PLOS | ONE

# Distinct MicroRNAs Expression Profile in Primary Biliary Cirrhosis and Evaluation of miR 505-3p and miR197-3p as Novel Biomarkers

**Masashi Ninomiya[1], Yasuteru Kondo[1]*, Ryo Funayama[2], Takeshi Nagashima[2], Takayuki Kogure[1], Eiji Kakazu[1], Osamu Kimura[1], Yoshiyuki Ueno[3], Keiko Nakayama[2], Tooru Shimosegawa[1]**

**1** Division of Gastroenterology, Tohoku University Hospital, Sendai, Japan, **2** Division of Cell Proliferation, Tohoku University School of Medicine, Sendai, Japan, **3** Department of Gastroenterology, Yamagata University Faculty of Medicine, Yamagata, Japan

## Abstract

***Background and Aims:*** MicroRNAs are small endogenous RNA molecules with specific expression patterns that can serve as biomarkers for numerous diseases. However, little is known about the expression profile of serum miRNAs in PBC.

***Methods:*** First, we employed Illumina deep sequencing for the initial screening to indicate the read numbers of miRNA expression in 10 PBC, 5 CH-C, 5 CH-B patients and 5 healthy controls. Comparing the differentially expressed miRNAs in the 4 groups, analysis of variance was performed on the number of sequence reads to evaluate the statistical significance. Hierarchical clustering was performed using an R platform and we have found candidates for specific miRNAs in the PBC patients. Second, a quantitative reverse transcription PCR validation study was conducted in 10 samples in each group. The expression levels of the selected miRNAs were presented as fold-changes ($2^{-\Delta\Delta Ct}$). Finally, computer analysis was conducted to predict target genes and biological functions with MiRror 2.0 and DAVID v6.7.

***Results:*** We obtained about 12 million 32-mer short RNA reads on average per sample and the mapping rates to miRBase were 16.60% and 81.66% to hg19. In the statistical significance testing, the expression levels of 81 miRNAs were found to be differentially expressed in the 4 groups. The heat map and hierarchical clustering demonstrated that the miRNA profiles from PBC clustered with those of CH-B, CH-C and healthy controls. Additionally, the circulating levels of hsa-miR-505-3p, 197-3p, and 500a-3p were significantly decreased in PBC compared with healthy controls and the expression levels of hsa-miR-505-3p, 139-5p and 197-3p were significantly reduced compared with the viral hepatitis group.

***Conclusions:*** Our results indicate that sera from patients with PBC have a unique miRNA expression profile and that the down-regulated expression of hsa-miR-505-3p and miR-197-3p can serve as clinical biomarkers of PBC.

**Citation:** Ninomiya M, Kondo Y, Funayama R, Nagashima T, Kogure T, et al. (2013) Distinct MicroRNAs Expression Profile in Primary Biliary Cirrhosis and Evaluation of miR 505-3p and miR197-3p as Novel Biomarkers. PLoS ONE 8(6): e66086. doi:10.1371/journal.pone.0066086

**Editor:** Aftab A. Ansari, Emory University School of Medicine, United States of America

**Received** November 28, 2012; **Accepted** May 3, 2013; **Published** June 12, 2013



**Funding:** Financial support for this study was provided by Grant-in-Aid for Young Scientists (B) (23790765). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* E-mail: yasuteru@ebony.plala.or.jp

## Introduction

MicroRNAs (miRNAs) are small endogenous RNA molecules of 19 to 24 nucleotides that control the translation and transcription of targeting mRNAs by base-pairing to the complementary sites [1] [2] [3] [4]. The expression of miRNAs in serum is reported to be stable, reproducible and consistent among individuals of the same species [5]. So far, specific expression patterns of serum miRNAs were identified as a fingerprint for numerous diseases and cancers [6] [5]. The serum miR-122 levels are elevated in patients with liver damage due to chronic hepatitis B (CH-B) and C infection (CH-C) [7] [8]. In addition, the miR-122 and miR-34a levels are positively correlated with the disease severity in CH-C and non-alcoholic fatty-liver disease [9]. However, there are some reports that miR-122 expression in healthy controls was significantly higher than that in patients with hepatitis C virus (HCV) infection [10]. Li et. al. described that 13 miRNAs were differentially expressed in hepatitis B virus (HBV) serum and that miR-25, miR-375 and let-7f could be used as biomarkers to separate a HBV-positive hepatocellular carcinoma (HCC) group from HBV-negative HCC [11]. However, little is known about the expression profile of miRNAs in autoimmune disease such as primary biliary cirrhosis (PBC).

PBC is female predominant, progressive autoimmune disease characterized by immune-mediated destruction of the intrahepatic bile ducts. The serological marker of PBC is the presence of anti-mitochondrial antibody (AMA) directed against the E2 subunit of the pyruvate dehydrogenase enzyme complexes located in the inner mitochondrial membrane [12] [13] [14]. The etiology of PBC is considered to be a combination of genetic predisposition and environmental triggers [15]. Particularly, concerning genetic predisposition, previous studies reported that common genetic variants at the HLA class II, IL12A, IL12RB2, STAT4, IRF5-

TNPO3 and IKZF3 had significant associations [16] [17] [18] [19] [20] [21]. Recently, genome-wide association study in Japanese population showed TNFSF15 and POU2AF1 as susceptibility loci [22]. Several GWAS data suggested the important contributions of several immune pathways to the development of PBC. However, the results have differed among the study groups [21]. The diagnosis of PBC is established based on the following criteria: (1) biochemical evidence of cholestasis; (2) the presence of AMA; and (3) histopathologic evidence of nonsuppurative cholangitis and destruction of the interlobular bile ducts [23]. Though diagnostic criteria have been determined, the eventual progression to biochemically and clinically apparent disease is unpredictable. Many patients are recognized at an earlier stage of disease and respond well to medical therapy, while some patients will require liver transplantation [24] [25]. To revolutionize the diagnosis, treatment and prognosis of PBC, new biomarkers seem to be feasible, and miRNAs are emerging as highly tissue-specific biomarkers with potential clinical applicability [26] [6].

In this study, we employed a strategy of using Illumina small-RNA sequencing for the initial screening followed by quantitative reverse transcription PCR (qRT-PCR) validation to analyze serum samples, which were arranged in multiple trial and testing sets. Additionally, computer analysis was conducted to predict target genes and biological functions from the differentially expressed miRNAs in PBC. The results demonstrate that the unique expression pattern of serum miRNAs can serve as a noninvasive biomarker for the diagnosis of PBC.

## Results

### Global analysis of miRNAs by deep sequencing

Circulating miRNAs were detected from human serum in 10 patients diagnosed with PBC and 15 non-PBC subjects with CH-B, CH-C and healthy controls, using Illumina GA IIx sequencing. Detailed clinical information is shown in table 1. We analyzed three samples by single-end deep sequencing on one lane and obtained about 12 million 32-mer short RNA reads on average per sample. After trimming the reads to exclude adaptor and tag sequences by using cutadapt, 9,245,752 high-quality reads per sample were subjected to analysis [27]. The mapping rates to miRBase were 16.60% on average and those to hg19 were 81.66% (Table 2).

### miRNA expression profile in serum affected of PBC

We normalized the differential expression of miRNA count data using the trimmed mean of M values (TMM) normalization process and the number of individual miRNA reads was standardized by the total numbers of 1,000,000 reads in each sample [28]. Comparing the 4 groups (PBC, CH-C, CH-B and healthy control), the differential expression levels of miRNA were extracted using analysis of variance (ANOVA). 1594 miRNAs were detected by deep sequencing (Table S1). Due to the small number of miRNA detections, the statistical significance of 821 miRNAs could not be determined. The ANOVA was used to determine differentially expressed miRNAs and multiple comparisons procedure was applied to compare more than one pair of means at the same time. Therefore, in statistical significance testing in the remaining 773 miRNAs, the p-value by multiple comparisons was performed by calculating False discovery rate (FDR)$<0.1$. The expression levels of 81 miRNAs were found to be differentially expressed in the 4 groups. Although, several types of clinical data were shown to be different from PBC, M2 negative for PBC-8, ANA positive for PBC-10 or past HBV infection for PBC-6, the histologies of all samples were characterized by PBC of chronic, nonsuppurative cholangitis affecting the interlobular and septal ducts. The heat map and hierarchical clustering demonstrated that the miRNA profiles from PBC clustered with those of CH-B, CH-C and healthy controls (Figure 1). Of note, CH-B and healthy controls were not clearly distinguished.

In the result of Illumina sequencing, 3 miRNAs were up-regulated (>2-fold) in PBC patients, with hsa-miR-1273g-5p being most enriched. The relative levels of 6 dysregulated miRNAs were down-regulated (<0.5-fold), with hsa-miR-766-5p being the smallest (Table 3).

### miRNA validation study

We used qRT-PCR to verify the data obtained from the Illumina sequencing. The relative expression levels from 9 differentially expressed miRNAs were analyzed with the TaqMan MicroRNA assay (Applied Biosystems) or miRCURY LNA microRNA PCR system (Exiqon). In addition to the 25 samples used for Illumina deep sequencing, 10 cases of miRNA expression in PBC, CH-C, CH-B and healthy controls were quantified (Table 4). The expression levels of selected miRNAs detected by qRT-PCR were normalized to spiked-in cel-miR-39 and presented as fold-changes ($2^{-\Delta\Delta Ct}$) above those of the CH-C-5. Among 9 miRNAs, the quantity of four miRNA expressions could be determined. The circulating levels of hsa-miR-505-3p, miR-197-3p and miR-500a-3p ($p<0.01$) were significantly decreased in PBC compared with healthy controls and the expression levels of hsa-miR-505-3p, miR-139-5p and miR-197-3p were significantly reduced compared with the viral hepatitis group (Figure 2). The quantities of miRNA expression by qRT-PCR were supported by the data of Illumina sequencing. Of note, for samples with only trace amounts, quantification of the reaction products after 30 cycles is uncertain. The quantity of 5 remaining miRNA expressions indicated the Ct values under 30 cycles even in the higher amount group or were undetected.

### miRNA target genes and Gene ontology (GO) analysis

To determine possible target genes for the differentially expressed miRNAs, we searched the miRror 2.0 database for predicted miRNA targets in human. The ranking of miRror was according to the miRror Internal Score (miRIS). This score is a balance between (i) the proportions of predicting miRNA-target prediction databases (MDBs) out of all tested MDBs and (ii) the fraction of the potentially regulated genes from the entire input genes [29]. A total of 75 genes were predicted as targets for 9 differentially expressed miRNAs by Illumina sequencing of hsa-miR-1273g-5p, miR-33a-5p, miR-3960, miR-766-5p, miR-505, miR-30b, miR-139-5p, miR-197 and miR-500a in PBC serum (Table 5). Then, these 75 genes were submitted to Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7, which was used for the GO biological process categorization, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and BIOCARTA pathway. The predicted target genes involved biological processes such as blood circulation (5.4%), circulatory system process (5.4%), positive regulation of I-kappaB kinase/NF-kappaB cascade (4.1%), regulation of phosphoprotein phosphatase activity (2.7%), cell volume homeostasis (2.7%), regulation of phosphatase activity (2.7%) and cellular amino acid derivative catabolic process (2.7%), some cellular components such as cytosol (13.5%), cell fraction (12.2%), membrane fraction (9.5%) and protein serine/threonine phosphatase complex (4.1%), and some molecular functions such as ion binding (33.8%), Cation binding (33.8%), metal ion binding (33.8%) and transition metal ion binding (25.7%) (Table S2). The functional annotation analysis of

