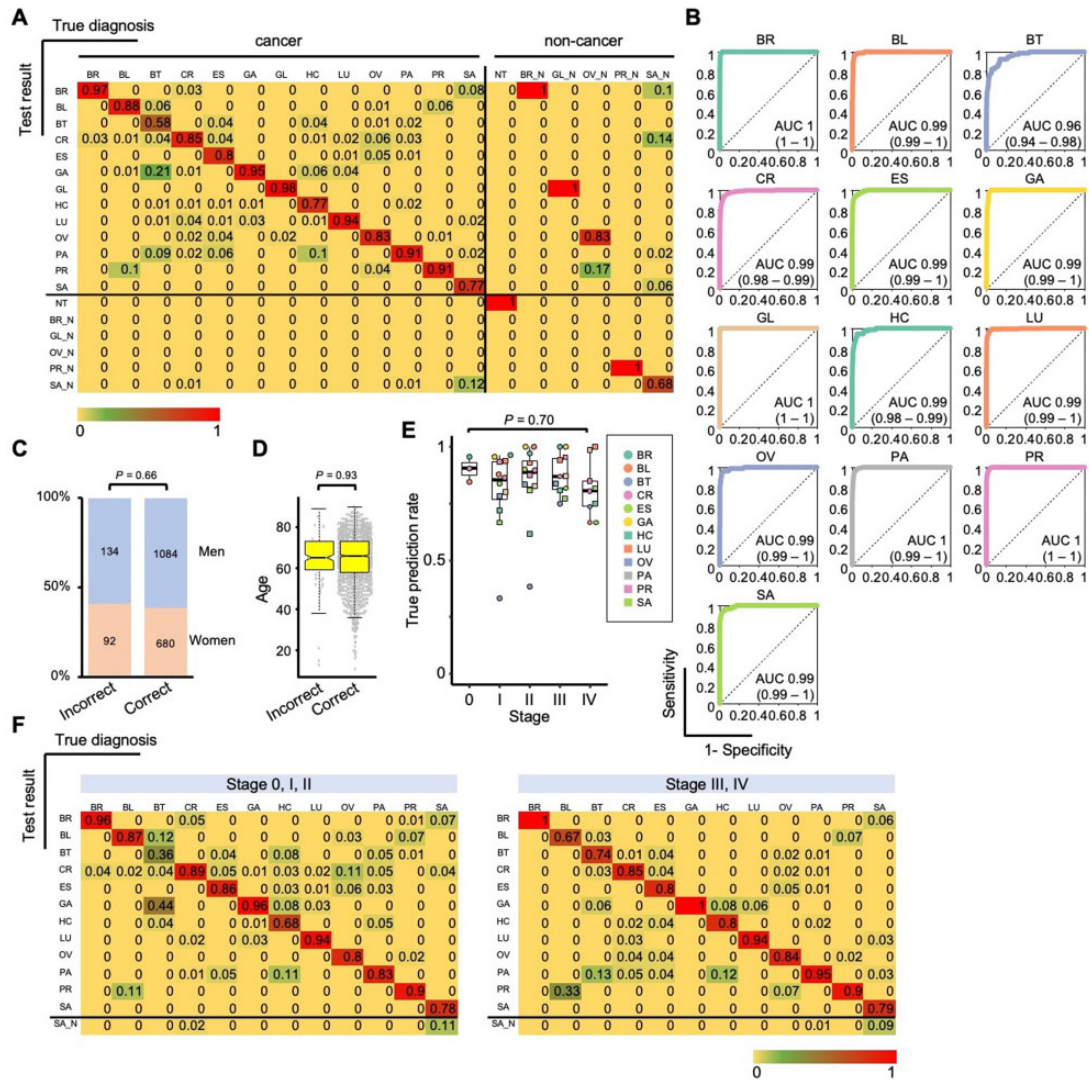


Figure 2. The HEAD model enables accurate discrimination of 13 cancer types in the validation set. **A)** The true prediction rate for each of 13 kinds of solid cancer was greater than 0.8 except for BT, HC, and SA in HEAD. NT samples and PR_N were perfectly discriminated as nontumor. BR_N, GL_N, and OV_N samples were mainly diagnosed as cancer samples in the corresponding organs. **B)** ROC curve analysis of the HEAD model for discrimination of each cancer type. The discrimination performance for each cancer type among all cancer samples and noncancer samples is indicated after the exclusion of NT control samples. The AUC for detecting each cancer type was greater than 0.95. Numbers inside parentheses indicate 95% confidence interval of AUC. **C)** The proportion of each sex did not differ between patients diagnosed correctly and incorrectly by HEAD. P , Fisher exact test. **D)** Age distribution did not differ between correctly and incorrectly diagnosed patients in HEAD. P , student t test. **E)** The diagnostic sensitivities calculated by HEAD were not associated with the disease stage of cancer samples, indicating that serum miRNA-based tests are feasible for early detection of cancer. P , one-way analysis of variance. **F)** The diagnostic performance for earlier stage cancers (stages 0, I, and II) and later stage cancers (stages III and IV). The true prediction rate was greater than 0.75 except for BT and HC even in the earlier stage. AUC = area under the ROC curve; BL = bladder cancer; BR = breast cancer; BT = biliary tract cancer; CR = colorectal cancer; ES = esophageal squamous cell carcinoma; GA = gastric cancer; GL = intraparenchymal brain tumor such as glioma; HC = hepatocellular carcinoma; HEAD = Hierarchical Ensemble Algorithm with Deep learning; LU = lung cancer; miRNA = microRNA; N = benign disease; NT = nontumor; OV = ovarian cancer; PA = pancreatic cancer; PR = prostate cancer; ROC = receiver operating characteristic; SA = sarcoma.

gene mutations (KRAS wild type or mutant) were reflected in the serum miRNomes (Figure 5, F).

