



(12) **EUROPEAN PATENT APPLICATION**  
 published in accordance with Art. 153(4) EPC

(43) Date of publication:  
**26.07.2017 Bulletin 2017/30**

(51) Int Cl.:  
**C12Q 1/68 (2006.01) C12M 1/00 (2006.01)**  
**C12M 1/34 (2006.01) C12N 15/09 (2006.01)**  
**G01N 33/53 (2006.01) G01N 37/00 (2006.01)**

(21) Application number: **15841913.5**

(22) Date of filing: **15.09.2015**

(86) International application number:  
**PCT/JP2015/076085**

(87) International publication number:  
**WO 2016/043170 (24.03.2016 Gazette 2016/12)**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**  
 Designated Extension States:  
**BA ME**  
 Designated Validation States:  
**MA**

(72) Inventors:  
 • **KOZONO, Satoko**  
**Kamakura-shi**  
**Kanagawa 248-8555 (JP)**  
 • **KONDOU, Satoshi**  
**Kamakura-shi**  
**Kanagawa 248-8555 (JP)**

(30) Priority: **16.09.2014 JP 2014187685**

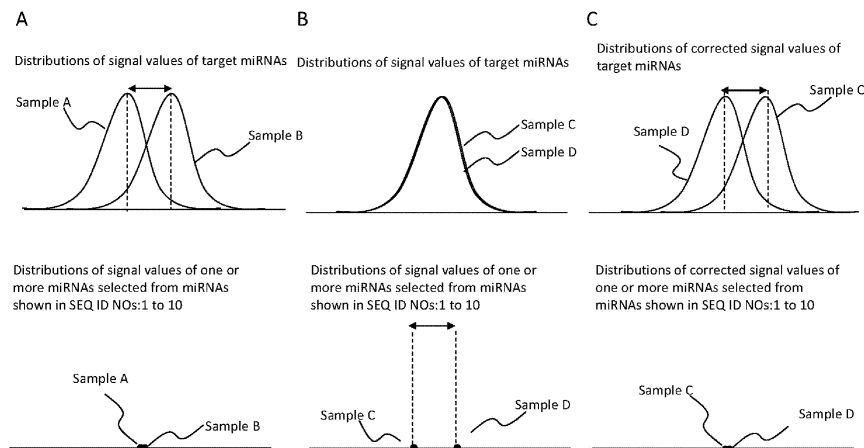
(74) Representative: **Prüfer & Partner mbB**  
**Patentanwälte · Rechtsanwälte**  
**Sohnckestraße 12**  
**81479 München (DE)**

(71) Applicant: **Toray Industries, Inc.**  
**Tokyo 103-8666 (JP)**

(54) **METHOD AND DEVICE FOR COMPARATIVE ANALYSIS OF miRNA EXPRESSION LEVEL**

(57) Disclosed is a method for comparative analysis of the expression levels of target miRNAs among a plurality of body fluid samples. In the method for comparative analysis of the present invention, the expression levels of the target miRNAs in each body fluid sample are corrected using the expression level(s) of a correcting endogenous miRNA(s) that is/are simultaneously meas-

ured with the expression levels of the target miRNAs in the sample. As the correcting endogenous miRNA(s), one or more miRNAs selected from specific 10 kinds of correcting endogenous miRNAs is/are used. By the present invention, comparative analysis of target miRNAs among body fluid samples can be carried out more accurately than by conventional techniques.



**Fig. 1**

**Description**

## TECHNICAL FIELD

5 **[0001]** The present invention relates to a method and a device for comparative analysis of the expression level(s) of a miRNA(s) contained in a plurality of body fluid samples.

## BACKGROUND ART

10 **[0002]** A miRNA (microRNA) is transcribed as an RNA (precursor) having a hairpin-like structure from genomic DNA. This precursor is cleaved by a particular enzyme, dsRNA cleavage enzyme (Drosha, Dicer) having RNase III cleavage activity, followed by conversion into a double-stranded form and then into single strands. It is thought that incorporation of the antisense strand, which is one of the strands, into a protein complex called RISC subsequently occurs, leading to its involvement in suppression of translation of mRNA. Thus, miRNA takes various forms in the stages after its transcription. Therefore, when targeting (detecting) a miRNA, various forms including the hairpin structure, double-stranded structure, and single-stranded structure need to be taken into account. A miRNA is consisting of an RNA of 15 to 25 bases, and the presence of miRNAs has been confirmed in various organisms.

15 **[0003]** In recent years, it has been suggested that a large amount of miRNAs are present in not only cells but also body fluids such as serum, plasma, urine, and spinal fluid, which are samples containing no cells, and that the expression levels of those miRNAs could be biomarkers for various diseases including cancers. As of June 2014, there are not less than 2500 kinds of miRNAs in human (miRBase release 20), and, when a highly sensitive assay system such as a DNA microarray is used, expression of more than 1000 kinds of miRNAs among them can be detected simultaneously in serum or plasma. Thus, studies are being carried out to find biomarkers in body fluids such as serum/plasma, urine, and spinal fluid using the DNA microarray method.

20 **[0004]** On the other hand, it is known that, in cases where gene expression analysis is carried out using a DNA microarray, the obtained data may include some errors depending on the sample, experimenter, and experimental conditions. Thus, data correction methods for correcting such errors have been devised.

25 **[0005]** The methods used for the correction are based on the principle that, when the expression data for a plurality of genes are treated as a single cluster to be regarded as a gene expression data group, the expression level in the whole gene expression data group is not different among any samples. Examples of such methods include the global normalization method, the quantile method, the lowess method, and the 75 percentile method.

30 **[0006]** There are also methods in which particular genes (for example, beta-actin and GAPDH) whose expression levels are the same among samples are used for correcting data from each sample such that the detected values of such genes become constant.

35 **[0007]** As methods for performing the correction such that the expression levels of particular miRNAs become constant, methods in which, among the non-coding RNAs expressed in samples, housekeeping RNAs (U1 snoRNA, U2 snoRNA, U3 snoRNA, U4 snoRNA, U5 snoRNA, U6 snoRNA, 5S rRNA, and 5.8S rRNA), whose expression levels are said to be constant among various samples, are used for the correction have been proposed (Patent Document 1 and Patent Document 2). That is, in Patent Document 1 and Patent Document 2, miRNA detection results are corrected by carrying out an operation in which the detected value of 5S rRNA, whose detection is carried out at the same time as the detection of miRNAs, becomes constant among all samples.

40 **[0008]** On the other hand, it is said that, in DNA microarray, a method in which only limited kinds of housekeeping RNAs are used to correct the data lacks accuracy since a large number of miRNAs are measured simultaneously. Thus, a method in which the expression levels of a plurality of mRNAs are measured at the same time as miRNAs, and the expression levels of these plurality of mRNAs are used for correcting the miRNAs has been proposed (Patent Document 3). In Patent Document 3, miRNA detection results are corrected by carrying out an operation in which the detected values of a plurality of mRNAs, whose detection is carried out at the same time as the detection of miRNAs, become constant among all samples.

45 **[0009]** In Non-patent Document 1, endogenous miRNAs were used for correction of the expression levels of miRNAs measured in mouse sera, and the efficiency of the correction was evaluated. "Endogenous miRNA" means a miRNA which is naturally present in the sample to be investigated (mouse serum, in the case of Non-patent Document 1) and derived from the organism from which the sample was provided. That is, in Non-patent Document 1, correction of miRNA expression levels in mouse sera was attempted using miR-16, miR-31, and miR-223 as examples of endogenous miRNAs for the correction. However, the expression levels of these miRNAs were not constant among individuals, and thus Non-patent Document 1 does not adopt the method in which correction of the whole data is carried out based on the expression levels of the endogenous miRNAs but proposes use of an externally added RNA as a reference substance for data correction. In short, Non-patent Document 1 concluded that correction of miRNA expression levels using these endogenous miRNAs is inappropriate.

**[0010]** Non-patent Document 2 proposes use of the combination of let-7d, let-7g, and let-7i as correcting endogenous miRNAs in detection of miRNAs in human serum by quantitative RT-PCR or sequencing. That is, by carrying out an operation such that the expression levels of let-7d, let-7g, and let-7i become constant among samples, the expression levels of the miRNAs to be measured were corrected.