**Table 1.** Clinical data for patients enrolled in Illumina sequencing analysis.

| Patient | Sex | Age | Biopsy finding[a] | T-bil (mg/dl) | ALT (IU/l) | ALP (IU/l) | Albumin (g/dl) | PT-INR | ANA | AMA | M2 (index) | HBV-DNA (Log copies/ml) | HCV-RNA (LogIU/ml) | Past treatment[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PBC-1 | M | 60 | II | 0.5 | 36 | 404 | 3.8 | 1.00 | - | 1:80 | 128 | | | URSO (−) |
| PBC-2 | M | 65 | I | 1.2 | 54 | 485 | 4.1 | 1.05 | - | 1:160 | 125 | | | URSO (−) |
| PBC-3 | F | 61 | II | 1.2 | 38 | 478 | 4.0 | 0.97 | - | 1:20 | 89.5 | | | URSO (−) |
| PBC-4 | F | 51 | III | 0.8 | 71 | 798 | 3.8 | 0.96 | - | - | 47.9 | | | URSO (−) |
| PBC-5 | F | 62 | II | 0.6 | 26 | 407 | 4.0 | 0.99 | - | 1:20 | 102 | | | URSO (−) |
| PBC-6 | F | 61 | I | 0.6 | 38 | 527 | 4.2 | 0.91 | - | - | 111.4 | | | URSO (−) |
| PBC-7 | F | 55 | II | 1.0 | 53 | 800 | 4.2 | 1.02 | 1:80 | - | 89.9 | | | URSO (−) |
| PBC-8 | F | 52 | II | 0.9 | 28 | 525 | 3.7 | 0.94 | 1:80 | - | - | | | URSO (−) |
| PBC-9 | F | 72 | I | 1.1 | 42 | 666 | 3.4 | 1.00 | - | 1:80 | 143.9 | | | URSO (−) |
| PBC-10 | F | 64 | I | 0.7 | 38 | 513 | 4.0 | 0.88 | >2561 | 1:80 | 79.4 | | | URSO (−) |
| CH-C-1 | M | 57 | 2 | 0.6 | 22 | 240 | 4.4 | 0.91 | | | | | 5.4 | IFN (+) |
| CH-C-2 | F | 55 | 3 | 1.5 | 73 | 403 | 4.0 | 1.06 | | | | | 5.1 | IFN (+) |
| CH-C-3 | F | 61 | 1 | 1.1 | 12 | 179 | 4.2 | 1.11 | | | | | 5.2 | IFN (+) |
| CH-C-4 | F | 49 | 2 | 0.8 | 42 | 261 | 3.6 | 1.05 | | | | | 6.3 | IFN (−) |
| CH-C-5 | F | 53 | 2 | 0.8 | 27 | 257 | 4.0 | 0.93 | | | | | 6.9 | IFN (+) |
| CH-B-1 | F | 35 | 1 | 0.7 | 585 | 283 | 4.1 | 1.03 | | | | 9.1 | | NA (−) |
| CH-B-2 | F | 72 | 2 | 0.9 | 64 | 211 | 3.8 | 1.03 | | | | 3.4 | | NA (−) |
| CH-B-3 | F | 38 | 2 | 0.8 | 283 | 268 | 3.1 | 0.99 | | | | 8.8 | | NA (−) |
| CH-B-4 | F | 43 | 3 | 0.6 | 87 | 239 | 3.7 | 1.01 | | | | 7.4 | | NA (−) |
| CH-B-5 | M | 47 | 1 | 0.6 | 40 | 156 | 4.1 | 1.12 | | | | 5.9 | | NA (−) |
| Healthy-3 | M | 37 | | 0.6 | 24 | 252 | 4.2 | 1.02 | | | | | | |
| Healthy-4 | F | 49 | | 0.7 | 21 | 274 | 4.1 | 1.08 | | | | | | |
| Healthy-5 | F | 33 | | 0.6 | 18 | 173 | 3.9 | 1.12 | | | | | | |
| Healthy-7 | M | 26 | | 0.8 | 35 | 218 | 4.3 | 1.08 | | | | | | |
| Healthy-8 | M | 28 | | 0.9 | 22 | 238 | 4.2 | 1.04 | | | | | | |

[a]We used the Scheuer score in PBC and fibrosis score of histological activity index (HAI) in CH-B and CH-C [53] [54] [55].
[b]URSO is the abbreviation for ursodeoxycholic acid, IFN for interferon and NA for nucleos(t)ide analogue.
doi:10.1371/journal.pone.0066086.t001

**Table 2.** The number of small RNAs in serum detected by Illumina sequencing.

| | | Cutadapt | | Mapping (miRBase) | | Mapping (hg19) | |
|---|---|---|---|---|---|---|---|
| Sample | Total | No. of read | % read | No. of read | % read | No. of read | % read |
| PBC-1 | 9,996,912 | 5,912,672 | 59.14 | 582,682 | 9.85 | 5,195,131 | 87.86 |
| PBC-2 | 17,103,184 | 8,147,153 | 47.64 | 837,575 | 10.28 | 6,484,515 | 79.59 |
| PBC-3 | 11,731,105 | 9,001,730 | 76.73 | 873,985 | 9.71 | 5,414,933 | 60.15 |
| PBC-4 | 12,785,162 | 10,700,147 | 83.69 | 1,463,341 | 13.68 | 8,227,422 | 76.89 |
| PBC-5 | 14,732,479 | 9,138,080 | 62.03 | 1,125,157 | 12.31 | 7,317,178 | 80.07 |
| PBC-6 | 12,139,379 | 7,256,738 | 59.78 | 1,274,462 | 17.56 | 5,612,279 | 77.34 |
| PBC-7 | 12,895,734 | 10,107,477 | 78.38 | 1,067,578 | 10.56 | 7,826,537 | 77.43 |
| PBC-8 | 18,786,941 | 11,711,844 | 62.34 | 967,802 | 8.26 | 8,323,607 | 71.07 |
| PBC-9 | 11,852,431 | 8,980,873 | 75.77 | 1,141,912 | 12.71 | 7,593,581 | 84.55 |
| PBC-10 | 18,224,562 | 12,752,337 | 69.97 | 756,930 | 5.94 | 8,912,212 | 69.89 |
| CH-C-1 | 10,874,814 | 7,820,291 | 71.91 | 2,870,316 | 36.70 | 6,924,685 | 88.55 |
| CH-C-2 | 10,242,500 | 8,138,815 | 79.46 | 3,250,159 | 39.93 | 6,976,754 | 85.72 |
| CH-C-3 | 19,183,649 | 12,135,107 | 63.26 | 1,878,487 | 15.48 | 9,056,579 | 74.63 |
| CH-C-4 | 18,750,568 | 15,952,136 | 85.08 | 1,877,654 | 11.77 | 14,829,719 | 92.96 |
| CH-C-5 | 12,702,304 | 10,306,186 | 81.14 | 1,592,422 | 15.45 | 9,269,394 | 89.94 |
| CH-B-1 | 5,861,013 | 4,732,185 | 80.74 | 1,126,751 | 23.81 | 3,708,378 | 78.37 |
| CH-B-2 | 7,164,871 | 5,937,732 | 82.87 | 1,213,392 | 20.44 | 5,076,436 | 85.49 |
| CH-B-3 | 7,029,349 | 6,357,274 | 90.44 | 858,562 | 13.51 | 5,378,310 | 84.60 |
| CH-B-4 | 8,077,025 | 6,788,987 | 84.05 | 1,581,475 | 23.29 | 6,142,836 | 90.48 |
| CH-B-5 | 10,255,895 | 9,101,104 | 88.74 | 1,952,400 | 21.45 | 7,687,217 | 84.46 |
| Healthy-3 | 11,111,254 | 7,843,011 | 70.59 | 1,248,499 | 15.92 | 6,778,655 | 86.43 |
| Healthy-4 | 12,449,813 | 10,690,833 | 85.87 | 2,410,128 | 22.54 | 9,204,163 | 86.09 |
| Healthy-5 | 13,597,339 | 8,914,365 | 65.56 | 2,214,635 | 24.84 | 6,571,819 | 73.72 |
| Healthy-7 | 11,351,494 | 9,825,442 | 86.56 | 1,880,266 | 19.14 | 8,618,696 | 87.72 |
| Healthy-8 | 14,349,933 | 12,891,294 | 89.84 | 2,320,100 | 18.00 | 11,618,231 | 90.12 |
| Total | 313,249,710 | 231,143,813 | 73.79 | 38,366,670 | 16.60 | 188,749,267 | 81.66 |

doi:10.1371/journal.pone.0066086.t002

BIOCARTA showed that the genes of catenin (cadherin-associated protein), alpha 1 and similar to breast cancer anti-estrogen resistance 1 predicted target genes of the listed miRNAs, and played a role in cell-to-cell adhesion signaling pathway (Figure S1). The KEGG pathway indicated that the genes of baculoviral IAP repeat-containing 2, protein phosphatase 3 catalytic subunit beta isoform and tumor necrosis factor ligand superfamily member 10 were related to apoptosis (Figure S2).

## Discussion

MiRNA changes in the liver have been reported in diseases such as HCC or chronic viral hepatitis. However, there is only limited information about their detection in blood and their correlations in PBC patients. The current study provides the first evidence that PBC is associated with altered miRNA expression. We have demonstrated that a number of miRNAs, especially hsa-miR-505-3p and miR-197-3p, were significantly differentially expressed in patients with PBC, leading to a unique miRNA expression profile in the diseased liver. Recently, many studies have examined several PBC associations with genes and there have been significant differences in the genetic risk loci reported [21] [30]. Therefore more carefully constructed studies will be needed to clarify, the pathogenesis of PBC, and the study of these differentially expressed miRNAs could serve in identifying biomarkers or lead to a better understanding of the underlying molecular mechanism that perpetuates PBC.

In our study, miRNA Illumina deep sequencing was first used to screen 10 PBC patients' sera. We were able to match the sample's sex because of particular importance for X-linked miRNA [31]. Then, qRT-PCR was used to confirm the result of deep sequencing.

Quantitative differential expression analysis identified a 81-miRNA signature distinguishing PBC, CH-C, CH-B and healthy controls. A hierarchical clustering analysis was performed utilizing the 81-miRNAs and their patterns separated the PBC from viral hepatitis and healthy controls. In addition, there were three subgroups (PBC-1,2,4,5, PBC-3,6,10 and PBC-7,8,9) in the PBC cluster. When comparing the subgroups, the PBC-3, 6,10 group showed an expression pattern that differed from those of the other two subgroups. As for the clinical background, PBC-6 patient that had been infected with hepatitis B virus were HBsAg negative and anti-HBc and anti-HBs positive, and PBC-10 patient had positive antinuclear antibodies (ANA) titers of 1:2560 or greater, while other patients had no serological evidence of HBV infection and no positive ANA titers of 1:160 or greater. However, there was no clear difference between the three subgroups in terms of the clinical stage. At present, there is no conclusive proof whether

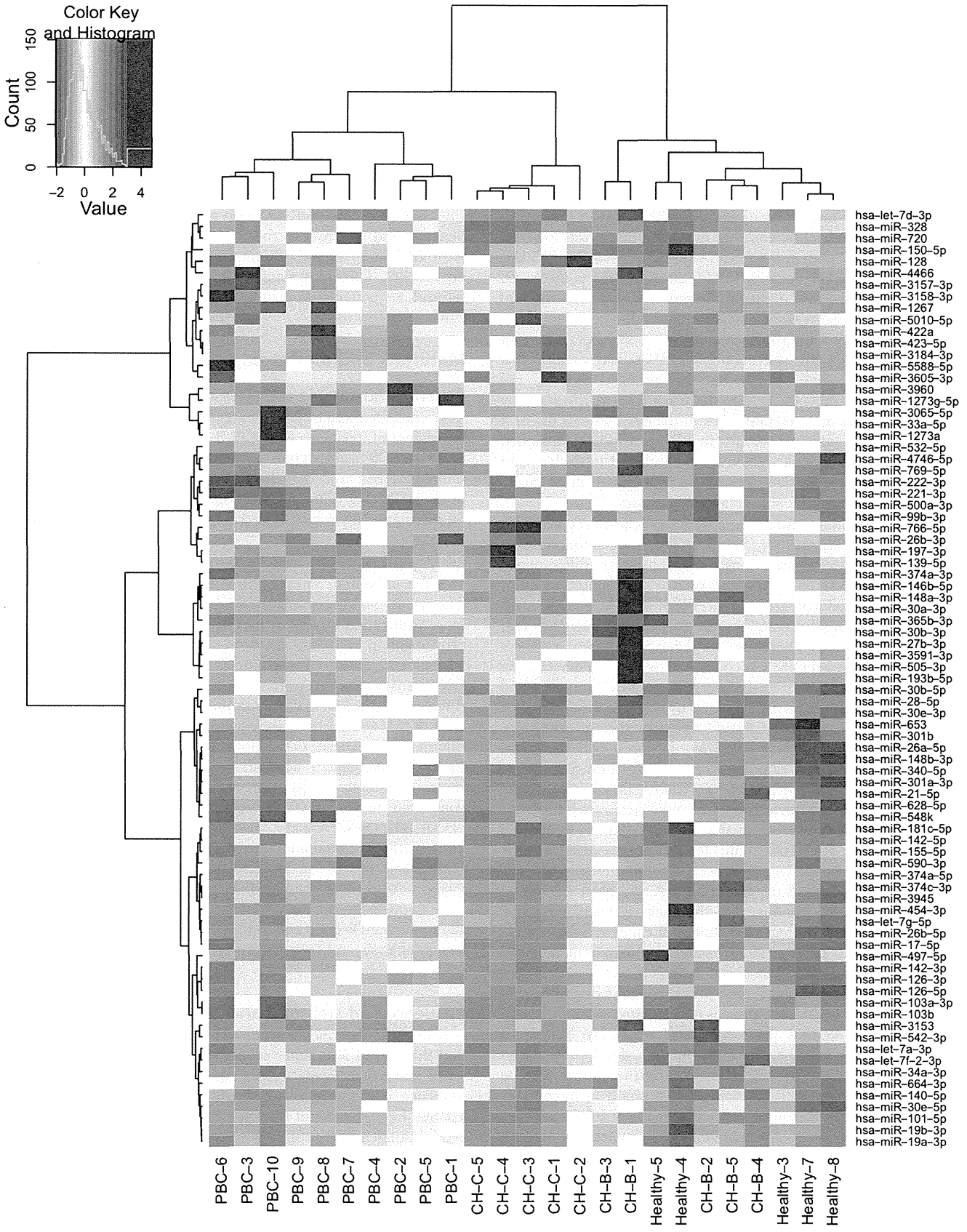


**Figure 1. Heat map and hierarchical clustering.** Individual miRNA expression were calculated by R platform and heat map was computed and described using a function of heatmap.2 in gplots. It uses hierarchical clustering with Euclidean distance; Pearson Linear Correlation and Ward's method to generate the hierarchical tree [56]. ANOVA was applied to extract differentially expressed miRNAs and adjustment of the p-value by

multiple comparisons was performed by calculating FDR. Those miRNAs with FDR<0.1 were presented. The red indicates high level of miRNA expression and the blue shows low.
doi:10.1371/journal.pone.0066086.g001

clinical or pathological differences can be found in clustering subgroups.