Evaluation of the major sources of serum miRNAs

Changes in serum miRNomes in cancer patients are thought to be derived from aberrant miRNA expression in cancer cells. First, we checked the cancer tissue miRNAs that were reported previously to contribute strongly to cancer type discrimination (25). However, of the 100 most prominent of these miRNAs, the majority (80%) was not among the 179 miRNAs identified in this study as strong signatures for cancer classification (Supplementary

Figure 5, available online). This suggests that different serum miRNomes among cancer types cannot be explained by the difference of cancer tissue miRNA expressions. Therefore, we investigated the major cell type sources of the miRNAs that contribute to the prediction of cancer tissue-of-origin by combining several online RNA sequence databases. In the following analysis, we focused on 18 miRNAs showing the highest contribution to cancer type discrimination (contribution >0.05 , red dots in Figure 5, C). Unsupervised clustering analysis of the contributions of the 18 miRNAs showed that 7 miRNAs (miR-6717-5p, miR-3131, miR-122-5p, miR-422a, miR-551b-5p, miR-125a-3p, and miR-1343-3p)

Table 2. Sensitivity of each cancer type at a specificity threshold of 0.95 or 0.99 in DDDmir-DB

Cancer type	With nontumor samples		Without nontumor samples	
	0.95 specificity Sensitivity (95% CI)	0.99 specificity Sensitivity (95% CI)	0.95 specificity Sensitivity (95% CI)	0.99 specificity Sensitivity (95% CI)
Breast cancer	1.00 (1.00 to 1.00)	0.98 (0.94 to 1.00)	1.00 (1.00 to 1.00)	0.98 (0.94 to 1.00)
Bladder cancer	0.99 (0.96 to 1.00)	0.94 (0.85 to 0.99)	0.99 (0.95 to 1.00)	0.88 (0.70 to 0.96)
Biliary tract cancer	0.84 (0.74 to 0.91)	0.64 (0.52 to 0.77)	0.79 (0.70 to 0.89)	0.60 (0.49 to 0.72)
Colorectal cancer	0.96 (0.94 to 0.98)	0.84 (0.78 to 0.88)	0.94 (0.91 to 0.97)	0.81 (0.71 to 0.87)
Esophageal squamous cell carcinoma	0.97 (0.93 to 1.00)	0.92 (0.86 to 0.96)	0.96 (0.90 to 0.99)	0.89 (0.81 to 0.95)
Gastric cancer	1.00 (1.00 to 1.00)	0.87 (0.82 to 0.97)	1.00 (0.99 to 1.00)	0.82 (0.68 to 0.88)
Intraparenchymal brain tumor ^a	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)
Hepatocellular carcinoma	0.94 (0.89 to 0.99)	0.81 (0.71 to 0.90)	0.93 (0.86 to 0.99)	0.77 (0.67 to 0.87)
Lung cancer	0.99 (0.98 to 1.00)	0.94 (0.91 to 0.97)	0.99 (0.97 to 1.00)	0.91 (0.87 to 0.95)
Ovarian cancer	0.96 (0.91 to 1.00)	0.90 (0.81 to 0.96)	0.96 (0.93 to 1.00)	0.88 (0.78 to 0.95)
Pancreatic cancer	0.99 (0.97 to 1.00)	0.92 (0.87 to 0.97)	0.98 (0.95 to 1.00)	0.88 (0.80 to 0.94)
Prostate cancer	1.00 (1.00 to 1.00)	0.96 (0.92 to 0.98)	1.00 (0.99 to 1.00)	0.93 (0.88 to 0.97)
Sarcoma	0.97 (0.92 to 1.00)	0.88 (0.78 to 0.97)	0.95 (0.88 to 1.00)	0.90 (0.80 to 0.97)

^a Intraparenchymal brain tumor such as glioma. CI = confidence interval; DDDmir-DB = Development and Diagnostic Technology for Detection of miRNA in Body Fluids.