PRIOR ART DOCUMENTS

PATENT DOCUMENTS

**[0011]**

Patent Document 1: JP 2007-75095 A  
 Patent Document 2: JP 2007-97429 A  
 Patent Document 3: JP 2014-7995 A

NON-PATENT DOCUMENTS

**[0012]**

Non-patent Document 1: Roberts, T. C., 2014, PLoS ONE, vol. 9 (2), e89237  
 Non-patent Document 2: Chen, X. et al., 2013, PLoS ONE, vol. 8 (11), e79652

SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

**[0013]** As described above, when detecting gene expression, methods in which the expression levels of the genes to be detected are corrected using the expression levels of "housekeeping genes", which show only small variations of the expression levels among samples, are generally employed. Examples of especially widely known housekeeping genes include ACTB and GAPDH. However, since the nucleic acid lengths and the absolute expression levels of these genes are largely different from those of miRNAs, it is not preferred to use these genes for correction of the expression levels of miRNAs.

**[0014]** Thus, conventionally, when the expression levels of miRNAs are to be detected, housekeeping RNAs such as U1 snoRNA, U2 snoRNA, U3 snoRNA, U4 snoRNA, U5 snoRNA, U6 snoRNA, 5S rRNA, and 5.8S rRNA have been mainly used for the correction. However, in cases where the expression levels of miRNAs in body fluids are measured and the resulting data are to be corrected, the expression levels of U1 snoRNA, U2 snoRNA, U3 snoRNA, U4 snoRNA, U5 snoRNA, U6 snoRNA, 5S rRNA, and 5.8S rRNA can be hardly detected in body fluids such as serum, plasma, urine, and spinal fluid, which are samples containing no cells, since these RNAs are present in nuclei and cytoplasm. It is therefore said that these cannot be used as indicators for detection and correction of the expression levels of miRNAs in body fluids.

**[0015]** In detection of the expression levels of miRNAs by quantitative RT-PCR, attempts are being made to find "housekeeping miRNAs in body fluid" showing no difference in the expression level among samples, by individually analyzing miRNAs present in serum and plasma. As described above, Non-patent Document 2 proposes use of the combination of let-7d, let-7g, and let-7i for detection of miRNAs in human serum by quantitative RT-PCR or sequencing. However, whether or not these miRNAs function as correcting endogenous miRNAs in the DNA microarray method is not clear. That is, in analysis of miRNAs in a body fluid using a DNA microarray, there has been no effective correction method using particular correcting endogenous miRNAs.

**[0016]** On the other hand, in cases where correction of measured values is to be carried out in DNA microarray, the correction is generally carried out using the total expression level (signal value) of a large number of (several hundreds to several ten thousands) genes that can be measured with the DNA microarray. Examples of such a method include the global normalization method and the quantile normalization method.

**[0017]** By the way, in cases where a reference value obtained from the expression levels of miRNAs is used as an indicator, there is a problem: how many miRNAs should be used as the original data for construction of the reference value to make it possible to obtain a universal reference value for the correction? In particular, in DNA microarrays for specific uses in which the number of miRNAs to be measured is as small as several to several tens, application of the global normalization method or the like is highly likely to result in insufficient correction results. In the method disclosed by Patent Document 3, the expression levels of a large number of mRNAs in a sample are measured to provide a reference value for the correction. The method is effective for RNAs derived from tissues and cells. However, the

expression levels of mRNAs in body fluids are extremely low, and therefore it is difficult to use them as indicators for correction of the expression levels of miRNAs in body fluids.

**[0018]** Thus, in measurement and comparative analysis of the expression levels of miRNAs contained in a plurality of body fluid samples, no effective method for correction of the measured values has been discovered. In particular, there is no method that enables effective correction by DNA microarray using a small number of endogenous miRNAs for correction. The present invention aims to solve such problems. In particular, the present invention aims to solve important problems in practical use of DNA microarrays for measurement of a relatively small number of miRNAs, such as those for uses including tests and diagnoses.

#### MEANS FOR SOLVING THE PROBLEMS

**[0019]** As a result of intensive study for solving the above problems, the present inventors completed the following inventions.

(1) A method for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, the method comprising:

a measurement step of simultaneously measuring the expression level(s) of a target miRNA(s) and the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 in the respective body fluid samples;

a representative-value-obtaining step of obtaining a representative value for each of said plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

a correction-factor-obtaining step of obtaining, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained in the representative-value-obtaining step; and

a correction step of correcting the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample.

(2) The method according to (1), wherein the correction in the correction step is carried out as follows:

(a) in the case where a value calculated by subtracting the reference value from the representative value is obtained as a correction factor in the correction-factor-obtaining step, the correction factor is subtracted from the measured value(s) of the expression level(s) of the target miRNA(s);

(b) in the case where a value calculated by subtracting the representative value from the reference value is obtained as a correction factor in the correction-factor-obtaining step, the correction factor is added to the measured value(s) of the expression level(s) of the target miRNA(s);

(c) in the case where a value calculated by dividing the representative value by the reference value is obtained as a correction factor in the correction-factor-obtaining step, the measured value(s) of the expression level(s) of the target miRNA(s) is/are divided by the correction factor; or

(d) in the case where a value calculated by dividing the reference value by the representative value is obtained as a correction factor in the correction-factor-obtaining step, the measured value(s) of the expression level(s) of the target miRNA(s) is/are multiplied by the correction factor.

(3) The method according to (1) or (2), wherein the reference value is a fixed value arbitrarily defined in connection with the expression level(s) of the correcting endogenous miRNA(s), or a representative value for the expression level(s) of the correcting endogenous miRNA(s) obtained for a first body fluid sample, the first body fluid sample being arbitrarily selected from the plurality of body fluid samples.

(4) The method according to any one of (1) to (3), wherein the measurement step comprises carrying out hybridization by bringing nucleic acid probes for capturing a plurality of target miRNAs and a probe(s) for capturing the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, the probes being immobilized on a support, into contact with a nucleic acid sample derived from each of the body fluid samples, the nucleic acid sample being labeled with a labeling substance, and obtaining the expression levels of the target miRNAs and the one or more correcting endogenous miRNAs as signal intensity measurement values.

(5) The method according to any one of (1) to (4), wherein the body fluid sample is blood, serum, or plasma.

(6) The method according to any one of (1) to (5), wherein the representative value is an average or a median

expressed as a logarithmic value calculated from the measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10.

(7) A miRNA expression analysis device for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, the device comprising:

memory means which memorizes measured values of the expression level(s) of a target miRNA(s) and measured values of the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, both of the measured values being obtained by measurements on a plurality of body fluid samples;

representative-value-obtaining means which obtains a representative value for each of the plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

correction-factor-obtaining means which obtains, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained by the representative-value-obtaining means;

correction means which corrects the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample; and

output means which outputs a result(s) of comparison of the expression level(s) of the target miRNA(s) among at least two body fluid samples based on the corrected expression level(s) of the target miRNA(s).

(8) The device according to (7), wherein the correction means carries out the correction as follows:

- (a) in the case where a value calculated by subtracting the reference value from the representative value is obtained as a correction factor in the correction-factor-obtaining means, the correction factor is subtracted from the measured value(s) of the expression level(s) of the target miRNA(s);
- (b) in the case where a value calculated by subtracting the representative value from the reference value is obtained as a correction factor in the correction-factor-obtaining means, the correction factor is added to the measured value(s) of the expression level(s) of the target miRNA(s);
- (c) in the case where a value calculated by dividing the representative value by the reference value is obtained as a correction factor in the correction-factor-obtaining means, the measured value(s) of the expression level(s) of the target miRNA(s) is/are divided by the correction factor; or
- (d) in the case where a value calculated by dividing the reference value by the representative value is obtained as a correction factor in the correction-factor-obtaining means, the measured value(s) of the expression level(s) of the target miRNA(s) is/are multiplied by the correction factor.

(9) The device according to (7) or (8), wherein the measured values of the expression level(s) of the target miRNA(s) contained in the plurality of body fluid samples and the measured values of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 that are memorized in the memory means are values obtained by carrying out hybridization by bringing probes for capturing a plurality of target miRNAs and a probe(s) for capturing the one or more correcting endogenous miRNAs, said probes being immobilized on a support, into contact with a nucleic acid sample derived from each of the body fluid samples, said nucleic acid sample being labeled with a labeling substance, and obtaining the expression levels of the target miRNAs and the one or more correcting endogenous miRNAs as signal intensity measurement values.

(10) A program(s) for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, the program(s) causing one or more computers to execute:

a measurement step of simultaneously measuring the expression level(s) of a target miRNA(s) and the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 in the respective body fluid samples;

a representative-value-obtaining step of obtaining a representative value for each of the plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

a correction-factor-obtaining step of obtaining, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the repre-

sentative value for said each body fluid sample obtained in the representative-value-obtaining step; and a correction step of correcting the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample.