Nine miRNAs were confirmed to be significantly differentially expressed between the PBC group and viral hepatitis group or healthy control by Illumina deep sequencing. Among these 9 miRNAs, the serum levels of hsa-miR-505-3p and miR-197-3p were significantly lower in patients with PBC than in those with viral hepatitis and healthy controls, hsa-miR-139-5p was lower in patients with PBC than in those with viral hepatitis and miR-500a-3p were lower in patients with PBC than in healthy controls. Of note, we conducted qRT-PCR on the sera of some PBC who had already been treated with ursodeoxycholic acid. The serum levels of miRNA showed improvements in some samples (data not shown). Accordingly, quantifying these miRNAs may yield reliable diagnostic information. However, one problem is that the quantification of miRNA in this study used standardization by the total numbers of 1,000,000 reads in the deep sequencing or a comparative method in qRT-PCR. In other words, it was assumed that the same amount of miRNA was contained in each serum sample. Therefore, if one miRNA is quantified in a single specimen, we will not be able to accurately assess the result.

Specific circulating miRNA profiles have been reported for various diseases [32] [5] [33] [34] [35]. These circulating miRNA profiles have been described as correlating with differentially expressed miRNA in diseased tissue, such as liver injured by drugs or stomach afflicted with gastric cancer [36] [37]. Moreover, some disease-specific profiles can inform both the diagnosis and prognosis [38] [39]. Therefore, to determine if any of these diferentially expressed miRNAs could lead to better understanding of the molecular mechanism that perpetuates PBC, we examined for gene targets that may be reflected by this particular miRNA expression signature. Of the several target prediction algorithms prepared, we selected mirror 2.0. There has been evidence that a seed region of miRNA positioned within a limited range in the 3′ UTR of a target gene degrades the mRNA function [40]. We predicted 75 genes as targets for 9 differentially expressed miRNAs and conducted a functional analysis of DAVID. This analysis revealed that the genes of catenin (cadherin-associated protein), alpha 1 and similar to breast cancer anti-estrogen resistance 1 predicted target genes of the listed miRNA and played a role in cell-to-cell adhesion signaling, and the genes of baculoviral IAP repeat-containing 2, protein phosphatase 3 catalytic subunit beta isoform and tumor necrosis factor ligand superfamily member 10 were related to apoptosis. The onset of autoimmune disorders with PBC can be linked to apoptosis. A previous report described that the expression of TRAIL receptors is up-regulated by an increased bile acid level and that the serum level of soluble TRAIL is elevated, which may be involved in the development and progression of PBC [41,42]. However, further work will be required so that these miRNAs can serve not only as biomarkers but also for the elucidation of the pathogenesis of PBC.

GO analysis provides representations of biological annotations using precisely defined terms [43]. A previous report has described a number of genes involved in the signaling, regulation of I-kappaB kinase/NF-kappaB cascade and homeostasis that are associated with PBC [44] [45,46]. Additionally, our study indicated the biological processes, cellular component and molecular functions affected by the target genes included those associated with cell or membrane fraction, various kinds of ion binding and protein serine/threonine phosphatase complex, all of which are potentially related to PBC. Further studies will need to examine the relationship between differentially expressed miRNAs, both in serum and liver tissue, and target genes, which may provide more insights into the role of miRNAs in the pathology of PBC.

In conclusion, our results indicate that sera from patients with PBC have a unique miRNA expression profile compared to viral hepatitis and healthy controls and down-regulated expression of hsa-miR-505-3p and 197-3p may represent new clinical biomarkers in PBC. This study suggests that the amounts of miRNAs in serum have potential as diagnostic and prognostic biomarkers for PBC.

## Materials and Methods

### Patients and sample processing

We included sera of 10 patients with PBC who were treatment-naïve, sera of 5 patients with CH-B, sera of 5 patients with CH-C and sera of 5 healthy controls in this study. Initially these serum samples were enrolled to be analyzed by the Illumina miRNA deep sequencing (Illumina). The diagnosis of all cases was based on internationally established criteria [23].

### Library preparation and Illumina sequencing

A ten ml venous sample was collected from each participant. The whole blood was separated into serum and cellular fractions by centrifugation at 2,500 r.p.m. for 10 min, followed by 10 min centrifugation at 10,000 r.p.m. to completely remove cell debris. The supernatant serum was stored at −20°C until analysis. Total RNA was extracted from 800 µl of serum using Trizol LS (Invitrogen, Carlsbad, CA). The libraries were constructed from total RNA using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, CA) following the manufacturer's protocol. Briefly, RNA 3′ and 5′ adapters were ligated to target microRNAs in two separate steps. Reverse transcription reaction was conducted to the ligation products to create single stranded cDNA. The cDNA was amplified by PCR using a common primer and a primer containing the index sequence. One µl of each library was loaded on an Agilent Bioanalyzer (Agilent, Santa Clara, CA) to check the size, purity, and concentration. Libraries were sequenced on an Illumina GA IIx (SCS 2.8 software; Illumina, SanDiego, CA), with a 32-mer single end sequence. Image analysis and base calling were performed using RTA 1.8 software.

### Sequence and statistical analysis

Raw miRNA sequence reads were conducted as a quality check and the 3′ and 5′ adapter sequences were removed by cutadapt while discarding reads shorter than 20 nucleotides [27]. The sequence reads were mapped with miRBase (Release 18) and UCSC (hg19) by use of bwa (0.5.9-r16), allowing one nucleotide base mismatch [47] [48].

Digital expression levels were normalized by taking into account the length of miRNAs and the total number of miRNA reads generated in each library using TMM normalization [28]. Read counts of each identified miRNA was normalized to the total number of miRNA reads, and then the ratio was multiplied by a constant set to $1\times10^6$ in this study. ANOVA was applied to extract differentially expressed miRNAs among the four groups. Adjustment of the p-value by multiple comparisons was performed by

**Table 3.** Differentially expressed miRNAs in serum from PBC patients compared with the second group (CH-C, CH-B, Healthy).

| miRNA | Expression | PBC | CH-C | CH-B | Healthy | Fold change | p-value |
|---|---|---|---|---|---|---|---|
| | | The mean no. of reads ±SE | The mean no. of reads ±SE | The mean no. of reads ±SE | The mean no. of reads ±SE | | |
| hsa-miR-1273g-5p | Up | 6.79±0.50 | 1.88±0.26 | 0.51±0.10 | 1.21±0.08 | 3.61 | 9.93E-03 |
| hsa-miR-33a-5p | Up | 6.19±1.59 | 0.10±0.04 | 1.58±0.23 | 2.20±0.18 | 2.82 | 4.07E-03 |
| hsa-miR-3960 | Up | 11.26±0.59 | 4.85±0.51 | 3.63±0.27 | 1.86±0.24 | 2.32 | 4.27E-03 |
| hsa-miR-766-5p | Down | 0.17±0.04 | 2.92±0.53 | 1.55±0.10 | 0.64±0.12 | 0.27 | 4.61E-03 |
| hsa-miR-505-3p | Down | 5.05±0.22 | 16.23±0.89 | 26.81±3.99 | 16.73±1.54 | 0.31 | 3.40E-03 |
| hsa-miR-30b-3p | Down | 0.41±0.08 | 3.76±0.38 | 8.77±1.00 | 1.30±0.26 | 0.31 | 1.01E-03 |
| hsa-miR-139-5p | Down | 19.73±0.77 | 77.72±9.44 | 61.29±6.57 | 82.86±6.06 | 0.32 | 6.86E-03 |
| hsa-miR-197-3p | Down | 226.99±10.32 | 1067.05±106.41 | 589.88±60.38 | 823.16±66.17 | 0.38 | 7.76E-03 |
| hsa-miR-500a-3p | Down | 36.01±1.66 | 74.61±1.95 | 86.52±5.35 | 99.59±3.00 | 0.48 | 2.29E-03 |

doi:10.1371/journal.pone.0066086.t003

calculating FDR [49]. Those miRNAs with FDR<0.1 were extracted as differentially expressed and used in the following analysis. Hierarchical clustering was performed using an R platform and a heat map described as using a function of heatmap.2 in gplots [50].

## qRT-PCR validation study

In addition to 25 samples analyzed by Illumina sequencing, five more serum samples of CH-B, CH-C and healthy controls (a total of 10 samples in each group) were used in qRT-PCR validation study. We followed the protocol previously reported by Mitchell *et al.* to determine the endogenous miRNA levels with spiked-in miRNA. Spiked-in miRNA was designed against *Caenorhabditis elegans* microRNA-39 (cel-miR-39) (5′- UCA CCG GGU GUA AAU CAG CUU -3′) and was synthesized by Sigma Aldrich Japan [32]. After total RNA isolation from 300 µl serum, reverse transcription was conducted using a TaqMan miRNA RT kit for identification of the cel-miR-39 expression (Applied Biosystems) with 5 fmol/µl for the internal control. qRT-PCR were conducted for detection of hsa-miR-1273g-5p, miR-505-3p and miR-139-5p in 20 µl PCR reactions using TaqMan MicroRNA assay with StepOne Plus detection system at 50°C for 2 min and 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min (Applied Biosystems). For detection of hsa-miR-33a-5p, miR-3960, miR-766-5p, miR-30b-3p, miR-197-3p and miR-500a-3p expression, we used the Exiqon system. Total RNA was reverse transcribed using the miRCURY LNA$^{TM}$ Universal RT miRNA PCR, Polyadenylation and cDNA synthesis kit (Exiqon). cDNA diluted 50× was assayed in 10 µl PCR reactions according to the protocol for miRCURY LNA$^{TM}$ Universal RT miRNA PCR with StepOne Plus detection system at 95°C for 10 min, followed by 40 cycles of 95°C for 10 s and 60°C for 1 min (Exiqon). The data were analyzed by the $2^{-\Delta\Delta Ct}$ method.

## Statistical methods

Expression levels of the selected miRNAs detected by qRT-PCR were normalized to cel-miR-39 and presented as the fold-change ($2^{-\Delta\Delta Ct}$) above the control (CH-C-5): $\Delta\Delta Ct = (Ct_{miRNA}\text{-}Ct_{cel\text{-}miR\text{-}39})_{patients}\text{-}(Ct_{miRNA}\text{-}Ct_{cel\text{-}miR\text{-}39})_{CH\text{-}C\text{-}5}$. Results for normally distributed continuous variables are given as means (±standard errors of the mean) and compared between groups by Student's t-test. Results for non-normally distributed continuous variables are summarized as medians (interquartile ranges) and were compared by Mann-Whitney U test.

## *In silico* analysis of miRNA target gene

For computational prediction of miRNA target genes, we used an algorithm: miRror 2.0 (June 2010 release, http://www.proto.cs.huji.ac.il/mirror/) [51]. MiRror 2.0 encompasses most of the available miRNA-target prediction tools covering human miRNAs. The algorithms used are collectively called miRNA-target prediction databases (MDBs): (i) PITA (Kartez); (ii) PicTar 4 (Krek); (iii) TargetRank (Nielsen); (iv) TargetScan (Lewis); (v) microCosm (John); (vi) miRanda (Betel); (vii) DIANA-microT (Maragkakis); (viii) MirZ (Hausser); (ix) miRDB (Wang); (x) RNA 22 (Miranda); (xi) MAMI (Sethupathy); (xii) miRNAMap2 (Hsu). The number of candidate genes and the number of miRNAs are indicated for each of the major MDBs. We selected as the search mode: miR2Gene; and as the search parameters of organism: human; and of selected tissue: all. Advanced parameters were inputted, cutoff: 0.01; database hits: 2; and target counts: 3. We created a list of common target genes for miRNAs. Then, these common targets were annotated by an annotation tool at the DAVID v6.7 (January 2010 release, http://david.abcc.ncifcrf.