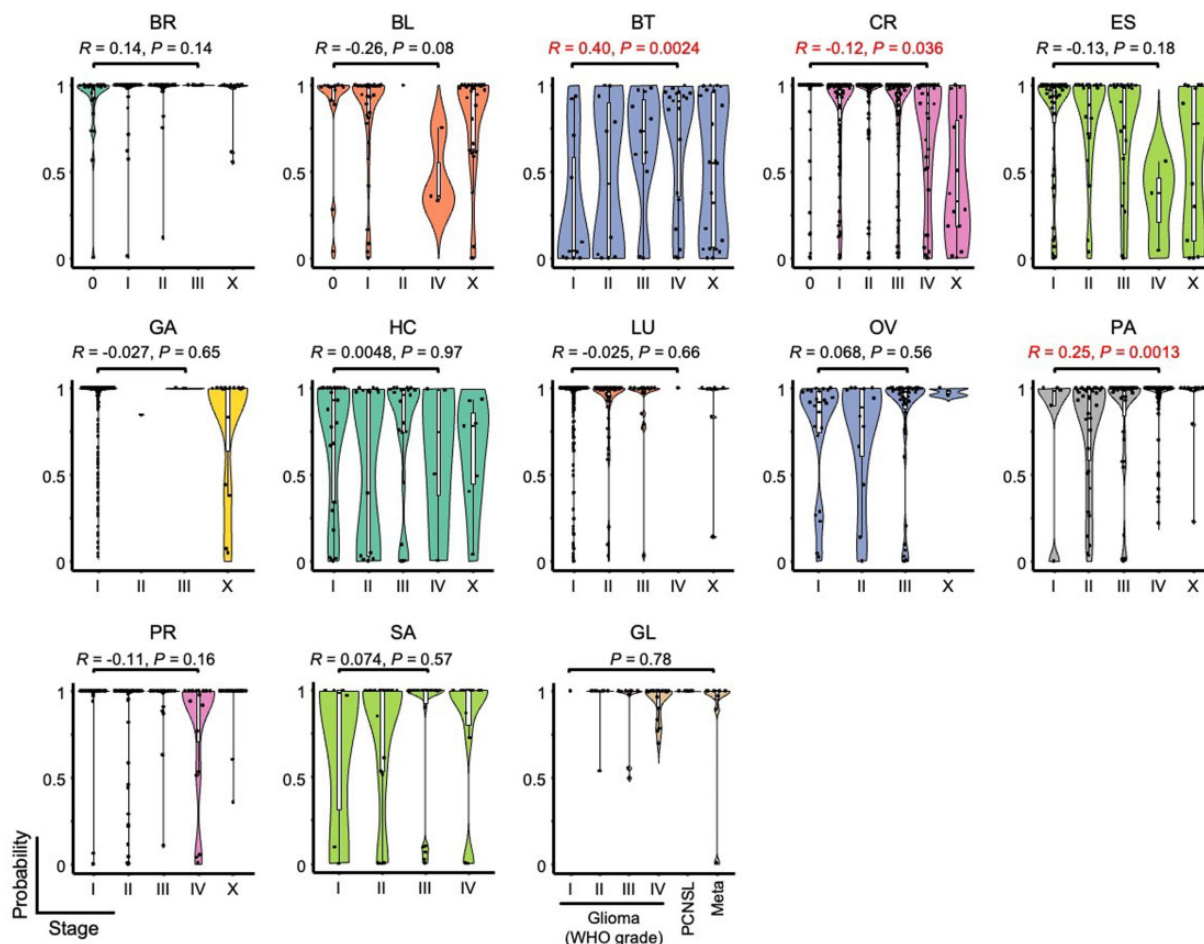


Figure 3. Correlations between the probability scores and disease stages. Violin plots indicating the distribution of the probability scores in each disease stage in each cancer type. For GL, a violin plot indicating the distribution of the probability scores in each histological subgroup. In BT and PA, the scores were lower in the early stage than in the late stage. Correlation coefficient (R) and P values were calculated by Pearson correlation analysis. Statistically significant R and P values are shown in red. P, one-way analysis of variance. BL = bladder cancer; BR = breast cancer; BT = biliary tract cancer; CR = colorectal cancer; ES = esophageal squamous cell carcinoma; GA = gastric cancer; GL = intraparenchymal brain tumor such as glioma; HC = hepatocellular carcinoma; LU = lung cancer; Meta = metastatic brain tumor; OV = ovarian cancer; PA = pancreatic cancer; PCNSL = primary central nervous system lymphoma; PR = prostate cancer; SA = sarcoma; WHO = World Health Organization.

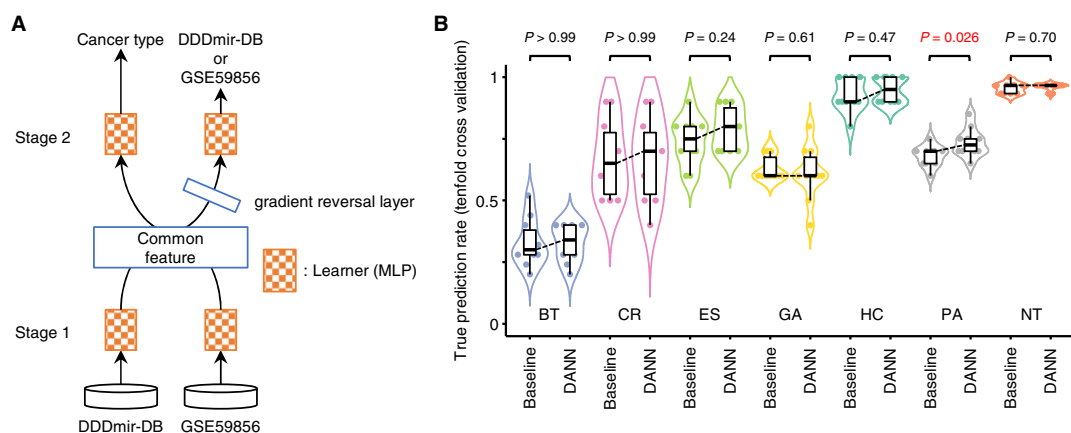


Figure 4. Schematic view of the domain adversarial neural network (DANN). **A**) The DANN consists of a common feature extraction network (stage 1) and a combination of a classifier for cancer diagnosis prediction and a domain classifier for predicting the source of the dataset (stage 2). A gradient reversal layer reverses the sign of the error back to propagation from the domain prediction bifurcation thus reducing the accuracy of the domain prediction as much as possible. This enables extracting the characteristics of the cancer regardless of the influence of the domain. **B**) Differences in the true prediction rate for each cancer type between before and after DANN analysis in the GSE59856 dataset. Transfer learning of the DDDmir-DB improved the diagnostic performance in the GSE59856 dataset. Statistically significant *P* values are shown in red. *P*, student *t* test. BT = biliary tract cancer; CR = colorectal cancer; DDDmir-B = Development and Diagnostic Technology for Detection of miRNA in Body Fluids; ES = esophageal squamous cell carcinoma; GA = gastric cancer; HC = hepatocellular carcinoma; miRNA = microRNA; MLP = multilayer perceptron; NT = nontumor; PA = pancreatic cancer.