(11) A program(s) for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, the program(s) causing one or more computers to function as:

memory means which memorizes measured values of the expression level(s) of a target miRNA(s) and measured values of the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, the target miRNA(s) and the one or more correcting endogenous miRNAs being contained in each of a plurality of body fluid samples; representative-value-obtaining means which obtains a representative value for each of the plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10; correction-factor-obtaining means which obtains, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained by the representative-value-obtaining means; correction means which corrects the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample; and output means which outputs a result(s) of comparison of the expression level(s) of the target miRNA(s) among at least two body fluid samples based on the corrected expression level(s) of the target miRNA(s).

(12) A computer-readable recording medium in which the program(s) according to (10) or (11) is/are recorded.

(13) A chip for analysis of miRNA expression, comprising a support on which probes for capturing a plurality of target miRNAs and a probe(s) for capturing one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 are immobilized.

## EFFECT OF THE INVENTION

**[0020]** According to the present invention, when comparative analysis of the expression level(s) of a target miRNA(s) contained in a plurality of body fluid samples is carried out, especially when the expression levels of a large number of miRNAs are measured using a microarray or the like and compared among samples, correction of the expression levels of the miRNAs can be carried out more accurately compared to conventional methods. The present invention enables more accurate comparative analysis of a target miRNA(s) among samples.

## BRIEF DESCRIPTION OF THE DRAWINGS

### **[0021]**

Fig. 1 is a conceptual diagram illustrating the method of the present invention.

Fig. 2 is a block diagram illustrating an outline of the constitution of the analysis device of the present invention.

Fig. 3 is an example of the flow chart of the correction process for the expression levels of target miRNAs according to the present invention.

Fig. 4-1 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Example 2. A: before the correction; B: corrected by hsa-miR-6085; C: corrected by hsa-miR-1227-5p; D: corrected by hsa-miR-2861; E: corrected by hsa-miR-149-3p; F: corrected by hsa-miR-4463; G: corrected by hsa-miR-4508.

Fig. 4-2 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Example 2. A: before the correction; H: corrected by hsa-miR-6090; I: corrected by hsa-miR-6775-5p; J: corrected by hsa-miR-6803-5p; K: corrected by hsa-miR-5787.

Fig. 5 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Example 2.

Fig. 6 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Comparative Example 1.

Fig. 7 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Comparative Example 2.

Fig. 8-1 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured

and corrected by the method described in Example 4. A: before the correction; B: corrected by the combination of hsa-miR-149-3p, hsa-miR-1227-5p, and hsa-miR-2861; C: corrected by the combination of hsa-miR-149-3p, hsa-miR-1227-5p, and hsa-miR-4508; D: corrected by the combination of hsa-miR-149-3p, hsa-miR-1227-5p, and hsa-miR-4463; E: corrected by the combination of hsa-miR-149-3p, hsa-miR-2861, and hsa-miR-4508.

Fig. 8-2 shows histograms of miRNA signal values obtained by DNA microarray detection, which were measured and corrected by the method described in Example 4. A: before the correction; F: corrected by the combination of hsa-miR-149-3p, hsa-miR-2861, and hsa-miR-4463; G: corrected by the combination of hsa-miR-1227-5p, hsa-miR-2861, and hsa-miR-4508; H: corrected by the combination of hsa-miR-1227-5p, hsa-miR-2861, and hsa-miR-4463; I: corrected by the combination of hsa-miR-1227-5p, hsa-miR-4508, and hsa-miR-4463.

## MODE FOR CARRYING OUT THE INVENTION

**[0022]** First, the concept of the correction method for a target miRNA(s) that is carried out by the present invention is described based on Fig. 1. Fig. 1 shows histograms of signal values that are measured values of the expression levels of miRNAs. The histograms schematically illustrate detection results obtained by labeling RNA in a body fluid sample and performing detection using a microarray on which probes for capturing a plurality of kinds of target miRNAs (hereinafter also referred to as "probes for capturing target miRNAs") and a probe(s) for capturing one or more miRNAs (hereinafter also referred to as "correcting endogenous miRNA(s)") selected from the miRNAs shown in SEQ ID NOs:1 to 10 (hereinafter also referred to as "correcting endogenous miRNA-capture probe(s)") are immobilized. The probes for capturing the target miRNAs and the probes for capturing the correcting endogenous miRNAs are hereinafter also collectively referred to as "miRNA capture probes" or, simply, "probes".

**[0023]** Fig. 1A shows histograms illustrating results of DNA microarray analysis of target miRNAs in samples extracted from a body fluid sample A and a body fluid sample B, respectively. Histograms of signal values obtained from a plurality of target miRNA capture probes loaded on the microarray, and histograms of signal values obtained from a correcting endogenous miRNA-capture probe(s) of the present invention, are shown. The body fluid sample A and the body fluid sample B show miRNA histograms largely shifting from each other. Thus, it can be interpreted that there is a large difference in the expression levels of miRNAs between the samples. On the other hand, another interpretation is possible: the difference could be due to an experimental error. It is impossible to determine which interpretation is correct based merely on the histograms of the miRNAs.

**[0024]** Here, assuming that a correcting endogenous miRNA(s) whose abundance in the body fluid is constant is/are present, the signal value(s) of the correcting endogenous miRNA(s) should be the same between the samples.

**[0025]** In Fig. 1A, the distribution of the histogram of the signal value(s) obtained from the correcting endogenous miRNA-capture probe(s) is almost the same between the body fluid sample A and the body fluid sample B. That is, it can be determined that the body fluid sample A and the body fluid sample B were correctly subjected to the experiment, and hence that there is no experimental error. In such a case, it follows that there is a large difference in the expression levels of the miRNAs between the body fluid samples A and B, and that correction of the signal values of the miRNAs is unnecessary for comparison between the body fluid samples.

**[0026]** Fig. 1B schematically shows results of analysis of a body fluid sample C and a body fluid sample D using a DNA microarray. Signal values obtained from target miRNA capture probes and signal values obtained from a correcting endogenous miRNA-capture probe(s) are shown.

**[0027]** Histograms of miRNAs from the body fluid sample C and the body fluid sample D show similar distributions. On the other hand, the histograms of the signal values obtained from the correcting endogenous miRNA-capture probe(s) show a large shift between the body fluid sample C and the body fluid sample D. Thus, it can be understood that the detection results from the body fluid sample C and the body fluid sample D include an experimental error due to some reason. In such a case, the signal values of the miRNAs need to be appropriately corrected for comparison between the body fluid samples C and D.

**[0028]** Histograms after correction of the signal values of the target miRNAs in accordance with the present invention are shown in Fig. 1C. The specific method of the correction is as described later. The data from the body fluid sample C were corrected such that the histograms of the signal values obtained from the correcting endogenous miRNA-capture probe(s) generated from the body fluid sample C and the body fluid sample D are matched to each other. As a result of this correction, the body fluid sample C and the body fluid sample D come to have the same histogram of the signal values obtained from the correcting endogenous miRNA-capture probe(s), and the histograms of the signal values from the target miRNA-capture probes corrected using the same correction factor come to be largely shifted from each other. That is, it follows that there is a large difference in the expression levels of the target miRNAs also between the samples C and D.

**[0029]** Although a method for correcting signal values of miRNAs between two body fluid samples was shown here, the number of body fluid samples to be compared is not limited to two, and the comparison may be carried out among an unlimited number of body fluid samples. For example, in cases where the comparison is carried out among three or

more body fluid samples, a signal value obtained from a correcting endogenous miRNA-capture probe is preliminarily hypothesized to be a fixed value (constant), and the difference or the ratio of the signal value obtained from the correcting endogenous miRNA-capture probe in each of the body fluid samples against this constant is then calculated. By adding the difference to, or subtracting the difference from, the signal values of the target miRNAs in the respective body fluid samples, or by multiplying or dividing the signal values of the target miRNAs in the respective body fluid samples by the reciprocal of the ratio, correction can be easily carried out among a large number of samples.

**[0030]** In the present invention, comparative analysis of the expression level(s) of a target miRNA(s) is carried out among a plurality of samples. The number of the samples may be two, or may be three or more.

**[0031]** "miRNA" is a non-coding RNA (ncRNA), which means a short-chain RNA produced in the body having a chain length of about 15 to 25 bases. It is thought to have a function to regulate expression of mRNA. A miRNA is transcribed as an RNA (precursor) having a hairpin-like structure from genomic DNA. This precursor is cleaved by a particular enzyme, dsRNA cleavage enzyme (Drosha, Dicer) having RNase III cleavage activity, and converted into a double-stranded form and then into single strands. It is thought that incorporation of the antisense strand, which is one of the strands, into a protein complex called RISC subsequently occurs, leading to its involvement in suppression of translation of mRNA. Thus, miRNA takes various forms in the stages after its transcription. Therefore, when targeting (detecting) a miRNA, various forms including the hairpin structure, double-stranded structure, and single-stranded structure need to be taken into account. The presence of miRNAs has been confirmed in various organisms.