**Table 4.** Clinical information of patients enrolled in the validation study.

| | PBC (n = 10) | CH-C (n = 10) | CH-B (n = 10) | Healthy (n = 10) |
|---|---|---|---|---|
| Male/Female | 2/8 | 4/6 | 6/4 | 6/4 |
| Age range | 51–72 | 47–70 | 27–72 | 26–62 |
| Histological findings[a] | | | | |
| Scheuer score | I (4) II (5) III (1) | | | |
| HAI | | 1 (1) 2 (8) 3 (1) | 1 (3) 2 (4) 3 (3) | |
| T-bil (mg/dl)[b] | 0.9 (0.5–1.2) | 0.9 (0.6–1.5) | 0.8 (0.5–1.1) | 0.8 (0.6–1.0) |
| ALT (IU/l)[b] | 42.4 (26–71) | 47.1 (12–193) | 139.1 (13–585) | 25.8 (18–38) |
| ALP (IU/l)[b] | 560.3 (404–800) | 238.7 (167–403) | 221.7 (111–319) | 230.4 (148–302) |
| Albumin (g/dl)[b] | 3.9 (3.4–4.2) | 4.0 (3.5–4.2) | 3.9 (3.7–4.4) | 4 (3.8–4.3) |
| PT-INR[b] | 0.97 (0.88–1.05) | 1.02 (0.91–1.11) | 1.07 (0.99–1.18) | 1.07 (1.02–1.13) |
| AMA positivity | 6 | | | |
| M2 positivity | 9 | | | |
| HBV-DNA (Log copies/ml)[b] | | | 6.2 (3.4–9.1) | |
| HCV-RNA (LogIU/ml)[b] | | 6.4 (5.1–7.3) | | |

[a]The numbers of patients are indicated in the parentheses.
[b]The range of laboratory data is indicated in the parentheses.
doi:10.1371/journal.pone.0066086.t004

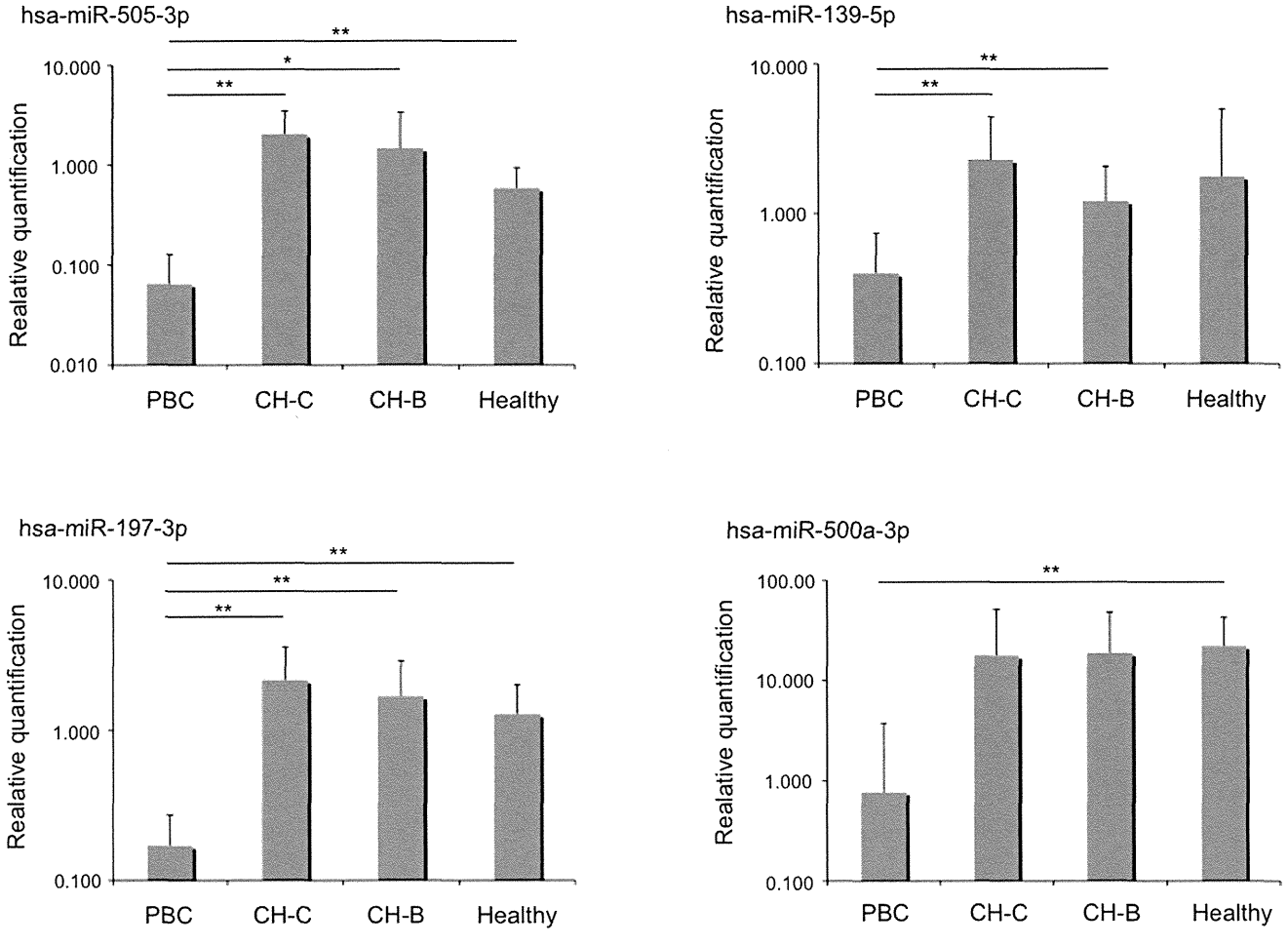


**Figure 2. Validation of deep sequencing results for selected miRNAs.** We have registered 10 samples in each group listed on Table S3. The threshold cycle for each miRNA primer/probe set were normalized with spiked in cel-miR-39 primer/probe pair and compared to CH-C-5. Result for normally distributed continuous variables are given as means and compared between groups by Student's t-test. Results for non-normally distributed continuous variables are summarized as medians and were compared by Mann-Whitney U test. Statistical significance indicates by one asterisk ($p<0.05$) and two ($p<0.01$).
doi:10.1371/journal.pone.0066086.g002

**Table 5.** Predicted target genes of 9 differentially expressed miRNA by Illumina sequencing in PBC.

| Targets | Description | miRIS* | P-value |
|---|---|---|---|
| NM_000350 | ATP-binding cassette, sub-family A (ABC1), member 4 | 0.250 | 1.83E-04 |
| NM_000434 | sialidase 1 (lysosomal sialidase) (NEU1), mRNA. | 0.306 | 6.06E-04 |
| NM_000491 | complement component 1, q subcomponent, B chain | 0.250 | 6.64E-04 |
| NM_000663 | 4-aminobutyrate aminotransferase (ABAT), nuclear gene | 0.417 | 9.55E-03 |
| NM_000767 | cytochrome P450, family 2, subfamily B, polypeptide 6 | 0.361 | 3.72E-03 |
| NM_000878 | interleukin 2 receptor, beta (IL2RB), mRNA. | 0.250 | 7.10E-03 |
| NM_001001716 | nuclear factor of kappa light polypeptide gene | 0.306 | 8.66E-03 |
| NM_001007214 | calcyclin binding protein (CACYBP), transcript variant | 0.306 | 1.90E-03 |
| NM_001012320 | zinc finger protein 302 (ZNF302), transcript variant | 0.306 | 3.32E-03 |
| NM_001029997 | zinc finger protein 181 (ZNF181), transcript variant | 0.292 | 7.83E-03 |
| NM_001033557 | protein phosphatase 1B (formerly 2C), | 0.250 | 7.51E-03 |
| NM_001033910 | TNF receptor-associated factor 5 (TRAF5), transcript | 0.361 | 5.99E-03 |
| NM_001034 | ribonucleotide reductase M2 (RRM2), mRNA. | 0.361 | 6.78E-03 |
| NM_001098831 | MORN repeat containing 4 (MORN4), transcript variant | 0.306 | 4.68E-03 |
| NM_001111125 | IQ motif and Sec7 domain 2 (IQSEC2), transcript | 0.361 | 2.49E-03 |
| NM_001128932 | cytochrome P450, family 4, subfamily F, polypeptide 11 | 0.361 | 8.47E-04 |
| NM_001135146 | solute carrier family 39 (zinc transporter), member 8 | 0.389 | 1.13E-03 |
| NM_001136216 | transmembrane protein 51 (TMEM51), transcript variant | 0.250 | 2.75E-03 |
| NM_001142289 | mahogunin, ring finger 1 (MGRN1), transcript variant | 0.361 | 7.27E-03 |
| NM_001142353 | protein phosphatase 3 (formerly 2B), catalytic | 0.250 | 1.72E-03 |
| NM_001142610 | unc-51-like kinase 2 (C. elegans) (ULK2), transcript | 0.347 | 5.97E-03 |
| NM_001143944 | LEM domain containing 2 (LEMD2), transcript variant 2, | 0.250 | 2.08E-03 |
| NM_001146699 | RNA binding motif protein 19 (RBM19), transcript | 0.306 | 6.81E-03 |
| NM_001159322 | phospholipase A2, group IVC (cytosolic, | 0.292 | 1.32E-03 |
| NM_001166 | baculoviral IAP repeat-containing 2 (BIRC2), mRNA. | 0.250 | 5.10E-03 |
| NM_001293 | chloride channel, nucleotide-sensitive, 1A (CLNS1A), | 0.250 | 1.32E-03 |
| NM_001337 | chemokine (C-X3-C motif) receptor 1 (CX3CR1), mRNA. | 0.417 | 1.39E-03 |
| NM_001678 | ATPase, Na+/K+ transporting, beta 2 polypeptide | 0.417 | 4.07E-04 |
| NM_001903 | catenin (cadherin-associated protein), alpha 1, 102 kDa | 0.389 | 3.97E-03 |
| NM_002298 | lymphocyte cytosolic protein 1 (L-plastin) (LCP1), | 0.417 | 4.64E-03 |
| NM_003810 | tumor necrosis factor (ligand) superfamily, member 10 | 0.250 | 1.95E-03 |
| NM_004414 | regulator of calcineurin 1 (RCAN1), transcript variant | 0.250 | 5.21E-03 |
| NM_004642 | cyclin-dependent kinase 2 associated protein 1 | 0.306 | 8.18E-04 |
| NM_005046 | kallikrein-related peptidase 7 (KLK7), transcript | 0.306 | 7.63E-03 |
| NM_005371 | methyltransferase like 1 (METTL1), transcript variant | 0.250 | 2.91E-03 |
| NM_005517 | high-mobility group nucleosomal binding domain 2 | 0.417 | 8.69E-03 |
| NM_005736 | ARP1 actin-related protein 1 homolog A, centractin | 0.361 | 8.46E-03 |
| NM_006598 | solute carrier family 12 (potassium/chloride | 0.361 | 5.04E-03 |
| NM_014012 | RAS (RAD and GEM)-like GTP-binding 1 (REM1), mRNA. | 0.250 | 7.27E-03 |
| NM_014452 | tumor necrosis factor receptor superfamily, member 21 | 0.250 | 1.53E-03 |
| NM_014567 | breast cancer anti-estrogen resistance 1 (BCAR1), | 0.250 | 6.32E-03 |
| NM_014718 | calsyntenin 3 (CLSTN3), mRNA. | 0.250 | 3.78E-03 |
| NM_014784 | Rho guanine nucleotide exchange factor (GEF) 11 | 0.306 | 7.74E-03 |
| NM_015278 | SAM and SH3 domain containing 1 (SASH1), mRNA. | 0.417 | 8.14E-03 |
| NM_015352 | protein O-fucosyltransferase 1 (POFUT1), transcript | 0.472 | 1.47E-03 |
| NM_016033 | family with sequence similarity 82, member B (FAM82B), | 0.458 | 3.43E-03 |
| NM_016332 | selenoprotein X, 1 (SEPX1), mRNA. | 0.306 | 4.68E-03 |
| NM_019072 | small glutamine-rich tetratricopeptide repeat | 0.417 | 8.71E-03 |
| NM_019860 | 5-hydroxytryptamine (serotonin) receptor 7 (adenylate | 0.250 | 7.39E-04 |

**Table 5.** Cont.

| Targets | Description | miRIS* | P-value |
|---|---|---|---|
| NM_020211 | RGM domain family, member A (RGMA), mRNA. | 0.250 | 3.32E-03 |
| NM_021131 | protein phosphatase 2A activator, regulatory subunit 4 | 0.403 | 5.94E-03 |
| NM_021939 | FK506 binding protein 10, 65 kDa (FKBP10), mRNA. | 0.306 | 1.13E-03 |
| NM_021943 | zinc finger, AN1-type domain 3 (ZFAND3), mRNA. | 0.306 | 1.60E-03 |
| NM_022497 | mitochondrial ribosomal protein S25 (MRPS25), nuclear | 0.417 | 8.60E-03 |
| NM_024025 | dual specificity phosphatase 26 (putative) (DUSP26), | 0.403 | 1.66E-03 |
| NM_024596 | microcephalin 1 (MCPH1), mRNA. | 0.361 | 1.00E-03 |
| NM_024637 | galactose-3-O-sulfotransferase 4 (GAL3ST4), mRNA. | 0.361 | 1.73E-03 |
| NM_024667 | vacuolar protein sorting 37 homolog B (S. cerevisiae) | 0.306 | 3.92E-03 |
| NM_024898 | DENN/MADD domain containing 1C (DENND1C), mRNA. | 0.250 | 7.54E-03 |
| NM_025108 | chromosome 16 open reading frame 59 (C16orf59), mRNA. | 0.250 | 3.39E-03 |
| NM_031287 | splicing factor 3b, subunit 5, 10kDa (SF3B5), mRNA. | 0.306 | 4.48E-03 |
| NM_032139 | ankyrin repeat domain 27 (VPS9 domain) (ANKRD27), | 0.306 | 9.38E-03 |
| NM_032497 | zinc finger protein 559 (ZNF559), mRNA. | 0.347 | 4.45E-03 |
| NM_080678 | ubiquitin-conjugating enzyme E2F (putative) (UBE2F), | 0.347 | 3.60E-03 |
| NM_138396 | membrane-associated ring finger (C3HC4) 9 (MARCH9), | 0.306 | 5.18E-04 |
| NM_138799 | membrane bound O-acyltransferase domain containing 2 | 0.361 | 7.92E-03 |
| NM_145168 | short chain dehydrogenase/reductase family 42E, member | 0.306 | 8.59E-03 |
| NM_147202 | chromosome 9 open reading frame 25 (C9orf25), mRNA. | 0.417 | 4.67E-03 |
| NM_173509 | family with sequence similarity 163, member A | 0.361 | 2.90E-03 |
| NM_175839 | *spermine oxidase* (SMOX), *transcript variant* 1, mRNA. | 0.361 | 4.28E-03 |
| NM_178468 | family with sequence similarity 83, member C (FAM83C), | 0.250 | 4.68E-03 |
| NM_178832 | MORN repeat containing 4 (MORN4), transcript variant | 0.306 | 4.68E-03 |
| NM_178835 | zinc finger protein 827 (ZNF827), mRNA. | 0.361 | 5.67E-03 |
| NM_182527 | calcium binding protein 7 (CABP7), mRNA. | 0.417 | 1.95E-03 |
| NM_198853 | tripartite motif-containing 74 (TRIM74), mRNA. | 0.403 | 1.82E-03 |

*miRIS :miRror Internal Score ranges from 0 top 1 by average 2 components (number of databases and input hits).

doi:10.1371/journal.pone.0066086.t005

gov/) [52]. DAVID can detect functional enrichment of a gene list based on the GO terms, KEGG pathway and BIOCARTA pathway. Differences were considered significant when the P value was less than 0.05.

