Comprehensive miRNA expression analysis in peripheral blood can diagnose liver disease

Yoshiki Murakami[1,†1,a)] Hidenori Toyoda[2] Toshihito Tanahashi[3] Junko Tanaka[4] Takashi Kumada[2] Yusuke Yoshioka[5] Nobuyoshi Kosaka[5] Takahiro Ochiya[5] Y-h. Taguchi[6,7,b)]

Abstract: *Background*: miRNAs circulating in the blood in a cell-free form have been acknowledged for their potential as readily accessible disease markers. Presently, histological examination is the golden standard for diagnosing and grading liver disease, therefore non-invasive options are desirable. Here, we investigated if miRNA expression profile in exosome rich fractionated serum could be useful for determining the disease parameters in patients with chronic hepatitis C (CHC). *Methodology*: Exosome rich fractionated RNA was extracted from the serum of 64 CHC and 24 controls with normal liver (NL). Extracted RNA was subjected to miRNA profiling by microarray and real-time qPCR analysis. The miRNA expression profiles from 4 chronic hepatitis B (CHB) and 12 non alcoholic steatohepatitis (NASH) patients were also established. The resulting miRNA expression was compared to the stage or grade of CHC determined by blood examination and histological inspection. *Principal Findings*: miRNAs implicated in chronic liver disease and inflammation showed expression profiles that differed from those in NL and varied among the types and grades of liver diseases. Using the expression patterns of twelve miRNAs, we classified CHC, CHB, NASH and NL with 95.25% accuracy. Additionally, we could link miRNA expression pattern with liver fibrosis stage and grade of liver inflammation in CHC. In particular, the miRNA expression pattern for early fibrotic stage differed greatly from that observed in high inflammation grades. *Conclusions*: We demonstrated that miRNA expression pattern in exosome rich fractionated serum shows a high potential as a biomarker for diagnosing the grade and stage of liver diseases

1. Introduction

MicroRNAs (miRNAs) are a gene family which is evolutionarily conserved and have important roles in the control of many biological processes, such as cellular development, differentiation, proliferation, apoptosis, and metabolism [1]. They are closely related progress of viral hepatitis infections, liver fibrosis, and hepatocarcinogenesis [2], [3], [4]. Recently, two independent groups showed that miR-122 plays a critical role in the maintenance of liver homeostasis and anti-tumor formation [5], [6]

Exosome in one of the endoplasmic reticulum carries mRNAs and miRNAs [7]. Recently, it has become clear that exosome perform intercellular signaling through miRNA. miRNAs are released through a ceramide-dependent secretory machinery and then transferable and functional in the recipient cells [8]. In a prior study using human blood and cultured cells, several miRNAs were selectively packaged into microvesicle (MV) and actively secreted [9]. In another study, miRNAs originating from Epstein-Barr virus (EBV) was transported by exosome and then participated in the immune response of host cells [10]. In HCC cells as well, this type of exosome-mediated miRNA transfer is an important mechanism of intercellular communication [11].

It has also become clear that exosome can adjust to immune function and control infection or carry the virus itself. Exosomes of T, B and dendritic immune cells contain miRNA repertoires that differ from those of their parent cells [12], [13] . Exosomes released from nasopharyngeal carcinoma cells harboring latent EBV were shown to contain LMP1, signal transduction molecules, and virus-encoded miRNAs [14]. Retroviruses evade adaptive immune responses by using nonviral or host exosome biogenesis pathways to form infectious particles and as a mode of infection[15].

Recent evidence has shown that expression pattern of serum or plasma miRNAs are altered in several diseases, in particular heart disease, sepsis, malignancies, and autoimmune diseases (reviewed in [16]). Tumour-associated miRNAs in diffuse large B-cell lymphoma patients were found in serum . Circulating miRNAs are detectable in serum and plasma in a form sufficiently stable to serve as biomarkers [17], [18]. One such example is that

[1] Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan
[2] Department of Gastroenterology, Ogaki Municipal Hospital, Ogaki 503-8502,Japan
[3] Department of Medical Pharmaceutics,Kobe Pharmaceutical University, Kobe 658-8558, Japan
[4] Department of Epidemiology, Infectious Disease Control and Prevention, Hiroshima University Graduate School of Biomedical Sciences, Hiroshima 734-8551, Japan
[5] Division of Molecular and Cellular Medicine, National Cancer Center Research Institute, Tokyo 104-0045, Japan
[6] Department of Physics, Chuo University, Tokyo 112-8551, Japan
[7] Corresponding Author
[†1] Presently with Department of Hepatology, Graduate School of Medicine, Osaka City University
[a)] m2079633@med.osaka-cu.ac.jp
[b)] tag@granular.com