**[0032]** The samples to which the present invention is applicable are body fluid samples separated from living bodies, and examples of the body fluid samples include, but are not limited to, body fluids such as blood, serum, plasma, urine, spinal fluid, saliva, swab, and various tissue fluids. The plurality of samples to be subjected to the comparative analysis may be a plurality of samples derived from different body fluids, or may be a plurality of body fluid samples derived from the same kind of body fluid separated from different living bodies.

**[0033]** RNA is extracted from these samples, and the extracted RNA is used for measuring the expression levels of miRNAs. Methods for extracting such RNA are known (for example, the method by Favaloro et al. (Favaloro et al., Methods Enzymol. 65: 718-749 (1980))), and various kits for such methods are commercially available (for example, miRNeasy, manufactured by QIAGEN; and "3D-Gene" RNA extraction reagent from liquid sample, manufactured by Toray Industries, Inc.).

**[0034]** The term "endogenous" as used in the present invention means being naturally present in a sample rather than being present due to artificial addition into the sample. For example, "endogenous miRNA" means a miRNA which is naturally present in the sample and derived from the organism from which the sample was provided.

#### <Measurement Step>

**[0035]** In the method for comparative analysis of the present invention, measurement of the expression levels of a plurality of kinds of target miRNAs in the samples is carried out at the same time as measurement of the expression level(s) of one or more correcting endogenous miRNAs. As described later, the expression level(s) of the correcting endogenous miRNA(s) is/are used for calculation of a correction factor for correcting the expression levels of the target miRNAs. The simultaneous measurement of the expression levels of the plurality of miRNAs can be carried out by, for example, hybridization assay using an array chip such as a microarray in which probes that specifically bind to the subject miRNAs are immobilized on a support. In the present invention, an array chip comprising a support on which a plurality of target miRNA-capture probes and one or more correcting endogenous miRNA-capture probes are immobilized may be used.

**[0036]** The "capture probe" or the "probe for capturing" means a substance capable of directly or indirectly, preferably directly, and selectively binding to the miRNA to be captured. Representative examples of such a probe include nucleic acids, proteins, sugars, and other antigenic compounds. In the present invention, nucleic acid probes may be preferably used. Examples of the nucleic acids that may be used include not only DNA and RNA, but also nucleic acid derivatives such as PNA (peptide nucleic acid) and LNA (Locked Nucleic Acid). In cases of nucleic acids, the "derivatives" herein means chemically modified derivatives such as labeled derivatives prepared using a fluorophore or the like; and derivatives comprising a modified nucleotide (a nucleotide containing halogen, or containing a group such as alkyl including methyl; alkoxy including methoxy; thio; or carboxymethyl; a nucleotide that has undergone reconstruction of the base, saturation of the double bonds, deamination, substitution of an oxygen molecule(s) into a sulfur molecule(s); and/or the like).

**[0037]** From the viewpoint of securing stability and specificity in the hybridization, the chain length of the nucleic acid probe is preferably not less than the length of the miRNA to be detected. Usually, when the chain length is about 17 to 25 bases, the probe can sufficiently exert the selective binding capacity to the subject miRNA. Such an oligonucleic acid probe having a short chain length can be easily prepared by a well-known chemical synthesis method or the like.

**[0038]** The nucleic acid probe preferably has the base sequence completely complementary to the miRNA to be detected. However, even in cases where there is a partial difference, the nucleic acid probe can be used as the capture

probe as long as the nucleic acid probe has a base sequence which is homologous enough to allow hybridization with the subject miRNA under stringent conditions.

5 [0039] The stringency in the hybridization is known to be a function of the temperature, the salt concentration, the chain length of the probe, the GC content of the nucleotide sequence of the probe, and the concentration of the chaotropic agent in the hybridization buffer. As the stringent conditions, those described in Sambrook, J. et al. (1998) *Molecular Cloning: A Laboratory Manual* (2nd ed.), Cold Spring Harbor Laboratory Press, New York, and the like may be employed. The stringent temperature condition is not less than about 30°C. Examples of other conditions include the hybridization time, the concentration of the washing agent (for example, SDS), and the presence or absence of carrier DNA. By combining these conditions, various stringencies can be set. Those skilled in the art can appropriately determine conditions for obtaining the function of the capture probe provided for detection of a desired sample RNA.

10 [0040] Sequence information of miRNA can be obtained from databases such as GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). Sequence information of miRNA can also be obtained from the website of miRBase (<http://www.mirbase.org/>). The correcting endogenous miRNA-capture probe(s) and the target miRNA-capture probe(s) can be designed based on information available from these sites.

15 [0041] The number of the miRNA capture probes immobilized on the support is not limited. For example, in the measurement of the expression levels of miRNAs, the number of the miRNA capture probes immobilized on the support may be a number which comprehensively covers all known miRNAs whose sequences have been identified. Or, the number of the miRNA capture probes immobilized on the support may be a desired number.

20 [0042] As the support on which the capture probes are to be aligned and immobilized, those which are the same as the supports used in known microarrays, macroarrays and the like may be used. Examples of the support include slide glasses, membranes, and beads. The support described in JP 4244788 B, which has a plurality of protruded portions on its surface, may also be used. Examples of the material of the support include, but are not limited to, inorganic materials such as glass, ceramic, and silicone; and polymers such as polyethylene terephthalate, cellulose acetate, polycarbonate, polystyrene, polymethyl methacrylate, and silicone rubber.

25 [0043] Examples of the known methods for immobilizing capture probes on a support include methods in which oligo-DNAs are synthesized on the surface of the support, and methods in which oligo-DNAs preliminarily synthesized are added dropwise to the surface of the support and then fixed thereon.

30 [0044] Examples of the former methods include the method of Ronald et al. (US 5705610 B), the method of Michel et al. (US 6142266 B), and the method of Francesco et al. (US 7037659 B). In these methods, an organic solvent is used in the DNA synthesis reaction, and therefore the material of the support is preferably a material resistant to organic solvents. In the method by Francesco et al., the DNA synthesis is controlled by irradiation with light from the back side of the support, and therefore the material of the support needs to be a light-transmitting material.

35 [0045] Examples of the latter methods include the method of Hirota et al. (JP 3922454 B) and methods using a spotter. Examples of the spotting method include the pin method, which is based on mechanical contact of a pin tip with a solid phase; the ink jet method, which utilizes the principle of ink jet printers; and the capillary method, which uses a capillary. If necessary, after the spotting treatment, posttreatment such as cross-linking by UV irradiation and/or surface blocking is carried out. For allowing immobilization of the oligo-DNAs through covalent bonds on the surface of the surface-treated support, functional groups such as amino groups and/or SH groups are introduced to the termini of the oligo-DNAs. The surface modification of the support is usually carried out by treatment with a silane coupling agent having an amino group and/or the like.

40 [0046] The hybridization with miRNA capture probes immobilized on the support is carried out by preparing a nucleic acid sample (nucleic acid sample derived from a body fluid sample) labeled with a labeling substance from RNA extracted from a body fluid sample, and bringing the resulting labeled nucleic acid sample with the probes. Examples of the "nucleic acid sample derived from a body fluid sample" include not only RNA extracted from a body fluid sample, but also cDNA prepared by reverse transcription reaction from the RNA, and crRNA. The labeled nucleic acid sample derived from the body fluid sample may be a sample prepared by directly or indirectly labeling the sample RNA with a labeling substance, or a sample prepared by directly or indirectly labeling cDNA or crRNA prepared from the sample RNA with a labeling substance.

45 [0047] Examples of the method for binding the labeling substance to the nucleic acid sample derived from the body fluid sample include methods in which the labeling substance is bound to the 3'-end of the nucleic acid sample, methods in which the labeling substance is bound to the 5'-end of the nucleic acid sample, and methods in which a nucleotide to which the labeling substance is bound is incorporated into the nucleic acid. In the methods in which the labeling substance is bound to the 3'-end and the methods in which the labeling substance is bound to the 5'-end, enzymatic reaction may be used. In the enzymatic reaction, T4 RNA Ligase, Terminal Deoxitidil Transferase, Poly A polymerase, or the like may be used. Any of the labeling methods may be carried out by reference to the methods described in "Shao-Yao Ying (ed.), *miRNA Experimental Protocols*, Yodosha Co., Ltd. (2008)". Various kits for directly or indirectly binding labeling substances to RNA termini are commercially available. Examples of kits for directly or indirectly binding a labeling substance to the 3'-end include "3D-Gene" miRNA labeling kit (Toray Industries, Inc.), miRCURY miRNA HyPower

labeling kit (Exiqon), NCode miRNA Labeling system (Life Technologies), and FlashTag Biotin RNA Labeling Kit (Genisphere).