### Ethics statement

This study was approved by the Ethics Committee of the Tohoku University School of Medicine (2010-404) and written informed consent was obtained from each individual.

## Supporting Information

**Figure S1 The pathway of cell-to-cell adhesion signaling.** The functional annotation analysis of BIOCARTA showed that the genes of catenin (cadherion-associated protein), alpha 1 and similar to breast cancer anti-estrogen resistance 1 played roles in this pathway. The stars indicate the related genes.
(TIFF)

**Figure S2 The pathway of apoptosis.** The functional annotation analysis of BIOCARTA showed that the genes of baculoviral IAP repeat-containing 2, protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform and tumor necrosis factor (ligand) superfamily, member 10 was related to apoptosis. The gene is indicated with the stars.
(TIFF)

**Table S1 The list of differential expression levels of miRNA in each sample.**
(XLS)

**Table S2 Biological function analysis in GO terms of predicted gene targets of differentially regulated miRNAs using DAVID.**
(DOC)

## Acknowledgments

We thank M. Tsuda, M. Kikuchi, N. Koshita and K. Kuroda for technical assistance. We also acknowledge the support of the Biomedical Research Core of Tohoku University Graduate School of Medicine.

Accession number of DNA data bank of Japan (DDBJ) for the deep-sequence data reported in this paper is DRA000933.

## Author Contributions

Conceived and designed the experiments: TS KN YU YK MN. Performed the experiments: MN RF. Analyzed the data: MN TN. Contributed reagents/materials/analysis tools: YK TK EK OK. Wrote the paper: MN YU.

## References


1. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294: 858–862.
2. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294: 853–858.
3. Lee RC, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science 294: 862–864.
4. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.
5. Chen X, Ba Y, Ma L, Cai X, Yin Y, et al. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Research 18: 997–1006.
6. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, et al. (2005) MicroRNA expression profiles classify human cancers. Nature 435: 834–838.
7. Bihrer V, Friedrich-Rust M, Kronenberger B, Forestier N, Haupenthal J, et al. (2011) Serum miR-122 as a biomarker of necroinflammation in patients with chronic hepatitis C virus infection. Am J Gastroenterol 106: 1663–1669.
8. Zhang Y, Jia Y, Zheng R, Guo Y, Wang Y, et al. (2010) Plasma microRNA-122 as a biomarker for viral-, alcohol-, and chemical-related hepatic diseases. Clin Chem 56: 1830–1838.
9. Cermelli S, Ruggieri A, Marrero JA, Ioannou GN, Beretta L (2011) Circulating MicroRNAs in Patients with Chronic Hepatitis C and Non-Alcoholic Fatty Liver Disease. PLoS One 6: e23937.
10. Morita K, Taketomi A, Shirabe K, Umeda K, Kayashima H, et al. (2011) Clinical significance and potential of hepatic microRNA-122 expression in hepatitis C. Liver Int 31: 474–484.
11. Li LM, Hu ZB, Zhou ZX, Chen X, Liu FY, et al. (2010) Serum microRNA Profiles Serve as Novel Biomarkers for HBV Infection and Diagnosis of HBV-Positive Hepatocarcinoma. Cancer Research 70: 9798–9807.
12. Nakanuma Y, Ohta G (1979) Histometric and serial section observations of the intrahepatic bile ducts in primary biliary cirrhosis. Gastroenterology 76: 1326–1332.
13. Gershwin ME, Mackay IR, Sturgess A, Coppel RL (1987) Identification and specificity of a cDNA encoding the 70 kd mitochondrial antigen recognized in primary biliary cirrhosis. J Immunol 138: 3525–3531.
14. Coppel RL, McNeilage LJ, Surh CD, Van de Water J, Spithill TW, et al. (1988) Primary structure of the human M2 mitochondrial autoantigen of primary biliary cirrhosis: dihydrolipoamide acetyltransferase. Proc Natl Acad Sci U S A 85: 7317–7321.
15. Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernizzi P, et al. (2004) Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology 127: 485–492.
16. Begovich AB, Klitz W, Moonsamy PV, Van de Water J, Peltz G, et al. (1994) Genes within the HLA class II region confer both predisposition and resistance to primary biliary cirrhosis. Tissue Antigens 43: 71–77.
17. Donaldson PT, Baragiotta A, Heneghan MA, Floreani A, Venturi C, et al. (2006) HLA class II alleles, genotypes, haplotypes, and amino acids in primary biliary cirrhosis: a large-scale study. Hepatology 44: 667–674.
18. Onishi S, Sakamaki T, Maeda T, Iwamura S, Tomita A, et al. (1994) DNA typing of HLA class II genes; DRB1*0803 increases the susceptibility of Japanese to primary biliary cirrhosis. J Hepatol 21: 1053–1060.
19. Hirschfield GM, Liu X, Xu C, Lu Y, Xie G, et al. (2009) Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. N Engl J Med 360: 2544–2555.
20. Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, et al. (2011) Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat Genet 43: 329–332.
21. Hirschfield G, Invernizzi P (2011) Progress in the Genetics of Primary Biliary Cirrhosis. Seminars in Liver Disease 31: 147–156.
22. Nakamura M, Nishida N, Kawashima M, Aiba Y, Tanaka A, et al. (2012) Genome-wide Association Study Identifies TNFSF15 and POU2AF1 as Susceptibility Loci for Primary Biliary Cirrhosis in the Japanese Population. Am J Hum Genet 91: 721–728.
23. Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, et al. (2009) Primary biliary cirrhosis. Hepatology 50: 291–308.
24. Corpechot C, Poupon R (2007) Geotherapeutics of primary biliary cirrhosis: Bright and sunny around the Mediterranean but still cloudy and foggy in the United Kingdom. Hepatology 46: 963–965.
25. Metcalf JV, Mitchison HC, Palmer JM, Jones DE, Bassendine MF, et al. (1996) Natural history of early primary biliary cirrhosis. Lancet 348: 1399–1402.
26. Gilad S, Meiri E, Yogev Y, Benjamin S, Lebanony D, et al. (2008) Serum microRNAs are promising novel biomarkers. PLoS One 3: e3148.
27. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal 17: 10–12.
28. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11: R25.
29. Balaga O, Friedman Y, Linial M (2012) Toward a combinatorial nature of microRNA regulation in human cells. Nucleic Acids Res 40: 9404–9416.
30. Juran BD, Lazaridis KN (2010) Update on the genetics and genomics of PBC. J Autoimmun 35: 181–187.
31. Zhang R, Peng Y, Wang W, Su B (2007) Rapid evolution of an X-linked microRNA cluster in primates. Genome Res 17: 612–617.
32. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proceedings of the National Academy of Sciences 105: 10513–10518.
33. Ai J, Zhang R, Li Y, Pu J, Lu Y, et al. (2010) Circulating microRNA-1 as a potential novel biomarker for acute myocardial infarction. Biochemical and Biophysical Research Communications 391: 73–77.
34. Chim SSC, Shing TKF, Hung ECW, Leung Ty, Lau Tk, et al. (2008) Detection and Characterization of Placental MicroRNAs in Maternal Plasma. Clinical Chemistry 54: 482–490.
35. Starkey Lewis PJ, Dear J, Platt V, Simpson KJ, Craig DGN, et al. (2011) Circulating microRNAs as potential markers of human drug-induced liver injury. Hepatology 54: 1767–1776.
36. Wang K, Zhang S, Marzolf B, Troisch P, Brightman A, et al. (2009) Circulating microRNAs, potential biomarkers for drug-induced liver injury. Proceedings of the National Academy of Sciences 106: 4402–4407.
37. Tsujiura M, Ichikawa D, Komatsu S, Shiozaki A, Takeshita H, et al. (2010) Circulating microRNAs in plasma of patients with gastric cancers. British Journal of Cancer 102: 1174–1179.
38. Kong X, Du Y, Wang G, Gao J, Gong Y, et al. (2010) Detection of Differentially Expressed microRNAs in Serum of Pancreatic Ductal Adenocarcinoma Patients: miR-196a Could Be a Potential Marker for Poor Prognosis. Digestive Diseases and Sciences 56: 602–609.
39. Silva J, Garcia V, Zaballos A, Provencio M, Lombardia L, et al. (2010) Vesicle-related microRNAs in plasma of nonsmall cell lung cancer patients and correlation with survival. European Respiratory Journal 37: 617–623.
40. Li M, Marin-Muller C, Bharadwaj U, Chow K-H, Yao Q, et al. (2008) MicroRNAs: Control and Loss of Control in Human Physiology and Disease. World Journal of Surgery 33: 667–684.
41. Berg CP, Stein GM, Keppeler H, Gregor M, Wesselborg S, et al. (2007) Apoptosis-associated antigens recognized by autoantibodies in patients with the autoimmune liver disease primary biliary cirrhosis. Apoptosis 13: 63–75.
42. Liang Y, Yang Z, Li C, Zhu Y, Zhang L, et al. (2008) Characterisation of TNF-related apoptosis-inducing ligand in peripheral blood in patients with primary biliary cirrhosis. Clinical and Experimental Medicine 8: 1–7.
43. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29.
44. Nakagome Y, Ueno Y, Kogure T, Fukushima K, Moritoki Y, et al. (2007) Autoimmune cholangitis in NOD.c3c4 mice is associated with cholangiocyte-specific Fas antigen deficiency. J Autoimmun 29: 20–29.
45. Singh R, Bullard J, Kalra M, Assefa S, Kaul AK, et al. (2011) Status of bacterial colonization, Toll-like receptor expression and nuclear factor-kappa B activation in normal and diseased human livers. Clinical Immunology 138: 41–49.
46. Kyriakou DS, Alexandrakis MG, Zachou K, Passam F, Stathakis NE, et al. (2003) Hemopoietic progenitor cells and bone marrow stromal cells in patients with autoimmune hepatitis type 1 and primary biliary cirrhosis. J Hepatol 39: 679–685.
47. Maglott D, Ostell J, Pruitt KD, Tatusova T (2010) Entrez Gene: gene-centered information at NCBI. Nucleic Acids Research 39: D52–D57.
48. Kozomara A, Griffiths-Jones S (2010) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Research 39: D152–D157.
49. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57: 289–300.
50. Team RDC (2011) R: a language and environment for statistical computing. Vienna, Austria: the R Foundation for Statistical Computing.
51. Friedman Y, Naamati G, Linial M (2010) MiRror: a combinatorial analysis web tool for ensembles of microRNAs and their targets. Bioinformatics 26: 1920–1921.
52. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4: P3.
53. Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, et al. (1981) Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology 1: 431–435.
54. Ludwig J, Dickson ER, McDonald GS (1978) Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 379: 103–112.
55. Scheuer P (1967) Primary biliary cirrhosis. Proceedings of the Royal Society of Medicine 60: 1257–1260.
56. Ward JH (1963) Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58: 236–244.