commonly contributed to discrimination of cancer type in GSE59856 and DDDmir-DB (ie, their serum levels varied the most among cancer types) (Figure 6, A). To estimate the cell sources of these miRNAs, tissue miRNA expression profiles of 6141 cancer tissue samples and 399 surrounding noncancer tissue samples from The Cancer Genome Atlas (TCGA) were used. We analyzed the correlation of the average levels of 18 miRNAs for each cancer type between sera (DDDmir-DB) and tissue (TCGA) (Figure 6, B). The analysis included 10 cancer types with available TCGA data for cancer and noncancer tissues. A positive correlation between serum and tissue, especially in cancer tissues, would suggest that cancer-specific aberrant miRNA expression influences serum miRNomes. However, the correlation coefficients for serum and tissue were higher in noncancer than in cancer tissues (HC, lung, GL, and BT) or similar between the two (esophageal squamous cell and bladder). For BR and pancreatic cancers, serum and tissue miRNA levels were negatively correlated in cancer and noncancer tissues. Colorectal and gastric cancers were the only samples in which cancer tissue-specific miRNA expression may affect the serum miRNomes. These results suggest that cancer cells are not the major source of serum miRNAs involved in the discrimination between cancer types.

To further estimate the sources of highly contributing serum miRNAs, the tissue levels of the 18 miRNAs were assessed in each organ using TCGA data (Figure 6, C). Notably, most of the 18 miRNAs were dominantly expressed in the brain (GL_N) or the liver (HC_N, BT_N). Among the 7 commonly contributing miRNAs, miR-1343 and miR-125A were brain-dominant miRNAs, whereas miR-122 was among liver-dominant miRNAs, consistent with the fact that the miR-125 family and miR-122-5p are expressed abundantly in the brain and in hepatocytes, respectively (26, 27). This result led us to speculate that extracellular miRNAs released from the brain (and/or the neuronal systems) and the liver, the largest organs derived from the ectoderm and the endoderm, respectively, may be more important than miRNAs derived from other organs to estimate the tissue-of-origin using serum miRNAs.

Nonparenchymal cell types, such as blood cells, should be considered as a potential source of serum miRNAs (28). We used the Database of Small Human Noncoding RNAs (v2.0) to obtain

miRNA expression data in various nonparenchymal cells as well as in liver and brain tissues (Figure 6, D). This analysis showed that some miRNAs highly contributing to the prediction of cancer tissue-of-origin were abundantly expressed in CD4⁺ T cells.

Finally, we focused on the well-known liver-dominant miRNA—miR-122-5p. Serum levels of miR-122-5p were highest in patients with BT or HC (Figure 6, E). Serum miR-122-5p levels did not differ between early and progressive stages in BT and HC (Figure 6, F). Considering that miR-122-5p acts as a tumor-suppressor miRNA and is downregulated in HC and cholangiocarcinoma cells (29, 30), the increase in serum miR-122-5p in patients with BT or HC is likely due to the release of this miRNA from hepatocytes damaged by the presence of a tumor.

Discussion

The present large-scale serum miRNome analysis demonstrated that the serum miRNA profile retains data on cancer tissue-of-origin, a concept that was confirmed using 2 independent datasets. High prediction performance was achieved for most cancer types, even in early stage cancers, showing a superior performance to that of ctDNA and conventional tumor markers such as cancer antigens 19-9 and 125 (5-7, 31). The circulating miRNome may be altered by changes in the secretion of miRNAs from tumor environmental cells and blood cells that sense the emergence of cancer, which would enable highly sensitive cancer diagnoses irrespective of disease stage. Despite the unsuccessful diagnosis of benign or cancer disease for BR, GL, and OV, serum miRNomes accurately predicted the disease location. This suggests that serum miRNomes are a useful tool for the determination of cancer tissue-of-origin.

Several key serum miRNAs for discriminating among cancer types that were identified in the present study could provide useful information regarding the major origins of ex-miRNAs. Among the identified key miRNAs, brain- or liver-predominant miRNAs, such as miR-125a-3p and miR-122-5p, were consolidated, suggesting that the brain (or neuronal system) and the liver could act as control towers for circulating miRNomes. Because recent studies pointed out that a liver-brain-gut neural