tumour-associated miRNAs were found in the serum of diffuse large B-cell lymphoma patients[19]. In other example, serum levels of miR-34a and miR-122 may represent histological disease severity in patients with CHC or non-alcoholic fatty-liver disease (NAFLD) [20]. In fact, the serum level of miR-122 strongly correlates with serum ALT activity and with necro-inflammatory activity in patients with CHC and elevated ALT levels. However, there seems to be no significant correlation with fibrosis stage and functional capacity of the liver [21]. The expression levels of miR-122 and miR-194 correlated negatively with the age in patients with CHB and Hepatitis B virus (HBV) associated acute-on-chronic liver failure [22]. The expression level of miR-122 and 192 in serum are closely related to drug induce liver injury [23]. Based on above, it comes as no surprise therefore that recently the expression profile from extracellular miRNA is being used clinically to diagnose various diseases.

Here, in order to obtain data with high resolution and reproducibility microarray analysis was comprehensively conducted after extracting the MVs in serum using exoquick. We attempted to diagnose Hepatitis C virus (HCV) infection, the degree of liver inflammation and fibrotic stage using exosome-rich fractioned miRNA. In short we investigate if serum-derived miRNAs have the potential to serve as non-invasive markers for various liver diseases [24].

2. Materials and Methods

2.1 miRNA expression

The data presented in this manuscript was partially deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series access number GSE33857: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33857. miRNA expression (gProcessedSignal) of each sample was normalized so as to have zero mean and common standard deviation of 1.

2.2 "*in silico* " patients resampling

Insufficient number of sampling patients can cause the following two types of problems.

(1) *Sampling bias*: Although the number of sampling patients is large enough, the sampling does not cover the cohort but a limited part of the cohort.

(2) *Sampling error*: Although sampling is not biased, sampling error can occur if a set of sampling patients does not cover the cohort throughout.

The example of sampling bias is age. Patients are often older than normal control. The difference between patients and healthy control may be caused by not the disease but the aging. On the other hands, the number of patients necessary for the aimed research often has minimum. Suppose that we would like to investigate the progress of a disease year by year over ten years. Then we need at least twenty patients, since there must be a male and a female patient for each of ten years. If we have only one patient for each of ten years, we can have insufficient number of samples, e.g., a male patient for the first year, a female patient for the second year, and so on. Then the outcome is inevitably biased.

In order to compensate this problem, we should try to collect enough number of patients throughout. However, if it is impossible to gather large enough number of patients, *in silico* resampling can weaken these two sampling problems. Suppose we have N patients with M types of clinical information $c_{ij}, i = 1, \ldots, N, j = 1, \ldots, M$. Then the *in silico* resampling for the ℓth resampled patient can be done by picking one $c_j^{(\ell)}$ for each j from $\{c_{ij}, i = 1, \ldots, N\}$. This gives us a set of patients having clinical information $\{c_j^{(\ell)}, j = 1, \ldots, M\}$, $\ell = 1, \ldots, L$, where L is the number of resampled patients

In order to infer the miRNA expression attributed to the ℓth resampled patient, we employed the Markov Chain Monte Carlo (MCMC) regression based upon the MCMCregress function in the MCMCpack[*1].

Before the inference, we must first model the relationship between the miRNA expression and the clinical information. We assume that the ith patient's kth miRNA expression, miRNA_{ik}, is related to clinical information as follows,

$$\text{miRNA}_{ik} = \beta_0 + \sum_{j=1}^{M} \beta_j c_{ij} + \epsilon_{ik}.$$

The MCMC regression generates β_j and ϵ_{ik}^2 that obey the following distributions,

$$\beta_j \sim \mathcal{N}(b_0, B_0^2),$$

and

$$\epsilon_{ik}^2 \sim \mathcal{G}amma(c_0/2, d_0/2),$$

where $\mathcal{N}$ and $\mathcal{G}amma$ are the Gaussian distribution and the Gamma distribution respectively. Using MCMC, we generate the parameter set, $\{\beta_j^{(\ell)}, j = 1, \ldots, M; \epsilon^{(\ell)^2}\}$, for the inference of the miRNA expression of the ℓth resampled patient, $\text{miRNA}_k^{(\ell)}$. Then $\text{miRNA}_k^{(\ell)}$ is generated by the following formula,

$$\text{miRNA}_k^{(\ell)} = \beta_0 + \sum_{j=1}^{M} \beta_j^{(\ell)} c_j^{(\ell)} \pm \epsilon^{(\ell)},$$

where ± takes + or − with equal probability.