Hepatology Research

*Hepatology Research* 2013 doi: 10.1111/hepr.12168

***Short Communication***

# Sequential analysis of amino acid substitutions with hepatitis B virus in association with nucleoside/nucleotide analog treatment detected by deep sequencing

Masashi Ninomiya,[1] Yasuteru Kondo,[1] Tetsuya Niihori,[2] Takeshi Nagashima,[3] Takayuki Kogure,[1] Eiji Kakazu,[1] Osamu Kimura,[1] Yoko Aoki,[2] Yoichi Matsubara[2] and Tooru Shimosegawa[1]

[1]*Division of Gastroenterology, Tohoku University Hospital,* [2]*Department of Medical Genetics, and* [3]*Division of Cell Proliferation, Tohoku University Graduate School of Medicine, Sendai, Japan*

Taking nucleoside/nucleotide analogs is a major antiviral therapy for chronic hepatitis B infection. The problem with this treatment is the selection for drug-resistant mutants. Currently, identification of genotypic drug resistance is conducted by molecular cloning sequenced by the Sanger method. However, this methodology is complicated and time-consuming. These limitations can be overcome by deep sequencing technology. Therefore, we performed sequential analysis of the frequency of drug resistance in one individual, who was treated with lamivudine on-and-off therapy for 2 years, by deep sequencing. The lamivudine-resistant mutations at rtL180M and rtM204V and the entecavir-resistant mutation at rtT184L were detected in the first subject. The lamivudine- and entecavir-resistant strain was still detected in the last subject. However, in the deep sequencing analysis, rt180 of the first subject showed a mixture in 76.9% of the methionine and in 23.1% of the leucine, and rt204 also showed a mixture in 69.0% of the valine and 29.8% of the isoleucine. During the treatment, the ratio of resistant mutations increased. At rt184, the resistant variants were detectable in 58.7% of the sequence, with the replacement of leucine by the wild-type threonine in the first subject. Gradually, entecavir-resistant variants increased in 82.3% of the leucine in the last subject. In conclusion, we demonstrated the amino acid substitutions of the serial nucleoside/nucleotide analog resistants. We revealed that drug-resistant mutants appear unchanged at first glance, but actually there are low-abundant mutations that may develop drug resistance against nucleoside/nucleotide analogs through the selection of dominant mutations.

**Key words:** amino acid substitutions, deep sequencing, hepatitis B virus, nucleoside/nucleotide analog resistants

Approximately 350–400 million patients are chronically infected with hepatitis B virus (HBV) globally, and the disease has caused epidemics in East Asia.[1,2] In Japan, approximately 1.5 million people are infected with HBV.[3] Chronic hepatitis B (CHB) increases the risk of liver cirrhosis and hepatocellular carcinoma.[4]

Hepatitis B virus is a DNA virus of 3.2 kb surrounded by an envelope of the surface protein (hepatitis B surface antigen [HBsAg]) and it has a circular genome of partially double-stranded DNA.[5] Once the HBV invades into hepatic cells, genomic DNA is transferred to the cell nucleus.[6] In the nucleus, the genomic DNA is converted to a stable intrahepatic reservoir of cccDNA. The cccDNA replicates through an RNA intermediate form by reverse transcription.[7] The purpose of CHB therapy is to achieve sustained suppression of HBV replication and the remission of liver disease. However, cccDNA is resistant to treatment and is not completely eradicated by currently available medications.[8,9] Taking nucleoside/nucleotide analogs (NA) is a major antiviral therapy for the treatment of CHB.[4] Therapies with NA available in Japan include lamivudine (LAM), adefovir dipivoxil (ADV) and entecavir (ETV). They inhibit viral polymerase activity by interfering with the priming of reverse transcription and elongation of the viral minus or plus strand DNA.[10–12] Most patients have chemical and virological responses, but these treatments are hampered by the selection of drug-resistant mutants, leading

*Correspondence: Dr Yasuteru Kondo, Division of Gastroenterology, Tohoku University Hospital, 1-1 Seiryo, Aoba-ku, Sendai 980-8574, Japan. Email: yasuteru@ebony.plala.or.jp*

The GenBank/EMBL/DDBJ accession numbers for the nucleotide sequences reported in this paper are AB820840-AB820852.

*Received 7 March 2013; revision 14 May 2013; accepted 20 May 2013.*

to a loss of efficacy, viral relapse and exacerbations of hepatitis after discontinuation.[13]

In the present study, the region of codon rt148 to rt 208 in the polymerase open reading frame was found to include the reported relevant NA-resistant amino acid substitutions. Resistance to LAM has been mapped to the YMDD locus in the C domain of rtM204I/V and is sometimes associated with compensatory mutations in the B domain of rtL180M and/or rtV173L.[14–16] The mutations common to LMV confer cross-resistance and reduced sensitivity to ETV but not to ADV. ETV resistance has been mapped to the B domain with rtI169T, rtL180M and/or rtT184S/A/I/L/F/G/C/M, C domain with rtS202C/G/I and rtM204V/I, and E domain of rtM250I/L/V. The three amino acid substitutions of rtL180 + rtM204 and either rtT184, rtS202 or rtM250 are required for ETV resistance to develop.[17–19] The mutations in the B domain of rtA181S/T/V and D domain of rtN236T were reported to be associated with ADV resistance.[20,21]

Currently, identification of HBV genotypic NA resistance is mainly conducted by polymerase chain reaction (PCR) amplification with Sanger direct sequencing. However, with this method it is difficult to measure the frequencies of each mutation, and it is impossible to detect several mutations combined in the same sequence. It is necessary to find the frequencies of NA-resistant amino acid substitutions, because the secondary compensatory mutations associated with primary NA resistance may restore replication defects or even give rise to multidrug-resistant variants.[19] As an alternative to Sanger direct sequencing, molecular cloning can analyze single viral DNA molecules. However, this methodology is complicated and time-consuming, because analysis of three-digit clones is required to detect variants present in several percent of quasispecies. These limitations can now be overcome by deep sequencing technology.[22] Genetic diversity plays a key role in the NA treatment of HBV infection. Therefore, using this technology, minor HBV variants can be detected, including those with combinations of nucleotide changes in the same period. Detecting low-frequency NA resistants will be important for choosing the NA treatment. In this study, we investigated the transition frequency of amino acid substitutions in NA resistants, in one case using deep sequencing during 2 years of ETV and ADV combination treatment.

A 42-year-old man, diagnosed with CHB infection in June 2005, was treated with LAM on-and-off therapy for 2 years. In HIV-infected patients, on-and-off antiretroviral therapy increased the risk of drug resistance compared with continuous therapy.[23] The HBV DNA levels had gradually increased and ETV treatment was begun in November 2007, but the HBV DNA levels did not decrease, and the treatment was stopped in November 2009. The level of serum transaminase was not elevated at that period. In April 2010, he was seen for the first time in our hospital. The patient was asymptomatic but showed an elevation of transaminase. The following clinical data was indicated: aspartate aminotransferase, 66 IU/L; alanine aminotransferase (ALT), 178 IU/L; positive for HBsAg, hepatitis B e-antigen (HBeAg) and hepatitis B core antibody; HBV genotype C; HBV DNA levels of $1.0 \times 10^8$ copies/mL; and negative for anti-hepatitis C virus and anti-HIV. The combination of ETV (0.5 mg/day) and ADV (10 mg/day) therapy was started at that time. After 3 months of therapy, the ALT elevation (119 IU/L) still persisted. Then, the dose of ETV was increased from 0.5 mg to 1.0 mg daily from August 2010. Three months after starting ETV at a dose of 1.0 mg, ALT was at 40 IU/L and HBV DNA had decreased to $6.3 \times 10^3$ copies/mL. At the most recent follow up, ALT has remained normal and HBV DNA is at nearly an undetectable level ($<1.3 \times 10^2$ copies/mL). Of note, the serum level of HBeAg is still positive. The changes in viral load and ALT levels are presented in Figure 1(a).

To investigate mutation involved in NA resistance, six serial serum samples of HBV-P10140 (stored in April 2010), HBV-P10193 (June 2010), HBV-P10234 (July 2010), HBV-P10264 (August 2010), HBV-P11021 (January 2011) and 12-009 (January 2012) were enrolled in this study. We extracted HBV DNA by the method described previously with slight modification.[24] Nucleic acids were obtained from 100 μL serum samples using SMITEST EX-R&D (Medical & Biological Laboratories, Nagoya, Japan) and nested PCR was conducted in the presence of PrimeSTAR HS DNA polymerase (TaKaRa Bio, Shiga, Japan) four with primers targeting the polymerase gene of HBV genomes (Table 1). The amplification product of the first-round PCR was 457 bp (nt 596–1052), and that of the second-round PCR was 390 bp (nt 610–999): the nucleotide numbers are in accordance with a genotype B HBV isolate of 3215 nt (accession no. AB010289). The first-round PCR with primers B034 and B037 was conducted for 30 cycles and the second-round PCR for 25 cycles with primers B035 and B036 (Table 1). The second-round PCR amplicons were sequenced directly on both strands using a BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on a 3500xL Genetic Analyzer (Applied Biosystems).

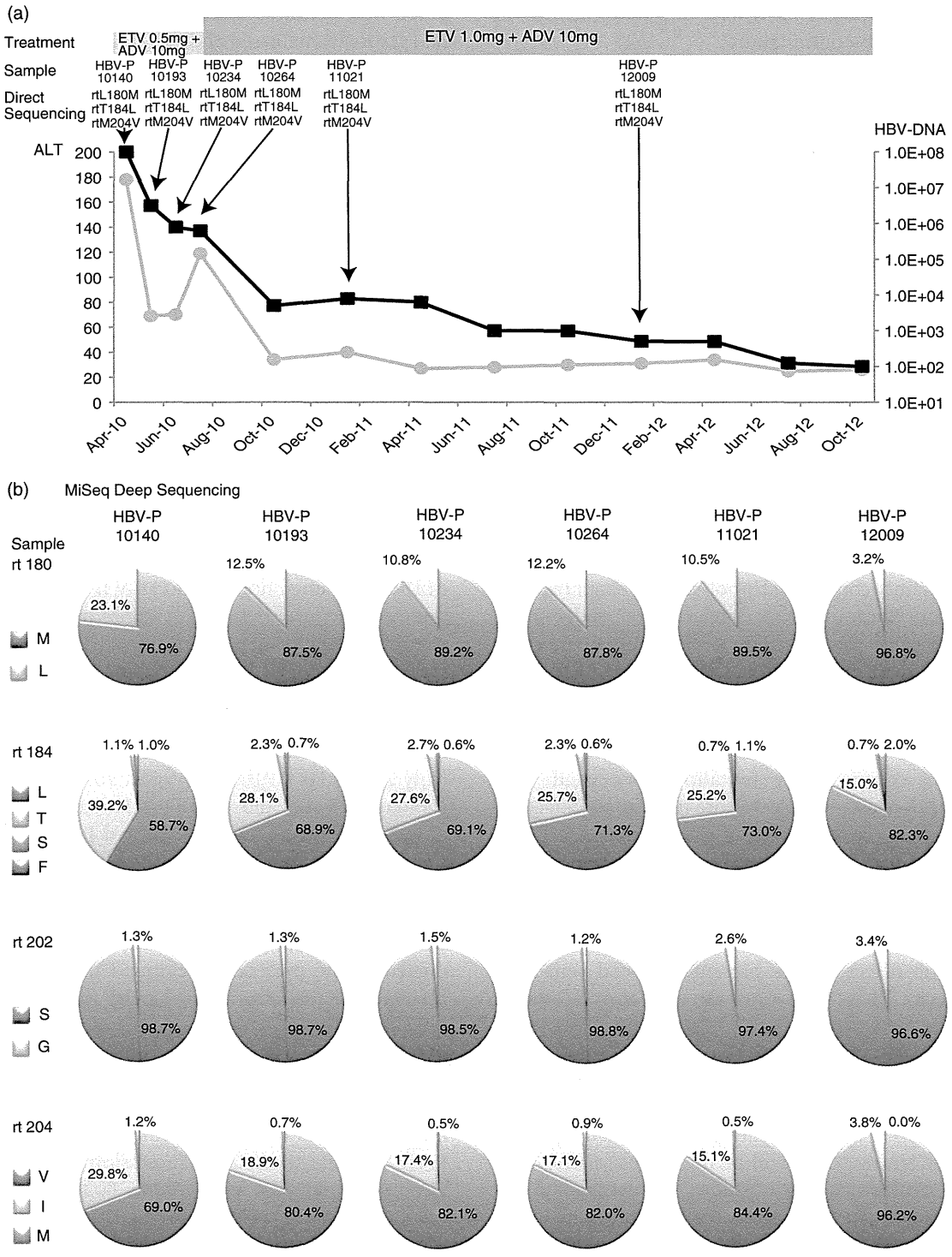


**Figure 1** Clinical stage and amino acid substitutions of hepatitis B virus (HBV) resistants. (a) Evolution of HBV DNA viral load, alanine aminotransferase (ALT) levels and resistant HBV variants sequenced directly with Sanger methods in a patient treated sequentially with entecavir (ETV) and adefovir dipivoxil (ADV). (b) Frequency of amino acid substitutions of nucleoside/nucleotide analog (NA)-resistant HBV by MiSeq deep sequencing. M denotes methionine, L denotes leucine, T denotes threonine, V denotes valine, S denotes serine, F denotes phenylalanine, G denotes glycine and I denotes isoleucine. ■, HBV-DNA; ●, ALT.