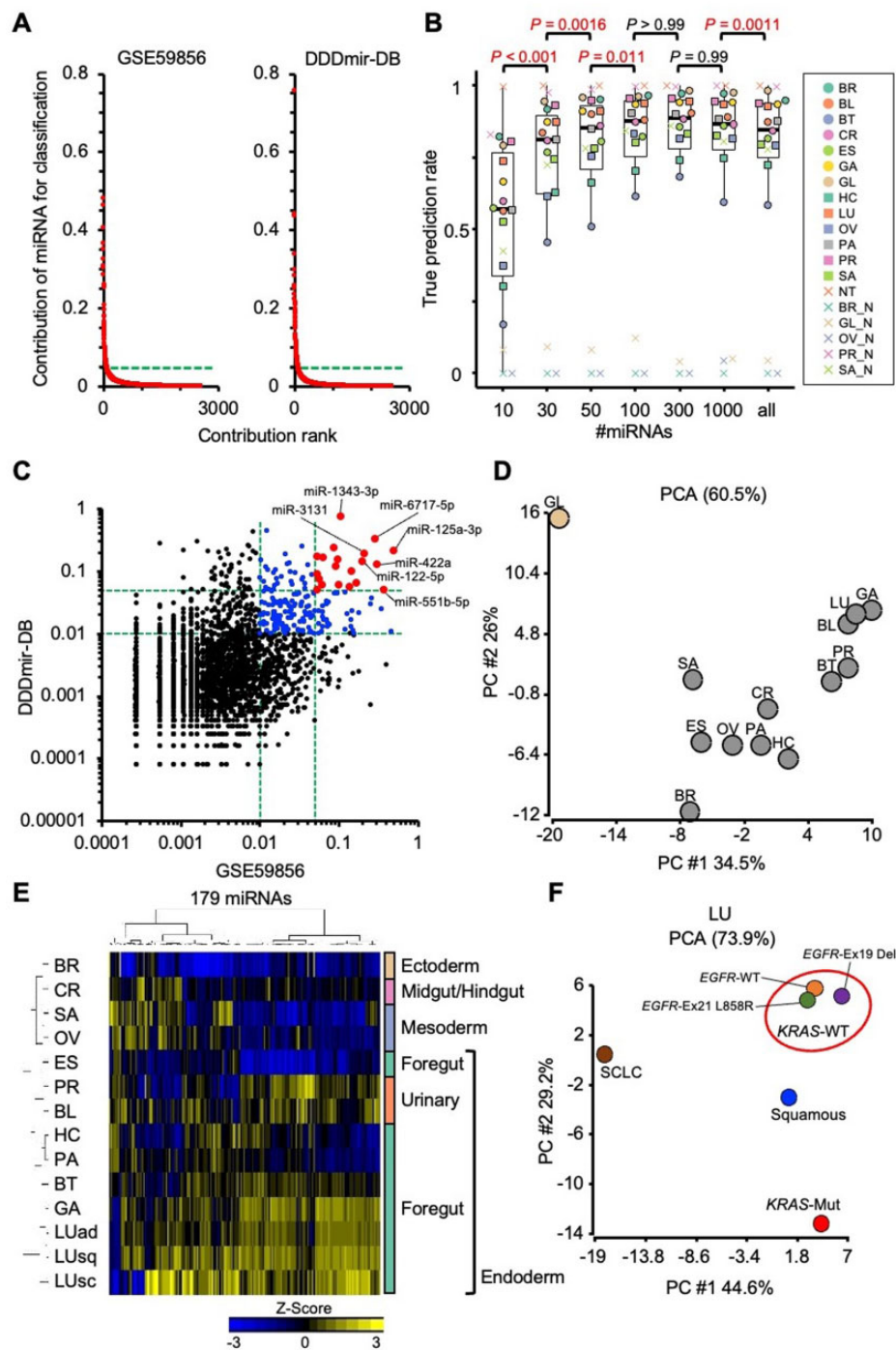


Figure 5. Extraction of highly contributing serum miRNAs for cancer classification. **A)** The contribution of each miRNA to multiclass discrimination was calculated based on the information obtained by splits in nodes in decision trees. The mean contributions in fivefold cross-validation were plotted. **B)** Diagnostic sensitivities computed by the HEAD model using the indicated number of strongly contributing miRNAs in DDDmir-DB. Sensitivities reached the optimal levels when 100 miRNAs were used. Statistically significant P values are shown in red. P , paired t test with Bonferroni correction. **C)** The correlation of the contribution to multiclass discrimination between 2 datasets. miRNAs with a contribution greater than 0.05 or 0.01 for both datasets were plotted as **red dots** or **blue dots**, respectively. **D)** PCA plot of the average serum miRNA levels in 13 cancer types after excluding GL. **E)** Heatmap with unsupervised clustering of the average serum miRNA levels in 12 cancer types after excluding GL. **F)** PCA plot of the average serum miRNA levels in histological subtypes of LU (with KRAS- and EGFR-mutation status). BL = bladder cancer; BR = breast cancer; BT = biliary tract cancer; CR = colorectal cancer; DDDmir-B = Development and Diagnostic Technology for Detection of miRNA in Body Fluids; KRAS = V-Ki-Ras2 Kirsten Rat Sarcoma Viral Oncogene Homolog; EGFR = Epidermal Growth Factor Receptor; ES = esophageal squamous cell carcinoma; GA = gastric cancer; HC = hepatocellular carcinoma; HEAD = Hierarchical Ensemble Algorithm with Deep learning; LU = lung cancer; LUad = lung adenocarcinoma; LUsc = lung small cell carcinoma; LUsc = lung squamous cell carcinoma; miRNA = microRNA; mut = mutation; N = benign disease; OV = ovarian cancer; PA = pancreatic cancer; PCA = principal component analysis; PR = prostate cancer; SA = sarcoma; WT = wild type.