5 [0048] In addition to the above, the same method as the conventional methods may also be used. That is, cDNA or cRNA may be synthesized from sample RNA in the presence of labeled deoxyribonucleotides or labeled ribonucleotides to prepare cDNA or cRNA in which a labeled substance is incorporated, and the resulting cDNA or cRNA may be hybridized with the probes on the array.

[0049] In the present invention, a plurality of samples are used. The same labeling substance may be used for all of the samples.

10 [0050] Examples of labeling substances that may be used in the present invention include various labeling substances that are also used in known microarray analyses. Specific examples of the labeling substances include, but are not limited to, fluorescent dyes, phosphorescent dyes, enzymes, and radioisotopes. Fluorescent dyes are preferred since they can be easily measured and their signals can be easily detected. Specific examples of the fluorescent dyes include, but are not limited to, known fluorescent dyes such as Cyanine (Cyanine 2), aminomethylcoumarin, fluorescein, indocarbocyanine (Cyanine 3), Cyanine 3.5, tetramethylrhodamine, rhodamine red, Texas red, indocarbocyanine (Cyanine 5), Cyanine 5.5, Cyanine 7, and Oyster.

15 [0051] As the labeling substance, luminescent semiconductor particles may also be used. Examples of such semiconductor particles include cadmium selenium (CdSe), cadmium tellurium (CdTe), indium gallium phosphide (InGaP), and silver indium zinc sulfide (AgInZnS).

20 [0052] The thus labeled nucleic acid sample derived from a body fluid sample is brought into contact with the miRNA capture probes on the support to allow hybridization of the nucleic acid sample with the probes. This hybridization step may be carried out in completely the same manner as the conventional hybridization step. The reaction temperature and the reaction time are appropriately selected depending on the chain length of the nucleic acid to be subjected to the hybridization. In cases of nucleic acid hybridization, the hybridization is usually carried out at about 30°C to 70°C for 1 minute to ten and several hours. After the hybridization and the washing, the signal intensity from the labeling substance in each area where each probe is immobilized on the substrate is detected. The detection of the signal intensity is carried out using an appropriate signal reader which is selected depending on the type of the labeling substance. When a fluorescent dye is used as the labeling substance, a fluorescence microscope or a fluorescence scanner may be used.

25 [0053] The detected signal value is compared with the surrounding noise. More specifically, the signal value obtained from the probe-immobilized area and the signal value obtained from a position other than the probe-immobilized area are compared with each other, and, in cases where the former value is higher than the latter value, the signal intensity is regarded as being detected (effectively judged positive).

30 [0054] In cases where the background noise is included in the detected signal value, the background noise may be subtracted from the detected signal value. The surrounding noise may be regarded as the background noise, and may be subtracted from the detected signal value. In addition, the method described in "Wataru Fujibuchi and Katsuhisa Horimoto (eds.), Microarray data statistical analysis protocols, Yodosha Co., Ltd. (2008)" may be used.

35 [0055] According to the above-described method, measured values of the expression levels of the correcting endogenous miRNA(s) and the target miRNA(s) are obtained as measured values of signal intensities.

40 <Representative-Value-Obtaining Step>

[0056] Subsequently, in the analysis method of the present invention, a representative value, preferably a representative value expressed as a logarithmic value, is obtained from the measured value(s) of the expression level(s) of the correcting endogenous miRNA(s) in each of the samples (representative-value-obtaining step). The "logarithmic value" in the present invention means a value converted to a logarithm with base 2.

45 [0057] When a plurality of correcting endogenous miRNAs are used, the average or the median, preferably the average or the median expressed as a logarithmic value, calculated from the measured values of the expression levels of the plurality of correcting endogenous miRNAs, is employed as the representative value. When only one correcting endogenous miRNA is used, the measured value, preferably the logarithmic value of the measured value, of the expression level of the correcting endogenous miRNA may be employed as the representative value as it is. Or, as described later, when an array chip on which a plurality of capture probes are spotted for one kind of correcting endogenous miRNA is used, the average or the median, preferably the average or the median expressed as a logarithmic value, may be calculated from the signal measurement values from these plurality of spots (probe-immobilized areas), to provide the representative value.

50 [0058] The "average expressed as a logarithmic value" means the average calculated from the logarithmic values obtained by converting measured values of the expression levels of the plurality of correcting endogenous miRNAs (for example, measured values of signal intensities obtained using a microarray) to logarithms with base 2. In cases where the representative value is the median, the "median expressed as a logarithmic value" means the median of the logarithmic

values obtained by converting measured values of the expression levels of the plurality of correcting endogenous miRNAs (for example, measured values of signal intensities obtained using a microarray) to logarithms with base 2, or means the logarithmic value obtained by conversion of the median of measured values of the expression levels of the correcting endogenous miRNAs to a logarithm with base 2. In cases of the median, the same value can be obtained irrespective

of whether the conversion of the measured values to logarithms is carried out in advance or not.  
**[0059]** The average or the median may be calculated using all of the measured values of the plurality of correcting endogenous miRNAs that were actually measured, or may be calculated using only a part of the measured values extracted from the measured values of the plurality of correcting endogenous miRNAs. For example, the average or the median may be calculated using the measured values of all of the correcting endogenous miRNAs obtained from the correcting endogenous miRNA-capture probes loaded on a microarray, or, a part of the entire correcting endogenous miRNA-capture probes (for example, in the case where 10 kinds of correcting endogenous miRNA-capture probes are loaded on a microarray, 3 out of 10 probes) may be extracted and used to calculate the average or the median. The measured values used in calculation of the average or the median may also be measured values obtained from a plurality of capture probes spotted on a plurality of spots for measuring one kind of correcting endogenous miRNA. For example, only a probe-immobilized area(s) in which a correcting endogenous miRNA-capture probe(s) is/are immobilized that is/are effectively judged positive across all the samples to be subjected to the comparative analysis may be extracted and used to obtain a representative value of the correcting endogenous miRNA(s).

**[0060]** Subsequently, using the representative value for the correcting endogenous miRNA(s) obtained in the representative-value-obtaining step for each sample and a reference value arbitrarily set for the expression level(s) of the correcting endogenous miRNA(s), a correction factor to be used for correction of the expression level(s) of the target miRNA(s) is obtained (correction-factor-obtaining step). To the correction-factor-obtaining step, the correction-factor-obtaining step-1 or the correction-factor-obtaining step-2 described below may be applied.

<Correction-Factor-Obtaining Step-1>

**[0061]** The correction-factor-obtaining step-1 is a method utilizing the difference between the representative value of the correcting endogenous miRNA(s) and the reference value. To this step, the following 1-1. reference sample-obtaining method or 1-2. fixed-value correction method may be applied.

1-1. Reference Sample-Obtaining Method

**[0062]** One sample (first sample) is arbitrarily selected from a plurality of body fluid samples in which miRNAs are to be detected, which sample is used as a "reference sample". The remaining one or more samples (subsequent sample(s)) are provided as a "sample(s) to be corrected".

**[0063]** In the present description, the term "subsequent sample(s)" includes the second sample. For example, in cases where the number of the plurality of samples to be compared is two, the sample to be corrected is only the second sample, and, in cases where the number of the plurality of samples to be compared is three, there are two samples to be corrected, that is, the second sample and the third sample.

**[0064]** In this method, the representative value of the correcting endogenous miRNA(s) in the reference sample is used as a "reference value". The difference between the reference value and the representative value of the correcting endogenous miRNA(s) in a certain one of the subsequent sample(s) (sample(s) to be corrected) is used as the correction factor for the certain one of the subsequent sample(s). Thus, the number of the correction factors obtained is the same as the number of the samples to be corrected.

**[0065]** More specifically, the correction factor is calculated according to Equation 1 or Equation 1'.

$$c_{1-1} = (\text{representative value of correcting endogenous miRNA(s) in reference sample (reference value)}) \\ - (\text{representative value of correcting endogenous miRNA(s) in sample to be corrected}) \dots \text{Equation 1}$$

$c_{1-1}' =$  (representative value of correcting endogenous miRNA(s) in sample to be corrected)

5

- (representative value of correcting endogenous miRNA(s) in reference sample (reference value)) ... Equation 1'

10 **[0066]** For example, in cases where measurement of the expression levels is carried out using a microarray, and the average is used as the representative value of the correcting endogenous miRNA(s), the correction factor for the sample to be corrected can be calculated according to Equation 2 or Equation 2'.