**Table 1** Nucleotide sequence of oligonucleotide primers

| Primer | Polarity | Usage | Nucleotide sequence |
|---|---|---|---|
| B034 | Sense | 1st PCR | ACYTGTATTCCCATCCCATC |
| B037 | Antisense | 1st PCR | AARGCAGGRTADCCACATTG |
| B035 | Sense | 2nd PCR | CCCATCRTCYTGGGCTTTCG |
| B036 | Antisense | 2nd PCR | AATTCKYTGACATACTTTCCAATC |
| P5_B035 | Sense | MiSeq sequence | AATGATACGGCGACCACCGAGATCTACACTC TTTCCCTACACGACGCTCTTCCGATCTCCCA *TCRTCYTGGGCTTTCG* |
| P7_Index1_B036 | Antisense | MiSeq sequence of HBV-P10140 | CAAGCAGAAGACGGCATACGAGAT***CGTGAT***G TGACTGGAGTTCAGACGTGTGCTCTTCCGA TCT*AATTCKYTGACATACTTTCCAATC* |
| P7_Index2_B036 | Antisense | MiSeq sequence of HBV-P10193 | CAAGCAGAAGACGGCATACGAGAT***ACATCG***G TGACTGGAGTTCAGACGTGTGCTCTTCCGAT *CTAATTCKYTGACATACTTTCCAATC* |
| P7_Index3_B036 | Antisense | MiSeq sequence of HBV-P10234 | CAAGCAGAAGACGGCATACGAGAT***GCCTAA***GT GACTGGAGTTCAGACGTGTGCTCTTCCGATCT *AATTCKYTGACATACTTTCCAATC* |
| P7_Index4_B036 | Antisense | MiSeq sequence of HBV-P10264 | CAAGCAGAAGACGGCATACGAGAT***TGGTCA***GTG ACTGGAGTTCAGACGTGTGCTCTTCCGATCTAA *TTCKYTGACATACTTTCCAATC* |
| P7_Index5_B036 | Antisense | MiSeq sequence of HBV-P11021 | CAAGCAGAAGACGGCATACGAGAT***CACTGT***GTG ACTGGAGTTCAGACGTGTGCTCTTCCGATCTAA *TTCKYTGACATACTTTCCAATC* |
| P7_Index6_B036 | Antisense | MiSeq sequence of HBV-P12009 | CAAGCAGAAGACGGCATACGAGAT***ATTGGC***GTGA CTGGAGTTCAGACGTGTGCTCTTCCGATCT*AATT CKYTGACATACTTTCCAATC* |

Y denotes C or T, R denotes A or G, D denotes A, G or T and K denotes G or T. Italic bold letters indicate the index by multiplex sequencing. Italic letters indicate the nucleotide sequence of B035 or B036.
PCR, polymerase chain reaction.

The LAM-resistant mutations at rtL180M and rtM204V, and the ETV-resistant mutation at rtT184L in HBV-P10140 were detected by direct sequencing. The HBV-mutated LAM- and ETV-resistant strain was still detected in HBV-P12009 and the amino acid substitutions were the same in all subjects (Fig. 1a).

Next, we investigated the frequency of amino acid bases with these same six PCR amplicons by MiSeq (Illumina, San Diego, CA, USA) deep sequencing. MiSeq is the only deep sequencer that quickly integrates amplification, sequencing and data analysis in a single instrument. Compared with HiSeq (Illumina) or Illumina GA IIx, MiSeq can produce longer paired-end reads ($<2 \times 250$) with fewer gigabases of data. However, it has a large enough data volume to sequence PCR amplicons.

A second PCR was performed to attach the required sequencing adaptor for Illumina Miseq sequencing protocol as well as barcode to allow multiplexing of multiple sample libraries per sequencing lane. Six pairs of primers tailed with the adaptor and the specific index were used to amplify six first PCR amplicons. Eight cycles of PCR were performed using PrimeSTAR HS DNA polymerase. All adaptor–barcode–primers are shown in Table 1. The PCR amplicons were purified using AMPure XP beads (Beckman Coulter, Danvers, MA, USA) and the quality and quantity were checked by size range analysis using a D1K ScreenTape on the 2200 Tapestation (Agilent Technologies, Santa Clara, CA, USA). Each library was pooled in equal amounts, and a PhiX control kit (Illumina) was added to 95% of the pooled libraries. Sequencing was performed on MiSeq (Illumina) using 251-bp paired-end reads. Fastq files including demultiplexed sequence reads were generated by Miseq reporter. 251-bp paired-end reads were stripped of low quality 3′-regions using trim galore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Sequence reads were aligned with the HBV reference sequence of Yamagata-1 (accession no. AB010289) using of Bowtie 2 (2.1.0).[25] The analysis

**Table 2** Number of reads and read depth of each sample

| Sample | HBV-P10140 | HBV-P10193 | HBV-P10234 | HBV-P10264 | HBV-P11021 | HBV-P12009 |
|---|---|---|---|---|---|---|
| No. of reads mapping on HBV reference | 187 336 | 186 386 | 166 257 | 158 781 | 169 709 | 137 139 |
| Read depth | 88 985 | 87 825 | 76 775 | 75 296 | 80 190 | 65 095 |

of the read depth and base count is displayed using GATK (www.broadinstitute.org/gatk/).

A total of 1 005 608 validated reads corresponding to six serial samples were analyzed by deep sequencing. The mean number of reads mapping on the HBV reference was 167 601 and the mean depth of each case was 79 028 (Table 2). The nine most variable amino acid substitutions, rt169, rt173, rt180, rt181, rt184, rt202, rt204, rt236 and rt250, known to be associated with NA resistance, were examined. The problem was how to differentiate true mutations from sequencing error in the obtained reads. Nishijima *et al.* tried to determine the PCR amplicon of the HBV genome derived from the expression plasmid to take the PCR-induced errors as well as sequencing errors into consideration. They described that the mean error rate was 0.034%, with the distribution of the per-nucleotide error rate ranging 0–0.13%.[26] Therefore, mutations over 1.0% were considered to be valid. First, rt180 of HBV-P10140 showed mixed variants in 76.9% of methionine with LAM-resistant mutations and in 23.1% of wild type with leucine, and rt204 also showed mixed variants in 69.0% of valine and 29.8% of isoleucine with LAM-resistant mutations, and in only 1.2% of wild-type methionine. During the NA treatment, the ratio of resistance mutations was increased, and rt180 of HBV-P12009 showed mixed variants in 96.8% of methionine with LAM-resistant mutations and rt204 also showed mixed variants in 96.2% of valine with LAM-resistant mutations. At rt184, related to ETV resistance, the resistant variants were detectable in 58.7% of the sequence, with replacement of leucine by wild-type threonine in HBV-P10140. Gradually, ETV-resistant variants increased and were detected in 82.3% of leucine in HBV-P12009. In addition, at rt202, the ration of serine with ETV mutants was increased from 1.3% to 3.4% during the treatment with NA (Fig. 1b). Of note, at rt169, rt173, rt181, rt236 and rt250, there were no NA-resistant mutations.

Then, we investigated the existence of drug-resistant HBV clones that could latently develop in patients. We enrolled seven patients treated with continuous NA for up to 2 years but in whom there was continual detection of viral copies in serum in this trial. Direct Sanger sequencing detected LAM-resistant mutations in five of the seven subjects. They had the following mutations: (i) rtL180M + rtM204V and (ii) rt180M and no mutations were found at rtV173L. However, by deep sequencing, in four subjects there was part of a mixture with rtL180L/M (35.4%/64.6%) and rtM204M/V (35.2%/64.8%) in HBV-P10036, with rtL180L/M (10.6%/86.8%) and rtM204M/V/I (0.0%/75.5%/24.5%) in HBV-P12043, with rtL180L/M (19.5%/80.5%) and rtM204M/V/I (4.4%/79.5%/16.1%) in HBV-P13064 and with rtL180L/M (23.4%/76.6%), rtM204M/V/I (23.2%/41.6%/35.2%) and one secondary minority variant of rtV173V/L (55.4%/44.6%) in HBV-P13061. Next, an ADV-resistant mutation of rtA181V was detected in two subjects by Sanger sequencing, and was present in 100.0% in HBV-P07061 and in 35.3% in HBV-P08463 by deep sequencing. In addition, minority variant ADV-resistant mutations were detected with 8.5% of rtA181V and with 14.5% of rtN236T in HBV-P10036, and with 2.1% of rtA181V in HBV-P13064. Of note, an ETV-resistant mutation appeared only in one subject with rtT184L by Sanger sequencing, while some low frequency ETV-resistant mutations were detected in four subjects by deep sequencing (Table 3). Therefore, the continual viremia of HBV treated with NA for long term may lead to NA-resistant mutations.

In conclusion, we demonstrated the amino acid substitutions of serial NA-resistant HBV by Sanger sequencing and MiSeq deep sequencing. We revealed that NA-resistant mutants appear unchanged at a glance but suggest the existence of low-abundant mutant clones that may develop drug resistance against NA through the selection of dominant mutations. The viral loads have decreased, but the effect of ETV is expected to be reduced. Further analysis by deep sequencing technologies is necessary to understand the significance and clinical relevance of viral mutations in the pathophysiology of NA-resistant HBV infection.

**Table 3** HBV NA-resistant mutations detected by Sanger sequencing and deep sequencing

| Sample | Treatment | HBV DNA (log copies/mL) | NA-resistant mutations | |
|---|---|---|---|---|
| | | | Sanger sequence | Deep sequence |
| HBV-P07061 | LAM + ADV | 5.0 | A181V | **A181V (100.0%)** |
| HBV-P10036 | ETV | 2.5 | L180M, M204V | V173M (1.3%), **L180M (64.6%)**, **A181V (8.5%)**, **T184S (4.5%)**, **S202C (2.3%)**, **M204V (64.8%)**, **N236T (14.3%)** |
| HBV-P08463 | ETV | 3.4 | A181V | **A181V (35.3%)** |
| HBV-P12043 | ETV | 4.7 | L180M, M204V | V173M (4.4%), **V173L (2.7%)**, **L180M (86.8%)**, L180Q (2.6%), **T184L (27.9%)**, **T184M (5.1%)**, **S202G (16.6%)**, **M204V (75.5%)**, **M204I (24.5%)** |
| HBV-P12081 | LAM + ADV | 2.9 | L180M, T184L, M204V | I169A (25.4%), **L180M (100.0%)**, **T184L (100.0%)**, **M204V (100.0%)** |
| HBV-P13064 | ETV | 7.1 | L180M, M204V | V173M (12.9%), **L180M (80.5%)**, **A181V (2.1%)**, **A181T (3.6%)**, **M204V (79.5%)**, **M204I (16.1%)**, **M250L (2.8%)** |
| HBV-P13061 | ETV | 2.3 | L180M | **V173L (44.6%)**, **L180M (76.6%)**, **M204V (41.6%)**, **M204I (35.2%)**, **M250L (1.3%)** |

Mutations in bold type at deep sequence section showed NA-resistant mutations that have already been reported.
ADV, adefovir dipivoxil; ETV, entecavir; HBV, hepatitis B virus; LAM, lamivudine; NA, nucleoside/nucleotide analog.

## REFERENCES


1 EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J Hepatol* 2012; **57**: 167–85.
2 Pan CQ, Hu KQ, Tsai N. Long-term therapy with nucleoside/nucleotide analogues for chronic hepatitis B in Asian patients. *Antivir Ther* 2012; doi: 10.3851/IMP2481. [Epub ahead of print].
3 Yokosuka O, Kurosaki M, Imazeki F *et al.* Management of hepatitis B: consensus of the Japan Society of Hepatology 2009. *Hepatol Res* 2011; **41**: 1–21.
4 Lok ASF, McMahon BJ. Chronic hepatitis B: update 2009. *Hepatology* 2009; **50**: 661–2.
5 Zuckerman AJ. Hepatitis viruses. In: Baron S, ed. *Medical Microbiology*, 4th edn. Galveston (TX): University of Texas Medical Branch at Galveston, 1996; Chapter 70.
6 Kann M, Bischof A, Gerlich WH. In vitro model for the nuclear transport of the hepadnavirus genome. *J Virol* 1997; **71**: 1310–16.
7 Locarnini S. Molecular virology of hepatitis B virus. *Semin Liver Dis* 2004; **24** (Suppl 1): 3–10.
8 Werle-Lapostolle B, Bowden S, Locarnini S *et al.* Persistence of cccDNA during the natural history of chronic hepatitis B and decline during adefovir dipivoxil therapy. *Gastroenterology* 2004; **126**: 1750–8.
9 Wong DK, Yuen MF, Ngai VW, Fung J, Lai CL. One-year entecavir or lamivudine therapy results in reduction of hepatitis B virus intrahepatic covalently closed circular DNA levels. *Antivir Ther* 2006; **11**: 909–16.
10 Severini A, Liu XY, Wilson JS, Tyrrell DL. Mechanism of inhibition of duck hepatitis B virus polymerase by (-)-beta-L-2',3'-dideoxy-3'-thiacytidine. *Antimicrob Agents Chemother* 1995; **39**: 1430–5.
11 Seigneres B, Aguesse-Germon S, Pichoud C *et al.* Duck hepatitis B virus polymerase gene mutants associated with resistance to lamivudine have a decreased replication capacity in vitro and in vivo. *J Hepatol* 2001; **34**: 114–22.
12 Seifer M, Hamatake RK, Colonno RJ, Standring DN. In vitro inhibition of hepadnavirus polymerases by the triphosphates of BMS-200475 and lobucavir. *Antimicrob Agents Chemother* 1998; **42**: 3200–8.
13 Inada M, Yokosuka O. Current antiviral therapies for chronic hepatitis B. *Hepatol Res* 2008; **38**: 535–42.
14 Allen MI, Deslauriers M, Andrews CW *et al.* Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. *Hepatology* 1998; **27**: 1670–7.
15 Fu L, Cheng YC. Role of additional mutations outside the YMDD motif of hepatitis B virus polymerase in L(-)SddC (3TC) resistance. *Biochem Pharmacol* 1998; **55**: 1567–72.
16 Fu L, Liu SH, Cheng YC. Sensitivity of L-(-)2,3-dideoxythiacytidine resistant hepatitis B virus to other antiviral nucleoside analogues. *Biochem Pharmacol* 1999; **57**: 1351–9.
17 Tenney DJ, Levine SM, Rose RE *et al.* Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to Lamivudine. *Antimicrob Agents Chemother* 2004; **48**: 3498–507.
18 Tenney DJ, Rose RE, Baldick CJ *et al.* Long-term monitoring shows hepatitis B virus resistance to entecavir in