2.2.1 "*in silico*" patients resampling for disease discriminant studies

On discriminations among CHC, NASH, CHB and normal control, we used *in silico* patients resampling. The clinical information considered for the modeling are age, gender and body mass index (BMI). The ith patient's miRNA expression is modeled as

$$\text{miRNA}_{ik} = \beta_0 + \beta_{\text{age}} \times \text{age} + \beta_{\text{gender}} \times \text{gender} + \beta_{\text{BMI}} \times \text{BMI} + \epsilon_{ik},$$

where age and BMI are actual values and gender is 1 (male) or 0 (female). For each miRNA selected for the discrimination between a pair of patient groups, MCMC is applied. Then a hundred resampled patients are **commonly used** to generate miRNA expression in each group. This means, virtually clinical information of patients in each group is same. Then linear discriminant analysis (LDA) with principal component analysis (PCA) is applied to discriminate two hundred patients into two groups.

[*1] http://mcmcpack.wustl.edu/, all parameters other than b_0 or B_0 are default values. b_0 and B_0 are set to be 1.0.

2.2.2 "*in silico*" patients resampling for qPCR validation of microarray results

On the validation of microarray results by qPCR, we used *in silico* patients resampling. In addition to clinical information employed in the previous section, we also used inflammation and fibrosis, i.e., miRNA expression by qPCR or microarray is modeled as,

$$\begin{aligned}\text{miRNA}_{ik} &= \beta_0 + \beta_{\text{age}} \times \text{age} + \beta_{\text{gender}} \times \text{gender}\\ &+ \beta_{\text{BMI}} \times \text{BMI} + \beta_{\text{fibrosis}} \times \text{fibrosis stage}\\ &+ \beta_{\text{inflammation}} \times \text{inflammation stage} + \epsilon_{ik},\end{aligned}$$

where both inflammation stages and fibrosis stages are 0, 1, 2, and 3. Based on this model, both qPCR and 16 microarray probes values are generated from **commonly used** one thousand resampled patients. The coincidence between qPCR and micorarray expression averaged over 16 probes is checked.

2.3 Semi-supervised learning for independent samples

In order to validate our method, we prepared independent samples that consist of 8 NASH, 31 CHC and 16 CHB samples (not deposited in GEO). Independent samples were normalized as well and were discriminated by semi-supervised learning. In semi-supervised learning, at first, PCA was applied to the joint set of original samples and independent samples. Then, PCA-based LDA was performed with the classification information of original sets, but without that of independent samples. Then, obtained discriminant function was used to discriminate independent samples. Performance was measured based on this result.

2.4 Discrimination between inflammation/fibrosis stages

In order to discriminate between disease progression stages, we employed P-values that reject null hypothesis that specific miRNA expression does not alter between a patients from normal controls in favor of the significance upregulation of the miRNA compared with normal control. P-values were computed by t test implemented in t.test function[25]. Then, method 2 PCA (see below) was applied to compute PC scores for each sample after obtained P-values were log-transformed, i.e., In method 2 PCA, x_{ji} is replaced with $\log P_{ji}$, where P_{ji} represents P-value attributed to the ith miRNA of the jth sample. LDA was performed PCs in combination with ages of patients.

2.5 Feature extraction based upon PCA

Suppose we have miRNA profiles $x_{ij}, (i = 1, \ldots, N, j = 1, \ldots, M)$, each of which corresponds to ith miRNA in jth sample. Samples are classified into L clinical sets, $G_l, (l = 1, \ldots, L)$. Then we have applied PCA to the set of $\{x_{ij}\}$ in two ways;

(1) Method 1 (miRNA based): Substitute $K_s(< M)$ principal component (PC) score x_{ik} to x_{ij}. In this case, PCA is applied to a matrix $\{x_{ij}\}$.

(2) Method 2 (sample based): Substitute $K_m(< N)$ principal component (PC) score x_{kj} to x_{ij}. In this case, PCA is applied to a transverse matrix $\{x_{ji}\}$.

The PCA based feature extraction is as follows;

(1) Step one : Choose a pair of clinical sets, l and l'.

(2) Step two : Compute x_{ik} with method 1 PCA from $\{x_{ij} \mid j \in G_l \cup G_{l'}\}$.

(3) Step three : Compute distance r_i,

$$r_i \equiv \sqrt{\sum_{k=1}^{K_s^0} x_{ik}^2},$$

where $K_s^0(< K_s)$ is the number of components to be used for feature selection.

(4) Step four : Select miRNAs i' with top $N_1(< N)$ r_is.