15 
$$c_{1-1} = \frac{1}{n} \sum_{j=1}^n \log_2 A_j - \frac{1}{n} \sum_{j=1}^n \log_2 X_j \quad \dots \text{Equation 2}$$

20 
$$c_{1-1}' = \frac{1}{n} \sum_{j=1}^n \log_2 X_j - \frac{1}{n} \sum_{j=1}^n \log_2 A_j \quad \dots \text{Equation 2'}$$

25 **[0067]** Here, in Equation 2 and Equation 2',

n represents the total number of probe-immobilized area(s) for capturing the correcting endogenous miRNA(s) on the support;

A<sub>j</sub> represents the signal measurement value from the probe-immobilized area for capturing the jth (1 ≤ j ≤ n) correcting endogenous miRNA in the reference sample; and

30 X<sub>j</sub> represents the signal measurement value from the probe-immobilized area for capturing the jth (1 ≤ j ≤ n) correcting endogenous miRNA in the second sample.

**[0068]** In cases where the probe(s) and the correcting endogenous miRNA(s) have a one-to-one relationship, n is equal to the number of the correcting endogenous miRNA(s) targeted by the correcting endogenous miRNA-capture probe(s) on the support.

35 **[0069]** In Equation 2 and Equation 2', n', the total number of probe-immobilized areas for capturing the correcting endogenous miRNA(s) that were effectively judged positive across all the samples to be compared, may be used instead of n.

40 1-2. Fixed-Value Correction Method

**[0070]** This method preliminarily assumes that the representative value of the correcting endogenous miRNA(s) is constant among all the samples. That is, the difference between a fixed value and the representative value of the correcting endogenous miRNA(s) derived from each of the samples is obtained, and this difference is utilized as the correction factor. In this method, the fixed value is used as the "reference value". Thus, the number of the correction factors obtained is the same as the number of the samples to be corrected. In this case, the "reference sample" described in 1-1. does not exist, and therefore all of the plurality of samples to be subjected to the miRNA detection are "samples to be corrected".

**[0071]** More specifically, the correction factor is calculated according to Equation 3 or Equation 3'.

50

$r_{1-2} =$  (fixed value (reference value))

- (representative value of correcting endogenous miRNA(s) in sample to be corrected) ... Equation 3

55

$r_{1-2}$ ' = (representative value of correcting endogenous miRNA(s) in sample to be corrected)

- (fixed value (reference value)) ... Equation 3'

[0072] For example, in cases where the average is used as the representative value of the correcting endogenous miRNA(s), the correction factor for the sample to be corrected can be calculated according to Equation 4 or Equation 4'.

$$r_{1-2} = \alpha - \frac{1}{n} \sum_{j=1}^n \log_2 Y_j \quad \dots \text{Equation 4}$$

$$r_{1-2}' = \frac{1}{n} \sum_{j=1}^n \log_2 Y_j - \alpha \quad \dots \text{Equation 4'}$$

[0073] Here, in Equation 4 and Equation 4',

$\alpha$  represents the reference value;

$n$  represents the total number of probe-immobilized area(s) for capturing the correcting endogenous miRNA(s) on the support; and

$X_j$  represents the signal measurement value from the probe-immobilized area for capturing the  $j$ th ( $1 \leq j \leq n$ ) correcting endogenous miRNA in a sample.

[0074] In cases where the probe(s) and the correcting endogenous miRNA(s) have a one-to-one relationship,  $n$  is equal to the number of the correcting endogenous miRNA(s) targeted by the correcting endogenous miRNA-capture probe(s) on the support.

[0075] In Equation 2 and Equation 2',  $n'$ , the total number of probe-immobilized areas for capturing the correcting endogenous miRNA(s) that were effectively judged positive across all the samples to be compared, may be used instead of  $n$ .

[0076] As the fixed value to be used as the reference value in the fixed-value correction method, any value (excluding 0) may be used as long as the same value is consistently used for all samples at least in one time of comparative analysis. By using the same expression measurement system, and always using the same value as the fixed value, comparative analysis can be carried out even between body fluid samples which were subjected to measurement of the expression levels on different days. The fixed value is not limited, and may be a value of an expression level which can be generally found for one or more correcting endogenous miRNAs to be used for obtaining the correction factor. Since, in general, such a value which can be generally found may vary depending on the system used for the measurement of the expression level, the fixed value may be selected without limitation depending on the system used.

[0077] For example, in cases where a single kind of correcting endogenous miRNA is used, the average of the expression levels of the correcting endogenous miRNA in a plurality of body fluid samples to be subjected to the comparative analysis may be determined, and the resulting average may be utilized as the fixed value. In cases where three kinds of correcting endogenous miRNAs are used, the average of the expression levels of all three kinds of miRNAs in a plurality of body fluid samples may be determined, and the resulting average may be utilized as the fixed value. A fixed value may be preliminarily determined using a large number of body fluid samples, and the determined fixed value may be repeatedly used in later analyses.

<Correction-Factor-Obtaining Step-2>

[0078] The correction-factor-obtaining step-2 is a method utilizing the ratio between the representative value of the correcting endogenous miRNA(s) and the reference value. To this step, the following 2-1. reference sample-obtaining method or 2-2. fixed-value correction method may be applied.

2-1. Reference Sample-Obtaining Method

**[0079]** One sample (first sample) is arbitrarily selected from a plurality of samples in which miRNAs are to be detected, which sample is used as a "reference sample". The remaining subsequent sample(s) is/are a "sample(s) to be corrected".

**[0080]** In this method, the representative value of the correcting endogenous miRNA(s) in the reference sample is used as a "reference value", and the ratio between the reference value and the representative value of the correcting endogenous miRNA(s) in a certain one of the subsequent sample(s) is used as the correction factor for the certain one of the subsequent sample(s). Thus, the number of the correction factors obtained is the same as the number of the samples to be corrected.

**[0081]** More specifically, the correction factor is calculated according to Equation 5 or Equation 5'.

$$c_{2-1} = \frac{\text{(representative value of correcting endogenous miRNA(s) in reference sample (reference value))}}{\text{(representative value of correcting endogenous miRNA(s) in sample to be corrected) ... Equation 5}}$$

$$c_{2-1}' = \frac{\text{(representative value of correcting endogenous miRNA(s) in sample to be corrected)}}{\text{(representative value of correcting endogenous miRNA(s) in reference sample (reference value)) ... Equation 5'}}$$

**[0082]** For example, in cases where measurement of the expression levels is carried out using a microarray, and the average is used as the representative value of the correcting endogenous miRNA(s), the correction factor for the second sample can be calculated according to Equation 6 or Equation 6'.

$$c_{2-1} = \frac{1}{n} \sum_{j=1}^n \log_2 A_j \div \frac{1}{n} \sum_{j=1}^n \log_2 X_j \quad \dots \text{Equation 6}$$

$$c_{2-1}' = \frac{1}{n} \sum_{j=1}^n \log_2 X_j \div \frac{1}{n} \sum_{j=1}^n \log_2 A_j \quad \dots \text{Equation 6'}$$

**[0083]** Here, in Equation 6 and Equation 6',

n represents the total number of probe-immobilized area(s) for capturing the correcting endogenous miRNA(s) on the support;

A<sub>j</sub> represents the signal measurement value from the probe-immobilized area for capturing the jth (1 ≤ j ≤ n) correcting endogenous miRNA in the reference sample; and

X<sub>j</sub> represents the signal measurement value from the probe-immobilized area for capturing the jth (1 ≤ j ≤ n) correcting endogenous miRNA in the second sample.

**[0084]** In cases where the probe(s) and the correcting endogenous miRNA(s) have a one-to-one relationship, n is equal to the number of the correcting endogenous miRNA(s) targeted by the correcting endogenous miRNA-capture probe(s) on the support.

**[0085]** In Equation 6 and Equation 6', n', the total number of probe-immobilized areas for capturing the correcting

endogenous miRNA(s) that were effectively judged positive across all the samples to be compared, may be used instead of n.

2-2. Fixed-Value Correction Method

5

**[0086]** This method preliminarily assumes that the representative value of the correcting endogenous miRNA(s) is constant among all samples. That is, the ratio between a fixed value and the representative value of the correcting endogenous miRNA(s) derived from each sample is obtained, and this ratio is utilized as the correction factor. In this method, the fixed value is used as the "reference value". Thus, the number of the correction factors obtained is the same as the number of the samples to be corrected. In this case, the "reference sample" described in 2-1. does not exist, and therefore all of the plurality of samples to be subjected to the miRNA detection are the "samples to be corrected".

10

**[0087]** More specifically, the correction factor is calculated according to Equation 7 or Equation 7'.

15

$$r_{2-2} = \frac{\text{(fixed value (reference value))}}{\text{(representative value of correcting endogenous miRNA(s) in sample to be corrected)}} \dots \text{Equation 7}$$

20

$$r_{2-2}' = \frac{\text{(representative value of correcting endogenous miRNA(s) in sample to be corrected)}}{\text{(fixed value (reference value))}} \dots \text{Equation 7'}$$

25

**[0088]** For example, in cases where the average is used as the representative value of the correcting endogenous miRNA(s), the correction factor for the sample to be corrected can be calculated according to Equation 8 or Equation 8'.