nucleoside-naive patients is rare through 5 years of therapy. *Hepatology* 2009; **49**: 1503–14.
19 Locarnini S. Primary resistance, multidrug resistance, and cross-resistance pathways in HBV as a consequence of treatment failure. *Hepatol Int* 2008; **2**: 147–51.
20 Bartholomeusz A, Tehan BG, Chalmers DK. Comparisons of the HBV and HIV polymerase, and antiviral resistance mutations. *Antivir Ther* 2004; **9**: 149–60.
21 Yatsuji H, Suzuki F, Sezaki H *et al.* Low risk of adefovir resistance in lamivudine-resistant chronic hepatitis B patients treated with adefovir plus lamivudine combination therapy: two-year follow-up. *J Hepatol* 2008; **48**: 923–31.
22 Ninomiya M, Ueno Y, Funayama R *et al.* Use of illumina deep sequencing technology to differentiate hepatitis C virus variants. *J Clin Microbiol* 2012; **50**: 857–66.
23 Danel C, Moh R, Chaix ML *et al.* Two-months-off, four-months-on antiretroviral regimen increases the risk of resistance, compared with continuous therapy: a randomized trial involving West African adults. *J Infect Dis* 2009; **199**: 66–76.
24 Takahashi M, Nishizawa T, Gotanda Y *et al.* High prevalence of antibodies to hepatitis A and E viruses and viremia of hepatitis B, C, and D viruses among apparently healthy populations in Mongolia. *Clin Diagn Lab Immunol* 2004; **11**: 392–8.
25 Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357–9.
26 Nishijima N, Marusawa H, Ueda Y *et al.* Dynamics of hepatitis B virus quasispecies in association with nucleos(t)ide analogue treatment determined by ultra-deep sequencing. *PLoS ONE* 2012; **7**: e35052.

*Human Molecular Genetics, 2014, Vol. 23, No. 4* **992–1001**
*doi:10.1093/hmg/ddt495*
*Advance Access published on October 8, 2013*

# RNA sequencing-based identification of aberrant imprinting in cloned mice

**Hiroaki Okae[1,3,†], Shogo Matoba[4,†], Takeshi Nagashima[2], Eiji Mizutani[4,5], Kimiko Inoue[4], Narumi Ogonuki[4], Hatsune Chiba[1,3], Ryo Funayama[2], Satoshi Tanaka[6], Nobuo Yaegashi[7], Keiko Nakayama[2], Hiroyuki Sasaki[3,8], Atsuo Ogura[4] and Takahiro Arima[1,3,*]**

[1]Department of Informative Genetics, Environment and Genome Research Center, [2]Division of Cell Proliferation, United Centers for Advanced Research and Translational Medicine, Tohoku University Graduate School of Medicine, Sendai 980-8575, Japan [3]JST, CREST, Saitama 332-0012, Japan [4]RIKEN BioResource Center, Tsukuba, Ibaraki 305-0074, Japan [5]Faculty of Life and Environmental Sciences, University of Yamanashi, Kofu 400-8510, Japan [6]Laboratory of Cellular Biochemistry, Department of Animal Resource Sciences/Veterinary Medical Sciences, The University of Tokyo, Tokyo 113-8657, Japan [7]Department of Obstetrics and Gynecology, Tohoku University Graduate School of Medicine, Sendai 980-8574, Japan and [8]Division of Epigenomics, Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan

Received May 8, 2013; Revised August 19, 2013; Accepted October 2, 2013

**Animals cloned by somatic cell nuclear transfer (SCNT) provide a unique model for understanding the mechanisms of nuclear epigenetic reprograming to a state of totipotency. Though many phenotypic abnormalities have been demonstrated in cloned animals, the underlying mechanisms are not well understood. In this study, we performed transcriptome-wide allelic expression analyses in brain and placental tissues of cloned mice. We found that *Gab1*, *Sfmbt2* and *Slc38a4* showed loss of imprinting in all cloned mice analyzed, which might be involved in placentomegaly of cloned mice. These three genes did not require *de novo* DNA methylation in growing oocytes for the establishment of imprinting, implying the involvement of a *de novo* DNA methylation-independent mechanism. Loss of *Dlk1-Dio3* imprinting was also observed in nearly half of cloned mouse embryos and showed a strong correlation with embryonic lethality. Our findings are essential to understand the underlying mechanisms of developmental abnormalities of cloned animals. We also emphasize that particular attention should be paid to specific imprinted genes for therapeutic and agricultural applications of SCNT.**

## INTRODUCTION

Animals cloned by somatic cell nuclear transfer (SCNT) provide a unique model for understanding the mechanisms of nuclear epigenetic reprograming to a state of totipotency (1,2). However, the cloning efficiency is very low regardless of the animal species or donor cell type. Though many phenotypic abnormalities have been demonstrated in cloned animals, including frequent embryonic and perinatal death and placentomegaly, the underlying mechanisms are not well understood (3).

Genomic imprinting is an epigenetic gene regulatory mechanism that leads to the preferential expression of the paternally or maternally inherited allele of a subset of genes. Most notably, DNA methylation, which occurs at discrete locations termed differentially methylated regions (DMRs) in the germline, initiates the imprinting process (4). The *de novo* methyltransferase Dnmt3a and the Dnmt3-related protein Dnmt3L are required for the establishment of germline DMRs (5,6). The majority of imprinted genes reside within imprinted domains, and these DMRs are known to control imprinted gene expression within the domain (7). Abnormal expression of some imprinted genes causes developmental defects (8), some of which are similar to those observed in cloned animals.

*To whom correspondence should be addressed at: Department of Informative Genetics, Environment and Genome Research Center, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8575, Japan. Tel: 81 227177844; Fax:+ 81 227177063; Email: tarima@med.tohoku.ac.jp
†H.O. and S.M. contributed equally to this work.

Abnormal imprinting of some imprinted genes has been reported in cloned animals (9–11). In particular, abnormal allelic expression or DMR methylation of *H19*, *Igf2r*, *Peg3*, *Snrpn*, *Ascl2* and *Xist* has been reported in cloned mice (12–16). However, the vast majority of the abnormalities are only partial loss of imprinting or donor cell type-specific. For example, aberrant imprinting of *H19* and *Igf2* was frequently observed in embryonic stem cell-derived cloned mice (14), but not in somatic cell-derived ones (17). Furthermore, the relationships between developmental defects and imprinting abnormalities are poorly understood in cloned mice. The only exception is ectopic expression of *Xist*, which occurs in nearly all cloned mouse embryos, and accounts for embryonic loss at the postimplantation stage (16,18). Importantly, Inoue *et al.* reported faithful monoallelic expression of seven well-known imprinted genes in cloned mouse embryos (17). Overall, the memory of genomic imprinting is believed to be stably maintained in cloned mice with the exception of *Xist*.

In the mouse, ~100 imprinted genes have been identified, but the imprinting status of most known imprinted genes has not been analyzed in cloned mice. Especially, we recently reported placenta-specific imprinted genes, which show imprinted expression in the placenta but are biallelically expressed in the fetus (19). It is possible that the nuclei of fetus-derived somatic cells do not acquire placenta-specific imprinting after SCNT. Therefore, we performed transcriptome-wide allelic expression analyses and found that three imprinted genes, including two placenta-specific imprinted genes, consistently showed loss of imprinting in all cloned mice analyzed (consistent loss of imprinting). We also verified the correlation between loss of *Dlk1-Dio3* imprinting and embryonic lethality of cloned mice.

## RESULTS

### Transcriptome-wide analyses of imprinted gene expression in cloned mouse placentas and brains

For whole transcriptome sequencing analyses, we selected the placenta and brain because many imprinted genes, including genes with tissue-specific imprinted expression, are known to be expressed in these tissues (19,20). 129X1/SvJ (129) females were mated with Japanese fancy 1 (JF1) males to generate [129xJF1]F1 mice from which cumulus and Sertoli cells were isolated and used for SCNT. F1 embryos were also generated by *in vitro* fertilization (IVF) to provide a comparison. Placentas and brains were dissected from embryonic day (E) 13.5 embryos and analyzed. Forty-five known imprinted genes had at least one single nucleotide polymorphism (SNP) site between 129 and JF1 and were expressed at a sufficient level to assess allelic expression in the placenta and/or brain (Supplementary Material, Table S1). For the analyses of allelic expression, the maternal read number divided by the paternal read number (M/P ratio) was calculated for each imprinted gene. The maternal expression and paternal expression are defined as [M/P ratio] $>2$ and [M/P ratio] $<0.5$, respectively. There were 28 and 25 genes that showed imprinted expression in all IVF-derived placentas and brains, respectively (Fig. 1). Of these, eight genes showed abnormal allelic expression in at least two cloned placentas or brains (boxed in Fig. 1, note that *Dlk1* showed aberrant expression in both tissues). Importantly, *Gab1*, *Sfmbt2* and *Slc38a4* showed biallelic expression in all cloned placentas analyzed (Fig. 1).

### Imprinting status of placenta-specific imprinted genes in cloned mouse placentas

Two paternally expressed placenta-specific imprinted genes, *Gab1* and *Sfmbt2*, were expressed biallelically in all cloned mouse placentas (Fig. 1). *Gab1* is mainly expressed via a placenta-specific promoter in the normal placenta and has a novel DMR that is acquired after implantation (Fig 2A and Supplementary Material, Fig. S1). *Gab1* was not detected in cumulus cells and showed biallelic expression in Sertoli cells (Fig. 2B). *Gab1* showed biallelic expression in all cloned placentas analyzed (Fig. 2B) and was expressed at a level twice as high as in the IVF-derived placentas (Fig. 2C). Maternal allele-specific methylation was lost in the cloned mouse placentas (Fig. 2D).

*Sfmbt2* was also biallelically expressed in the cloned mouse placentas, and the expression level was twice as high as that of the IVF-derived placentas (Fig. 2E and F). Though there is no direct DNA methylation of the *Sfmbt2* promoter region, we previously reported the enrichment of H3K9me2 and H3K27me3 on the maternal allele (19). This allele-specific accumulation of H3K9me2 and H3K27me3 was not observed in the cumulus-derived cloned mouse placenta and cumulus cells (Fig. 2G).

Unlike *Gab1* and *Sfmbt2*, other placenta-specific imprinted genes (*Ppp1r9a*, *Ascl2*, *Tssc4* and *Slc22a3*) showed normal maternal allele-specific expression in the cloned mouse placentas (Supplementary Material, Fig. S2A). Brain-specific imprinted genes also showed normal allelic expression in the cloned mouse brains (*Blcap* in Fig. 1; *Grb10* and *Ube3a* in Supplementary Material, Fig. S2B).

### Consistent loss of *Slc38a4* imprinting in cloned mice

*Slc38a4* shows paternal allele-specific expression both in embryonic and extraembryonic tissues and has a germline DMR (21,22). *Slc38a4* also showed paternal expression in donor cells but was biallelically expressed in the cloned mouse brains and placentas (Fig. 3A). Increased expression of *Slc38a4* was observed both in the cloned mouse brains and, to a lesser extent, in placentas (Fig. 3B). Maternal allele-specific methylation of the *Slc38a4* DMR was observed in cumulus cells but not in the cumulus-derived cloned mouse brain (Fig. 3C). Demethylation of the *Slc38a4* DMR was already evident in the cloned blastocysts (Fig. 3D) and also observed in adult cloned mice (Supplementary Material, Fig. S3A). Interestingly, the paternally biased expression was not lost in embryos obtained from *Dnmt3L*-deficient and oocyte-specific *Dnmt3a/3b*-deficient female mice (Fig. 3E). It is already reported that the *Slc38a4* DMR is not methylated at all in *Dnmt3L*-deficient oocytes (23). These data strongly suggested that during oocyte growth, *de novo* DNA methylation was dispensable for imprinted expression of *Slc38a4*.

### Correlation between loss of *Dlk1-Dio3* imprinting and embryonic lethality of cloned mice

Five imprinted genes showed abnormal allelic expression in some cloned placentas or brains (Fig. 1). Among them, we focused on *Dlk1* because the *Dlk1-Dio3* domain (Fig. 4A) is important for normal mouse development (24). *Dlk1* showed normal paternal expression in the donor cells (Fig. 4B). Biallelic expression of *Dlk1* was observed in the placentas and fetuses