N_1 miRNAs are a set of selected features to distinguish clinical sets l and l'. Thoughout this paper, K_s^0 is constantly taken to be 2, if not explicitly denoted. PCA is computed by prcomp function in R[25] base package.

One should notice that PCA based feature extraction do not make use of classification information at all. This method is classification free method and is very unique because of this point.

2.6 The PCA based linear discriminant analysis

The PCA based linear discriminant analysis (LDA) is as follows;

(1) Step one : Choose a pair of clinical sets, l and l'.

(2) Step two : If necessary, apply a feature extraction and reduce number of miRNAs used for LDA.

(3) Step three : Compute $x_{kj}, (k = 1, \ldots, K_m)$ using method 2 PCA.

(4) Step four : Divide sampled into training set and test set.

(5) Step five : Apply LDA to training set.

(6) Step six : Validate the performance of LDA using test set.

(7) Step seven : Repeat steps from four to six many times.

(8) Step eight : Compute performance with averaged values.

One should notice that division between training and test sets are done **AFTER** computation of PCA (and feature extraction if necessary). Thus, x_{kj} include the information of test sets, too. Feature extraction, if applied, is also before division, thus is sampling free. One may think that it is a fake since we do not know classification of test set. However, even if we do not have preknowledge about classifications, we can compute PCA, since we do not need classification information to compute x_{kj}. LDA is computed by lda function in R[25] base package.

3. Results

Fig. 1 shows the two dimensional embedding of paients samples with principal component analysisc (PCA). The twelve miRNAs, hsa-miR-1225-5p, 1275, 638, 762, 320c, 451, 1974_v14.0, 1207-5p, 630, 720, 1246, and 486-5p, were employed by PCA-based feature extraction method [26] in order to generate this plot. Among them, although 1974_v14.0 was discarded from the recent miRBase release, since microarray we used was Release 14 based, it was included into probes. The four patients class, i.e., CHC, CHB, NASH and NL do not look like well separated in Fig. 1, but the discrimination between these four was not bad (Table 1).

In order to compensate the small number of samples, we performed "*in silico*" resampling. Fig. 2 shows the two dimensional embedding of "*in silico*" resampled patients with principal

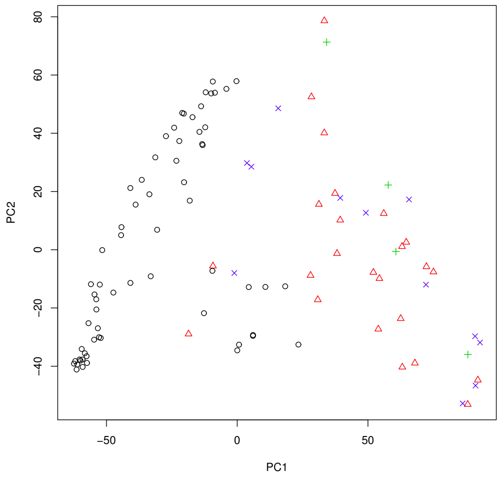


Fig. 1 Two dimensional embedding of patients using PCA. green +: CHB, black ○: CHC, blue ×:NASH and red △ : NL.

Table 1 Discrimination based on 12 selcted miRNAs expression using liner discrminant analysis (LDA). Leave-one-out-cross-validation was employed. Number of PCs used for discrilination is 13. Over all accuracty is 87.5

		result			
		CHB	CHC	NASH	NL
prediction	CHB	2	0	1	2
	CHC	0	64	1	3
	NASH	1	0	9	3
	NL	1	0	1	16

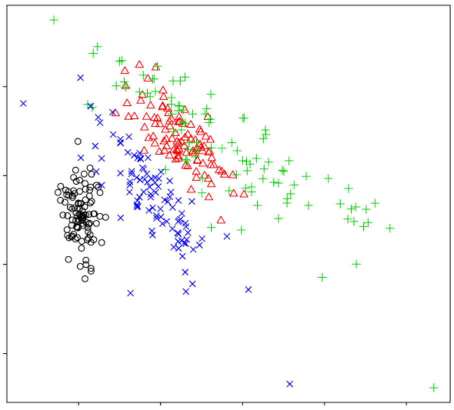


Fig. 2 Two dimensional embedding of patients using PCA of "*in silico*" resampled patients. green +: CHB, black ○: CHC, blue ×:NASH and red △ : NL.

component analysisc (PCA). Separation between different clinical classes was better than real data. Discrimination performance was also very good (Table 2).