30

$$r_{2-2} = \alpha \div \frac{1}{n} \sum_{j=1}^n \log_2 Y_j \dots \text{Equation 8}$$

35

$$r_{2-2}' = \frac{1}{n} \sum_{j=1}^n \log_2 Y_j \div \alpha \dots \text{Equation 8'}$$

40

**[0089]** Here, in Equation 8 and Equation 8',

$\alpha$  represents a fixed value;

n represents the total number of probe-immobilized area(s) for capturing the correcting endogenous miRNA(s) on the support; and

45

$Y_j$  represents the signal measurement value from the probe-immobilized area for capturing the jth ( $1 \leq j \leq n$ ) correcting endogenous miRNA in a sample.

**[0090]** In cases where the probe(s) and the correcting endogenous miRNA(s) have a one-to-one relationship, n is equal to the number of the correcting endogenous miRNA(s) targeted by the correcting endogenous miRNA-capture probe(s) on the support.

50

**[0091]** In Equation 8 and Equation 8', n', the total number of probe-immobilized areas for capturing the correcting endogenous miRNA(s) that were effectively judged positive across all the samples to be compared, may be used instead of n.

55

**[0092]** Details of the "fixed value" to be used here as the reference value are the same as those of the fixed value in "1-2. Fixed-Value Correction Method".

**[0093]** Subsequently, using the correction factor obtained by the correction-factor-obtaining step-1 or the correction-factor-obtaining step-2, correction of the expression level(s) of the target miRNA(s) in a sample(s) to be corrected is

carried out utilizing the method of the correction step-1 or the correction step-2.

<Correction Step-1>

5 **[0094]** The correction step-1 is a process in which correction of the expression level(s) of the target miRNA(s) is carried out using the correction factor obtained in the correction-factor-obtaining step-1, and the correction is carried out by adding the correction factor to the expression level(s) of the target miRNA(s) or by subtracting the correction factor from the expression level(s). In this step, there are two ways of carrying out the correction ; one corresponds to the reference sample-obtaining method, and the other corresponds to the fixed-value correction method, in the correction-factor-obtaining step-1.

#### 1-1. Reference Sample-Obtaining Method

15 **[0095]** Correction of the expression level(s) of the target miRNA(s) in the subsequent sample(s) is carried out using a correction factor obtained for each one of the subsequent sample(s). That is, when the expression level(s) of the target miRNA(s) in the second sample is/are corrected, the correction factor for the second sample ( $c_{2-1}$  or  $c_{2-1}'$ ) is used, and, when the expression level(s) of the target miRNA(s) in the third sample is/are corrected, the correction factor for the third sample ( $c_{3-1}$  or  $c_{3-1}'$ ) is used.

20 **[0096]** In the case where the difference calculated by subtracting the representative value of the correcting endogenous miRNA(s) in each one of the subsequent sample(s) from the representative value of the correcting endogenous miRNA(s) in the reference sample is used as the correction factor, that is, in the case of the above-described Equation 1, each individual correction factor is added to the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the subsequent sample(s), thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the subsequent sample(s). In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th miRNA in a "sample to be corrected" can be calculated according to the following Equation 9.

$$E_i = \log_2 W_i + c_{1-1} \dots \text{Equation 9}$$

30 **[0097]** Here,  $W_i$  represents the signal measurement value from the probe-immobilized area for capturing the  $i$ th miRNA.

35 **[0098]** In contrast, in the case where the difference calculated by subtracting the representative value of the correcting endogenous miRNA(s) in the reference sample from the representative value of the correcting endogenous miRNA(s) in each one of the subsequent sample(s) is used as the correction factor, that is, in the case of the above-described Equation 1', each individual correction factor is subtracted from the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the subsequent sample(s), thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the subsequent sample(s). In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 9'.

$$E_i = \log_2 W_i - c_{1-1}' \dots \text{Equation 9'}$$

40 **[0099]** Here, the definition of  $W_i$  is the same as in the above-described Equation 9.

45 **[0100]** When the expression level(s) of the target miRNA(s) measured in the second sample is corrected,  $c_2$  may be added to, or  $c_2'$  may be subtracted from, each of the logarithmic value(s) of the expression level(s) of the target miRNA(s) in the second sample. The same applies to the third and following samples. It should be noted that, although the difference between the representative value of the first sample, which is used as the reference sample, and the reference value is, of course, 0, the program(s) may be constituted such that the calculation of adding 0 to, or subtracting 0 from, each of the expression level(s) of the target miRNA(s) in the first sample is carried out.

#### 1-2. Fixed-Value Correction Method

55 **[0101]** Correction of the expression level(s) of the target miRNA(s) is carried out using each individual correction factor obtained from the difference between the representative value and a fixed value (reference value). That is, when the expression level(s) of the target miRNA(s) in a certain sample is/are corrected, the correction factor for the certain sample ( $r_{1-2}$  or  $r_{1-2}'$ ) is used.

**[0102]** In the case where the difference calculated by subtracting the representative value of the correcting endogenous

miRNA(s) in each one of the samples from the reference value is used as the correction factor, that is, in the case of the above-described Equation 3, each individual correction factor is added to the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the samples, thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the samples. In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 10.

$$E_i = \log_2 Wi + r_{1-2} \quad \dots \text{Equation 10}$$

[0103] Here,  $W_i$  represents the signal measurement value from the probe-immobilized area for capturing the  $i$ th miRNA.

[0104] In contrast, in the case where the difference calculated by subtracting the fixed value from the representative value of the correcting endogenous miRNA(s) in each one of the samples is used as the correction factor, that is, in the case of the above Equation 3', each individual correction factor is subtracted from the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the samples, thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the samples. In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 10'.

$$E_i = \log_2 Wi - r_{1-2}' \quad \dots \text{Equation 10'}$$

[0105] The definition of  $W_i$  is the same as in the above-described Equation 10.

<Correction Step-2>

[0106] The correction step-2 is a process in which correction of the expression level(s) of the target miRNA(s) is carried out using the correction factor obtained in the correction-factor-obtaining step-2, and the correction is carried out by dividing the expression level(s) of the target miRNA(s) by the correction factor or multiplying the expression level(s) by the correction factor. Also in this step, there are two ways of carrying out the correction; one corresponds to the reference sample-obtaining method, and the other corresponds to the fixed-value correction method, in the correction-factor-obtaining step-2.

#### 2-1. Reference Sample-Obtaining Method

[0107] Correction of the expression level(s) of the target miRNA(s) in the subsequent sample(s) is carried out using the correction factor obtained for each one of the subsequent sample(s). That is, when the expression level(s) of the target miRNA(s) in the second sample is/are corrected, the correction factor for the second sample ( $c_{2-1}$  or  $c_{2-1}'$ ) is used, and, when the expression level(s) of the target miRNA(s) in the third sample is/are corrected, the correction factor for the third sample ( $c_{3-1}$  or  $c_{3-1}'$ ) is used.

[0108] In the case where the ratio calculated by using the representative value of the correcting endogenous miRNA(s) in each one of the subsequent sample(s) as a denominator and using the representative value of the correcting endogenous miRNA(s) in the reference sample as a numerator is used as the correction factor, that is, in the case of the above-described Equation 5, the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the subsequent sample(s) is/are multiplied by each individual correction factor, thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the subsequent sample(s). In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 11.

$$E_i = \log_2 Wi \times c_{2-1} \quad \dots \text{Equation 11}$$

[0109] Here,  $W_i$  represents the signal measurement value from the probe-immobilized area for capturing the  $i$ th miRNA.

[0110] In contrast, in the case where the ratio calculated by using the representative value of the correcting endogenous miRNA(s) in the reference sample as a denominator and using the representative value of the correcting endogenous miRNA(s) in each one of the subsequent sample(s) as a numerator is used as the correction factor, that is, in the case of the above-described Equation 5', the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the subsequent sample(s) is/are divided by each individual correction factor, thereby

carrying out correction of the expression level(s) of the target miRNA(s) in each of the subsequent sample(s). In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th miRNA in a "sample to be corrected" can be calculated according to the following Equation 11'.

5

$$E_i = \log_2 Wi \div c_{2-1}' \quad \dots \text{Equation 11'}$$

[0111] Here, the definition of  $W_i$  is the same as in the above-described Equation 11.