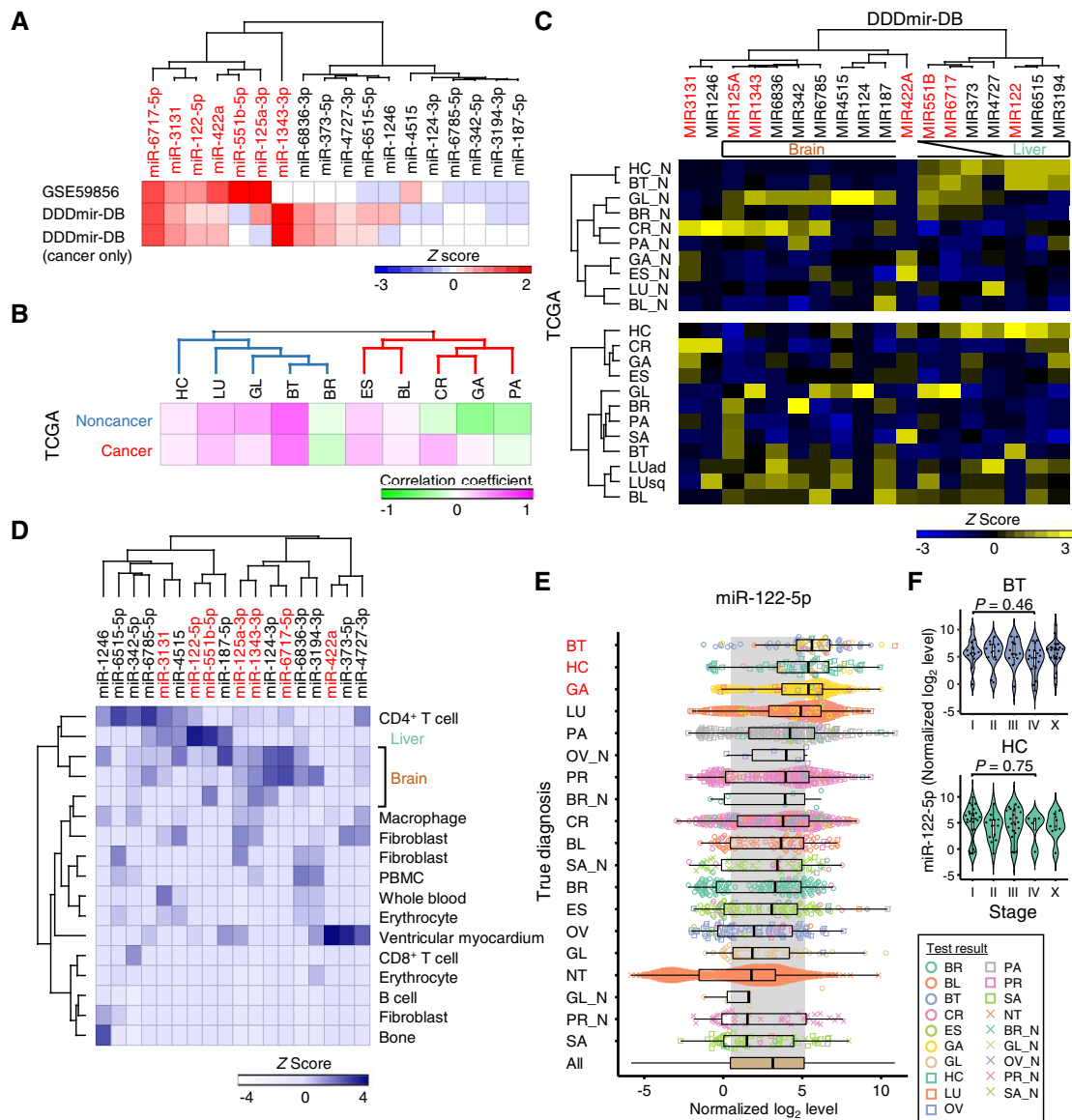


Figure 6. Comparison of miRNomes between serum and tissue. **A**) Unsupervised hierarchical clustering analysis of the contributions of the highly contributing 18 miRNAs for multiclass discrimination. Contributions were calculated for all-class in GSE59856, all-class in DDDmir-DB, or 13-class (among cancer samples) in DDDmir-DB. **B**) Clustering analysis of the correlation coefficient between tissue and serum miRNA levels for each cancer type. In the **blue-lined** cluster, the correlation coefficient between serum miRNAs and noncancer tissue miRNAs in the corresponding organ was higher than that between serum miRNAs and cancer tissue miRNAs in the corresponding organ. The opposite pattern was observed in the **red-lined** cluster. **C**) Clustering analysis of TCGA miRNA data. **Red-letter** miRNAs are the serum miRNAs that contributed the most to cancer discrimination. **D**) Clustering analysis of Database of Small Human Noncoding RNAs (v2.0) miRNA data. **Red-letter** miRNAs are the serum miRNAs that contributed the most to cancer discrimination. **E**) Distribution of miR-122-5p levels in each cancer type. **Gray background** indicates the upper quartile and median levels among all samples. x-axis labels indicate the true diagnosis (red letters = greater than upper quartile; blue letters = less than median among all samples). **Dot** colors and shapes indicate the test results. **F**) Serum levels of miR-122-5p in each stage in BT or HC participants. *P*, one-way analysis of variance. AUC = area under the ROC curve; BL = bladder cancer; BR = breast cancer; BT = biliary tract cancer; CR = colorectal cancer; DDDmir-DB = Development and Diagnostic Technology for Detection of miRNA in Body Fluids; ES = esophageal squamous cell carcinoma; GA = gastric cancer; GL = intraparenchymal brain tumor such as glioma; HC = hepatocellular carcinoma; HEAD = Hierarchical Ensemble Algorithm with Deep learning; LU = lung cancer; LUad = lung adenocarcinoma; LUSq = lung squamous cell carcinoma; miR = mature miRNA; miRNA = microRNA; N = benign disease; NT = nontumor; OV = ovarian cancer; PA = pancreatic cancer; PBMC = peripheral blood mononuclear cells; PR = prostate cancer; ROC = receiver operating characteristic; SA = sarcoma; TCGA = The Cancer Genome Atlas.