Table 2 Discrimination of "*in silico*" resampled patients based on 12 selcted miRNAs expression using liner discrminant analysis (LDA). Leave-one-out-cross-validation was employed. Number of PCs used for discrilination is 5. Over all accuracy is 95.25

		result			
		CHB	CHC	NASH	NL
prediction	CHB	94	0	0	6
	CHC	0	99	1	0
	NASH	6	1	97	3
	NL	0	0	2	91

Since one may think that good performance of "*in silico*" resampling is artifact, we validated this with independent samples. Tables 3, 4 and 5 shows the comparison among original, resampled and independent samples, for CHB vs NASH, CHB vs CHC and CHC vs NASH, respectively. The performances of discrimination for original and resampling was measured by leave-one-out-cross-validation and that for independent sample was measured by semi-supervised learning. Although performance predicted by resampling have tendency to be over estimated, these three, i.e., original, resampling and independent samples were discriminated well with a set of miRNAs (see Table 6) commonly selected by PCA-based feature extraction[26] for these three types of samples.

Table 3 Discrimination performance between CHB and NASH, for original, resmapling, and independent samples. The number of PCs used for discrimination is 11, 3 and 3, respectively.

		result					
		original		resampling		independent	
		CHB	NASH	CHB	NASH	CHB	NASH
prediction	CHB	3	1	91	0	11	0
	NASH	1	11	9	100	5	8
	Accuracy	82.5 %		95.5 %		79.2 %	

Table 4 Discrimination performance between CHB and CHC, for original, resmapling, and independent samples. The number of PCs used for discrimination is 4, 4 and 4, respectively.

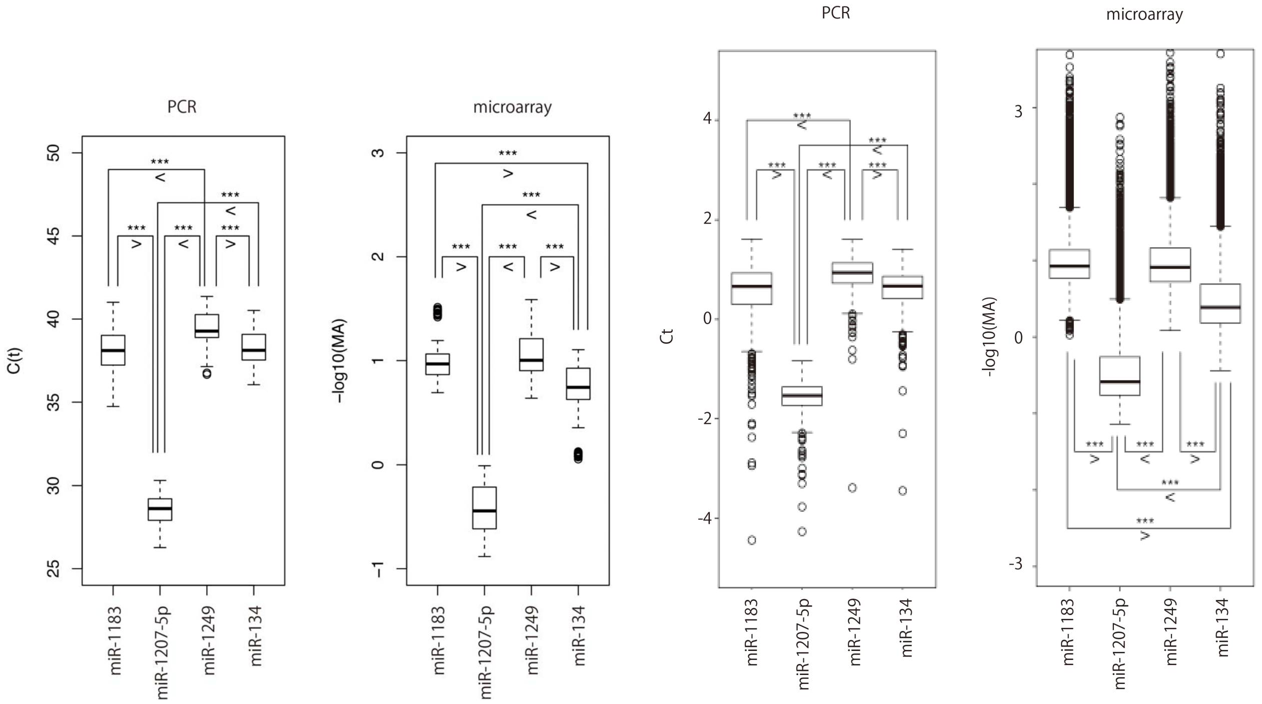
		result					
		original		resampling		independent	
		CHB	CHC	CHB	CHC	CHB	CHC
prediction	CHB	3	0	100	0	15	11
	CHC	1	64	0	100	1	20
	Accuracy	98.5 %		100 %		74.4 %	

Table 5 Discrimination performance between CHC and NASH, for original, resmapling, and independent samples. The number of PCs used for discrimination is 4, 4 and 74, respectively.

		result					
		original		resampling		independent	
		CHC	NASH	CHC	NASH	CHC	NASH
prediction	CHC	64	2	100	0	29	3
	NASH	0	10	0	100	2	5
	Accuracy	97.3 %		100 %		87 %	

In Tables 7 and 8, we showed the discrimination performance of inflammation/fibrosis stages. The performance was not bad. Although it should be validated for additional computation, we expect that a set of exsosome miRNA can diagnose inflammation and fibrosis stages, too.

It is very usual to measure miRNA expression by qPCR in order to validate microarray measurements. Fig. 3 shows the comparison between qPCR and microarray, for both original samples and "*in silico*" resampling. Measured miRNAs are miR-1183, miR-1207-5p, miR-1249 and miR-134. It is clear that "*in silico*" resampling can reproduce the original samples outcome. If we consider the above demonstrated ability of "*in silico*" resampling, we suggest that our results for the comparison between qPCR and microarray will be valid for larger samples, too.



(a) (b)

Fig. 3 Comparison between original sample and "*in silico*" resamplimg. (a) qPCR and microarray measurement of miR-1183, miR-1207-5p, miR-1249 and miR-134, original samples. (b) The same as (a), but "*in silico*" resamplimg.

Table 6 miRNAs selected by PCA-based feature extraction [26] for the discrimination between pairs of clinical classes. They were commonly employed for original, resampled and independent samples.

miRNAs	CHB vs NASH	CHB vs CHC	CHC vs NASH
1225-5p	○	○	○
1275	×	○	○
638	○	○	○
762	○	○	○
320c	○	○	○
1202	×	○	○
486-5p	○	○	○
451	○	○	○
1974	○	○	○
1915	○	○	○
630	○	○	○
483-5p	×	○	○
320b	×	○	○
1207-5p	○	○	○
16	○	×	×
720	○	○	○
1246	○	○	○
320d	○	○	○
92a	○	○	○
22	○	×	○
1202	○	×	×
1268	×	○	○

Table 7 Discrimination between inflammation stages. The number of PCs employed for discrimination is 2, 6, and 4 for A1 vs A0+A2+A3, A2 vs A0+A1+A3 and A3 vs A0+A1+A2, respectively.

		result		Accuracy *P*-value
prediction		A0+A2+A3	A1	
	A0+A2+A3	22	12	71.8 %
	A1	6	24	4.07×10^{-4}
		A0+A1+A3	A2	
	A0+A1+A3	34	4	75 %
	A2	12	14	2.26×10^{-4}
		A0+A1+A2	A3	
	A0+A1+A2	47	3	82.8 %
	A3	8	6	2.30×10^{-3}

Table 8 Discrimination between fibrosis stages. The number of PCs employed for discrimination is 2, 6, 5, and 3 for F0 vs F1+F2+F3, F1 vs F0+F2+F3, F2 vs F0+F1+F3 and F3 vs F0+F1+F2, respectively.

		result		Accuracy *P*-value
prediction		F1+F2+A3	F0	
	F1+F2+F3	54	1	87.5 %
	F0	7	2	4.95×10^{-2}
		F0+F2+F3	F1	
	F0+F2+F3	21	13	64.62 %
	F1	10	20	2.73×10^{-2}
		F0+F1+F3	F2	
	F0+F1+F3	33	4	70.31 %
	F2	15	12	3.24×10^{-3}
		F0+F1+F2	F3	
	F0+F1+F2	33	4	73.44 %
	F3	15	12	1.35×10^{-2}

4. Discussion

Although PCA-based LDA in combination with PCA-based feature extraction could work very well, there were some points to be resolved. For example, the number of PCs used for LDA can often become large. For example, in Table 1, the number of PCs used was as large as 13. This is possibly because of small number of NASH and CHB samples. This is apparent because in Table 2 the number of PCs used is only 5. Similar tendency can be found in Table 3. LDA for original samples that include only 4 CHB samples requires as many as eleven PCs, while LDA for both resampled and independent samples requires only three. This "LDA of small number of samples needs more PCs" tendency will be caused because the contributions of classes with small number of patients can be reflected only in higher order PCs. Since inclusion of too many PCs results in errors, it will be hopeful to find small number of PCs to discriminate classes with small number of patients.