arc plays an important role in cancer progression and inflammation (32, 33), the liver and brain might supervise the malignant status of all organs in the body. Although brain-derived miRNAs have been studied as biomarkers for neurological disorders, alterations in the circulating levels of these miRNAs in cancer patients have not been investigated in detail. Neuron-derived EVs are thought to regulate neurogenesis and angiogenesis (34, 35), suggesting that cancer could hijack these systems by ectopically

expressing miRNAs or somehow interacting with distant organs to release EVs (36, 37). T cell-derived EVs and miRNAs have also been studied as biomarkers for inflammatory disorders, such as hepatitis and myocarditis (38, 39). Further examination of miRNA networks between cancer cells and T cells is warranted considering the companion diagnostics tests for immune checkpoint inhibitors and the optimized EV therapy using genetically engineered T cells expressing a chimeric antigen receptor (40, 41).

The present study has the following limitations. First, this study is a case-control study using preserved samples, thus, we cannot exclude the possibility that differences in sample storage durations affected the analysis. Second, we were only able to collect 5 benign disease control samples, which are not enough to evaluate the cancer detection accuracy of serum miRNomes. Therefore, we focused on the prediction accuracy of cancer tissue-of-origin in this study. Third, the comprehensive miRNA analytical method used in our study is not the latest method, such as small RNA sequencing. Thus, further validations using optimal samples and methods are warranted. We have already begun prospective validation studies in Japan (2017-044, approved by the NCC Hospital institutional review board) and in the United States (NCT04671498).

In conclusion, this study demonstrates that serum miRNA analysis is a feasible strategy for predicting cancer tissue-of-origin even for early stage cancers. This concept should pave the way for further clinical and biological validation. Improving our understanding of the molecular mechanisms underlying the regulation of circulating miRNA profiles in the body could help propel the use of ex-miRNA diagnostics into clinical practice.

Funding

This study was supported by Japan Agency for Medical Research and Development grant 14ae0101011h0001, 15ae0101011h0002, 16ae0101011h0003, 17ae0101011h0004, and 18ae0101011h0005 (TOc) and the National Cancer Center Research and Development Fund 29-A-1 for the maintenance of the National Cancer Center Biobank.

Notes

Role of the funder: The funders provided financial assistance for this work. The funders did not play a role in the design; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Author disclosures: ST is employed by Toray Industries, the provider of the microarray system. KO, TA, and DO are employed by Preferred Networks, the developer of the machine learning system. NO is employed by Preferred Medicine and supervised the development of the machine learning system. The other authors declare no competing interests.

Author contributions: Conceptualization: JM, KK, ST, FT, NO. Formal Analysis: JM, KO, NT, NO. Funding acquisition: TOc. Investigation: JM, ST, KoSh, TK. Methodology: FT, HS, ST. Resources: KK, KSu, AS, KT, SS, TKi, HD, TW, HK, HO, YK, HT, SA, YS, NB, SK, HU, Tok, KaSh, YO, KA, YY, SW, NA, AK, MO, YN, MI, TKa, HF, SN, HN. Software: KO, NO, TA, DO. Supervision: TOc. Visualization: JM. Writing—original draft: JM, KO, NT, NO. Writing—review & editing: All authors

Acknowledgements: The authors thank Tomomi Fukuda, Takumi Sonoda, Hiroko Tadokoro, Megumi Miyagi, Tatsuya Suzuki, Junpei Kawauchi, Makiko Ichikawa, and Kamakura Techno-Science for performing the microarray assays; Satoshi Kondo for technical support; Noriko Abe for management of serum samples; Michiko Ohori for management of personal information; and Hiroshi Sato for management of intellectual property rights.

Data availability

The data that support the findings of this study are available in the main text or [supplementary materials](#). The microarray data have been deposited in the Gene Expression Omnibus (GEO) (GSE211692). The code used in this study is available on GitHub (https://github.com/pfnet-research/head_model).