Another point is that we could not apply PCA-based feature extraction to LDA for inflammation/fibrosis progression stages. This is because we had to use *P*-values in combination with ages. PCA-based feature extraction should be extended so as to be capable of the application to more general cases.

5. Conclusion

In this paper, we applied the previously proposed PCA-based feature extraction method [26] to liver diseases. PCA-based LDA in combination with PCA-based feature extraction could discriminate between diseases. The robustness of the method was confirmed with "*in silico*" resampling and semi-supervised learning applied to independent samples. Disease progression (inflammation and fibrosis) could be discriminated using *P*-value with PCA-based LDA.

Acknowledgments This research is supported by KAKENHI, 23300357.

References

[1] Ambros, V.: The functions of animal microRNAs, *Nature*, Vol. 431, No. 7006, pp. 350–355 (2004).

[2] Murakami, Y., Toyoda, H., Tanaka, M., Kuroda, M., Harada, Y., Matsuda, F., Tajima, A., Kosaka, N., Ochiya, T. and Shimotohno, K.: The progression of liver fibrosis is related with overexpression of the miR-199 and 200 families, *PLoS ONE*, Vol. 6, No. 1, p. e16081 (2011).

[3] Murakami, Y., Yasuda, T., Saigo, K., Urashima, T., Toyoda, H., Okanoue, T. and Shimotohno, K.: Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues, *Oncogene*, Vol. 25, No. 17, pp. 2537–2545 (2006).

[4] Braconi, C., Henry, J. C., Kogure, T., Schmittgen, T. and Patel, T.: The role of microRNAs in human liver cancers, *Semin. Oncol.*, Vol. 38, No. 6, pp. 752–763 (2011).

[5] Hsu, S. H., Wang, B., Kota, J., Yu, J., Costinean, S., Kutay, H., Yu, L., Bai, S., La Perle, K., Chivukula, R. R., Mao, H., Wei, M., Clark, K. R., Mendell, J. R., Caligiuri, M. A., Jacob, S. T., Mendell, J. T. and Ghoshal, K.: Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver, *J. Clin. Invest.*, Vol. 122, No. 8, pp. 2871–2883 (2012).

[6] Tsai, W. C., Hsu, S. D., Hsu, C. S., Lai, T. C., Chen, S. J., Shen, R., Huang, Y., Chen, H. C., Lee, C. H., Tsai, T. F., Hsu, M. T., Wu, J. C., Huang, H. D., Shiao, M. S., Hsiao, M. and Tsou, A. P.: MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis, *J. Clin. Invest.*, Vol. 122, No. 8, pp. 2884–2897 (2012).

[7] Valadi, H., Ekstrom, K., Bossios, A., Sjostrand, M., Lee, J. J. and Lotvall, J. O.: Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells, *Nat. Cell Biol.*, Vol. 9, No. 6, pp. 654–659 (2007).

[8] Kosaka, N., Iguchi, H., Yoshioka, Y., Takeshita, F., Matsuki, Y. and Ochiya, T.: Secretory mechanisms and intercellular transfer of microRNAs in living cells, *J. Biol. Chem.*, Vol. 285, No. 23, pp. 17442–17452 (2010).

[9] Zhang, Y., Liu, D., Chen, X., Li, J., Li, L., Bian, Z., Sun, F., Lu, J., Yin, Y., Cai, X., Sun, Q., Wang, K., Ba, Y., Wang, Q., Wang, D., Yang, J., Liu, P., Xu, T., Yan, Q., Zhang, J., Zen, K. and Zhang, C. Y.: Secreted monocytic miR-150 enhances targeted endothelial cell migration, *Mol. Cell*, Vol. 39, No. 1, pp. 133–144 (2010).

[10] Pegtel, D. M., Cosmopoulos, K., Thorley-Lawson, D. A., van Eijndhoven, M. A., Hopmans, E. S., Lindenberg, J. L., de Gruijl, T. D., Wurdinger, T. and Middeldorp, J. M.: Functional delivery of viral miRNAs via exosomes, *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 107, No. 14, pp. 6328–6333 (2010).

[11] Kogure, T., Lin, W. L., Yan, I. K., Braconi, C. and Patel, T.: Intercellular nanovesicle-mediated microRNA transfer: a mechanism of environmental modulation of hepatocellular cancer cell growth, *Hepatology*, Vol. 54, No. 4, pp. 1237–1248 (2011).

[12] Thery, C., Ostrowski, M. and Segura, E.: Membrane vesicles as conveyors of immune responses, *Nat. Rev. Immunol.*, Vol. 9, No. 8, pp. 581–593 (2009).

[13] Mittelbrunn, M., Gutierrez-Vazquez, C., Villarroya-Beltri, C., Gonzalez, S., Sanchez-Cabo, F., Gonzalez, M. A., Bernad, A. and Sanchez-Madrid, F.: Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells, *Nat Commun*, Vol. 2, p. 282 (2011).

[14] Meckes, D. G., Shair, K. H., Marquitz, A. R., Kung, C. P., Edwards, R. H. and Raab-Traub, N.: Human tumor virus utilizes exosomes for intercellular communication, *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 107, No. 47, pp. 20370–20375 (2010).

[15] Gould, S. J., Booth, A. M. and Hildreth, J. E.: The Trojan exosome hypothesis, *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 100, No. 19, pp. 10592–10597 (2003).

[16] Kosaka, N., Iguchi, H. and Ochiya, T.: Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis, *Cancer Sci.*, Vol. 101, No. 10, pp. 2087–2092 (2010).

[17] Mitchell, P. S., Parkin, R. K., Kroh, E. M., Fritz, B. R., Wyman, S. K., Pogosova-Agadjanyan, E. L., Peterson, A., Noteboom, J., O'Briant, K. C., Allen, A., Lin, D. W., Urban, N., Drescher, C. W., Knudsen, B. S., Stirewalt, D. L., Gentleman, R., Vessella, R. L., Nelson, P. S., Martin, D. B. and Tewari, M.: Circulating microRNAs as stable blood-based markers for cancer detection, *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 105, No. 30, pp. 10513–10518 (2008).

[18] Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K., Guo, J., Zhang, Y., Chen, J., Guo, X., Li, Q., Li, X., Wang, W., Zhang, Y., Wang, J., Jiang, X., Xiang, Y., Xu, C., Zheng, P., Zhang, J., Li, R., Zhang, H., Shang, X., Gong, T., Ning, G., Wang, J., Zen, K., Zhang, J. and Zhang, C. Y.: Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases, *Cell Res.*, Vol. 18, No. 10, pp. 997–1006 (2008).

[19] Lawrie, C. H.: MicroRNAs and haematology: small molecules, big function, *Br. J. Haematol.*, Vol. 137, No. 6, pp. 503–512 (2007).

[20] Cermelli, S., Ruggieri, A., Marrero, J. A., Ioannou, G. N. and Beretta, L.: Circulating microRNAs in patients with chronic hepatitis C and non-alcoholic fatty liver disease, *PLoS ONE*, Vol. 6, No. 8, p. e23937 (2011).

[21] Bihrer, V., Friedrich-Rust, M., Kronenberger, B., Forestier, N., Haupenthal, J., Shi, Y., Peveling-Oberhag, J., Radeke, H. H., Sarrazin, C., Herrmann, E., Zeuzem, S., Waidmann, O. and Piiper, A.: Serum miR-122 as a biomarker of necroinflammation in patients with chronic hepatitis C virus infection, *Am. J. Gastroenterol.*, Vol. 106, No. 9, pp. 1663–1669 (2011).

[22] Ji, F., Yang, B., Peng, X., Ding, H., You, H. and Tien, P.: Circulating microRNAs in hepatitis B virus-infected patients, *J. Viral Hepat.*, Vol. 18, No. 7, pp. e242–251 (2011).

[23] Starkey Lewis, P. J., Dear, J., Platt, V., Simpson, K. J., Craig, D. G., Antoine, D. J., French, N. S., Dhaun, N., Webb, D. J., Costello, E. M., Neoptolemos, J. P., Moggs, J., Goldring, C. E. and Park, B. K.: Circulating microRNAs as potential markers of human drug-induced liver injury, *Hepatology*, Vol. 54, No. 5, pp. 1767–1776 (2011).

[24] Murakami, Y., Toyoda, H., Tanahashi, T., Tanaka, J., Kumada, T., Yoshioka, Y., Kosaka, N., Ochiya, T. and Taguchi, Y.-h.: Comprehensive miRNA Expression Analysis in Peripheral Blood Can Diagnose Liver Disease, *PLoS ONE*, Vol. 7, No. 10, p. e48366 (online), DOI: 10.1371/journal.pone.0048366 (2012).

[25] R Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2012). ISBN 3-900051-07-0.

[26] Taguchi, Y.-h. and Murakami, Y.: Refined blood-borne miRNome of human diseases via PCA-based feature extraction, *IPSJ SIG Technical Report*, Vol. 2012, No. 13, pp. 1–6 (online), available from ⟨http://ci.nii.ac.jp/naid/110008803118/⟩ (2012).