References

1. Raof S, Kennedy CJ, Wallach DA, et al. Molecular cancer screening: in search of evidence. *Nat Med*. 2021;27(7):1139-1142.
2. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. 2018;359(6378):926-930.
3. Cristiano S, Leal A, Phallen J, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019;570(7761):385-389.
4. Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. *Science*. 2020;369(6499):eabb9601.
5. Shen SY, Singhania R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature*. 2018;563(7732):579-583.
6. Liu MC, Oxnard GR, Klein EA, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol*. 2020;31(6):745-759.
7. Klein EA, Richards D, Cohn A, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol*. 2021;32(9):1167-1177.
8. Murillo OD, Thistlethwaite W, Rozowsky J, et al. exRNA atlas analysis reveals distinct extracellular RNA cargo types and their carriers present across human biofluids. *Cell*. 2019;177(2):463-477.e15.
9. Yokoi A, Matsuzaki J, Yamamoto Y, et al. Integrated extracellular microRNA profiling for ovarian cancer screening. *Nat Commun*. 2018;9(1):4319.
10. Asakura K, Kadota T, Matsuzaki J, et al. A miRNA-based diagnostic model predicts resectable lung cancer in humans with high accuracy. *Commun Biol*. 2020;3(1):134.
11. Asano N, Matsuzaki J, Ichikawa M, et al. A serum microRNA classifier for the diagnosis of sarcomas of various histological subtypes. *Nat Commun*. 2019;10(1):1299.
12. Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. *Nature*. 2005;435(7043):834-838.
13. Rosenfeld N, Aharonov R, Meiri E, et al. MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol*. 2008;26(4):462-469.
14. Hoadley KA, Yau C, Hinoue T, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*. 2018;173(2):291-304.e6.
15. Telonis AG, Magee R, Loher P, et al. Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res*. 2017;45(6):2973-2985.
16. English TM. Stacked generalization and simulated evolution. *Biosystems*. 1996;39(1):3-18.
17. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63(1):3-42.
18. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. *Ann Stat*. 2000;28(2):400-407.
19. Pal SK, Mitra S. Multilayer perceptron, fuzzy sets, and classification. *IEEE Trans Neural Netw*. 1992;3(5):683-697.

20. Kuncheva LI, Whitaker CJ. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Mach Learn.* 2003;51(2):181-207.
21. Kojima M, Sudo H, Kawauchi J, et al. MicroRNA markers for the diagnosis of pancreatic and biliary-tract cancers. *PLoS One.* 2015; 10(2):e0118220.
22. Usuba W, Urabe F, Yamamoto Y, et al. Circulating miRNA panels for specific and early detection in bladder cancer. *Cancer Sci.* 2019;110(1):408-419.
23. Ohno M, Matsuzaki J, Kawauchi J, et al. Assessment of the diagnostic utility of serum microRNA classification in patients with diffuse glioma. *JAMA Netw Open.* 2019;2(12):e1916953.
24. Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks. *J Mach Learn Res.* 2016;17(1):2096-2030.
25. Lopez-Rincon A, Martinez-Archundia M, Martinez-Ruiz GU, et al. Automatic discovery of 100-miRNA signature for cancer classification using ensemble feature selection. *BMC Bioinformatics.* 2019;20(1):480.
26. Sempere LF, Freemantle S, Pitha-Rowe I, et al. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol.* 2004;5(3):R13.
27. Shifeng H, Danni W, Pu C, et al. Circulating liver-specific miR-122 as a novel potential biomarker for diagnosis of cholestatic liver injury. *PLoS One.* 2013;8(9):e73133.
28. Pritchard CC, Kroh E, Wood B, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila).* 2012;5(3):492-497.
29. Girard M, Jacquemin E, Munnich A, et al. miR-122, a paradigm for the role of microRNAs in the liver. *J Hepatol.* 2008;48(4): 648-656.
30. Liu N, Jiang F, He TL, et al. The roles of microRNA-122 over-expression in inhibiting proliferation and invasion and stimulating apoptosis of human cholangiocarcinoma cells. *Sci Rep.* 2015;5:16566.
31. Cwik G, Wallner G, Skoczylas T, et al. Cancer antigens 19-9 and 125 in the differential diagnosis of pancreatic mass lesions. *Arch Surg.* 2006;141(10):968-973; discussion 974.
32. Hondermarck H, Huang PS, Wagner JA. The nervous system: orchestra conductor in cancer, regeneration, inflammation and immunity. *FASEB Bioadv.* 2021;3(11):944-952.
33. Teratani T, Mikami Y, Nakamoto N, et al. The liver-brain-gut neural arc maintains the Treg cell niche in the gut. *Nature.* 2020; 585(7826):591-596.
34. Sharma P, Mesci P, Carrromeu C, et al. Exosomes regulate neurogenesis and circuit assembly. *Proc Natl Acad Sci U S A.* 2019; 116(32):16086-16094.
35. Ribeiro-Rodrigues TM, Laundos TL, Pereira-Carvalho R, et al. Exosomes secreted by cardiomyocytes subjected to ischaemia promote cardiac angiogenesis. *Cardiovasc Res.* 2017;113(11): 1338-1350.
36. Amit M, Takahashi H, Dragomir MP, et al. Loss of p53 drives neuron reprogramming in head and neck cancer. *Nature.* 2020; 578(7795):449-454.
37. Demir IE, Mota Reyes C, Alrawashdeh W, et al. Future directions in preclinical and translational cancer neuroscience research. *Nat Cancer.* 2021;1:1027-1031.
38. Kornek M, Lynch M, Mehta SH, et al. Circulating microparticles as disease-specific biomarkers of severity of inflammation in patients with hepatitis C or nonalcoholic steatohepatitis. *Gastroenterology.* 2012;143(2):448-458.
39. Blanco-Dominguez R, Sanchez-Diaz R, de la Fuente H, et al. A novel circulating microRNA for the detection of acute myocarditis. *N Engl J Med.* 2021;384(21):2014-2027.
40. Fu W, Lei C, Liu S, et al. CAR exosomes derived from effector CAR-T cells have potent antitumour effects and low toxicity. *Nat Commun.* 2019;10(1):4355.
41. Chen G, Huang AC, Zhang W, et al. Exosomal PD-L1 contributes to immunosuppression and is associated with anti-PD-1 response. *Nature.* 2018;560(7718):382-386.