[0112] When the expression level(s) of the target miRNA(s) measured in the second sample is corrected, each of the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in the second sample may be divided by  $c_{2-1}$ , or may be multiplied by  $c_{2-1}'$ . The same applies to the third and following samples. The finally obtained value of the corrected expression level  $E_i$  of target miRNA is the same between the procedures based on Equation 5 and Equation 11 and the procedures based on Equation 5' and Equation 11'. It should be noted that, although the ratio between the representative value for the first sample used as the reference sample and the fixed value (reference value) in this method is, of course, 1, the program(s) may be constituted such that the calculation of multiplying or dividing each of the expression level(s) of the target miRNA(s) in the first sample by 1 is carried out.

## 2-2. Fixed-Value Correction Method

[0113] Correction of the expression level(s) of the target miRNA(s) is carried out using each individual correction factor obtained from the ratio to a fixed value (reference value). That is, when the expression level(s) of the target miRNA(s) in a certain sample is/are corrected, the correction factor for the certain sample ( $r_{2-2}$  or  $r_{2-2}'$ ) is used.

[0114] In the case where the ratio calculated by using the representative value of the correcting endogenous miRNA(s) in each of the samples as a denominator and using the reference value as a numerator is used as the correction factor, that is, in the case of the above-described Equation 7, the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the samples is multiplied by each individual correction factor, thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the samples. In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 12.

30

$$E_i = \log_2 Wi \times r_{2-2} \quad \dots \text{Equation 12}$$

[0115] Here,  $W_i$  represents the signal measurement value from the probe-immobilized area for capturing the  $i$ th miRNA.

[0116] In contrast, in the case where the ratio calculated by using the representative value of the correcting endogenous miRNA(s) in each one of the samples as a denominator and using the reference value as a numerator is used as the correction factor, that is, in the case of the above-described Equation 7', the logarithmic value(s) of the measured value(s) of the expression level(s) of the target miRNA(s) in each one of the samples is divided by the correction factor, thereby carrying out correction of the expression level(s) of the target miRNA(s) in each of the samples. In this case, the correction can be expressed as an equation as follows. That is, the corrected expression level  $E_i$  of the  $i$ th target miRNA in a "sample to be corrected" can be calculated according to the following Equation 12'.

40

$$E_i = \log_2 Wi \div r_{2-2}' \quad \dots \text{Equation 12'}$$

45

[0117] Here, the definition of  $W_i$  is the same as in the above-described Equation 12.

### <Comparative Analysis Step>

[0118] Based on the corrected expression level(s) of the target miRNA(s), the expression level(s) of the target miRNA(s) is/are compared among a plurality of body fluid samples. In cases where the correction is carried out by the reference sample-obtaining method, the expression level(s) of the target miRNA(s) in the first sample, which is used as the reference sample, is/are not subjected to the correction. Therefore, for example, comparison between the first sample and the second sample is carried out by comparing the uncorrected expression level of each target miRNA in the first sample with the corrected expression level of each target miRNA in the second sample. Thus, at least one of the samples to be subjected to the comparison is always a corrected sample. Accordingly, the term "based on the corrected expression level(s) of the target miRNA(s), the expression level(s) of the target miRNA(s) is/are compared among a plurality of body fluid samples" includes modes in which comparison is made between an uncorrected reference sample and a corrected

55

sample(s).

**[0119]** The comparative analysis step itself can be carried out in the same manner as in conventional methods. For example, the comparative analysis results can be represented as a scatter diagram of expression level data, which is called scatter plot. For example, when carrying out comparison among three samples, two scatter plots based on comparative analysis between any one of the three samples and each one of the remaining samples (for example, a scatter plot based on comparative analysis between the first sample and the second sample, and a scatter plot based on comparative analysis between the first sample and the third sample) may be prepared, and, if necessary, an additional scatter plot based on comparative analysis between the remaining two samples (in the above-exemplified case, between the second sample and the third sample) may be prepared. Comparative analysis among four or more samples may also be carried out in the same manner. In the case of comparative analysis among three samples, a three-dimensional scatter plot may also be prepared. Even in the cases of the reference sample-obtaining method, comparison between the reference sample and each one of the remaining two samples is not necessarily required, and, for example, comparison between the second sample and the reference sample and comparison between the second sample and the third sample may be carried out.

**[0120]** The result of the comparative analysis may also be represented by log fold-change, which may be obtained by calculating, based on the corrected expression level of each target miRNA, the difference in the expression level of each target miRNA between any one of the samples and the remaining sample(s). For example, the difference between the expression level of each target miRNA in the reference sample (in the case where the reference sample-obtaining method is employed) or the corrected expression level of each target miRNA in the first sample (in the case where the fixed-value correction method is employed) and the corrected expression level of each target miRNA in each of the subsequent sample(s) may be calculated. Also in this case, similarly to the above-described cases, the calculation of the difference may be carried out not only between the first sample and the remaining sample(s), but also between any one of the subsequent sample(s) and the remaining sample(s).

**[0121]** The miRNA expression analysis device of the present invention is a device that carries out the comparative analysis method of the present invention described above, and comprises:

memory means which memorizes measured values of the expression level(s) of a target miRNA(s) and measured values of the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, both of which measured values are those obtained by measurements on a plurality of body fluid samples;

representative-value-obtaining means which obtains, for each body fluid sample, a representative value, preferably a representative value expressed as a logarithmic value, from the measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs;

correction-factor-obtaining means which obtains, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for the above-mentioned each body fluid sample obtained by the representative-value-obtaining means;

correction means which corrects the expression level(s) of the target miRNA(s) measured in each body fluid sample using each individual correction factor obtained by the correction-factor-obtaining means; and

output means which outputs a result(s) of comparison of the expression level(s) of the target miRNA(s) among at least two body fluid samples based on the corrected expression level(s) of the target miRNA(s).

**[0122]** In one embodiment, the reference value is the representative value of the correcting endogenous miRNA(s) in the first sample (reference sample) that is arbitrarily selected, and the expression levels of the target miRNAs measured in the subsequent sample(s) are corrected. That is, in this embodiment, the miRNA expression analysis device of the present invention comprises:

memory means which memorizes measured values of the expression levels of a plurality of target miRNAs and a measured value(s) of the expression level(s) of one or more correcting endogenous miRNAs that were simultaneously measured in each one of a plurality of samples;

representative-value-obtaining means which obtains, for each body fluid sample, a representative value, preferably a representative value expressed as a logarithmic value, from the measured value(s) of the expression level(s) of the correcting endogenous miRNA(s);

correction-factor-obtaining means which obtains, using an arbitrarily-selected first sample as a reference sample and using the representative value of the correcting endogenous miRNA(s) in the reference sample as a reference value, the difference or the ratio between the reference value and the representative value of the correcting endogenous miRNA(s) in each one of the remaining subsequent sample(s) as a correction factor for the before-mentioned each one of the subsequent sample(s);

correction means which corrects the expression levels of the target miRNAs measured in each of the subsequent sample(s) using the correction factor for each of the subsequent sample(s) obtained by the correction-factor-obtaining means; and

output means which outputs a result(s) of comparison of the expression levels of the target miRNAs among at least two body fluid samples based on the corrected expression levels of the target miRNAs.

**[0123]** In another embodiment, the reference value is a fixed value arbitrarily defined in connection with the expression level(s) of the correcting endogenous miRNA(s), and correction of the expression levels of the target miRNAs is carried out for all samples including the first sample. That is, in this embodiment, the miRNA expression analysis device of the present invention comprises:

memory means which memorizes measured values of the expression levels of a plurality of target miRNAs and a measured value(s) of the expression level(s) of one or more correcting endogenous miRNAs that were simultaneously measured in each one of a plurality of body fluid samples;

representative-value-obtaining means which obtains, for each sample, a representative value, preferably a representative value expressed as a logarithmic value, from the measured value(s) of the expression level(s) of the correcting endogenous miRNA(s);

correction-factor-obtaining means which obtains, using a fixed value as a reference value, the difference or the ratio between the reference value and the representative value of the correcting endogenous miRNA(s) in each one of the samples as a correction factor for the before-mentioned each one of the samples;

correction means which corrects the expression levels of the target miRNAs measured in each sample using the correction factor for each sample obtained by the correction-factor-obtaining means; and

output means which outputs a result(s) of comparison of the expression levels of the target miRNAs among at least two body fluid samples based on the corrected expression levels of the target miRNAs.

**[0124]** Fig. 2 shows a block diagram illustrating an outline of the constitution of an example of the analysis device of the present invention. An analysis device of the present invention 10 comprises an input unit 110, display unit 120, output unit 130, memory unit 140, control unit 150, conversion unit 160, and analysis unit 170. Fig. 3 shows an example of the flow chart of the correction process for the expression levels of target miRNAs according to the present invention.

**[0125]** The input unit 110 is means for inputting information on the operation by the analysis device 10. Conventionally known input means such as keyboards may be preferably used. The expression level data obtained by a hybridization assay using a microarray are, for example, read by reading means such as a scanner which is separate from the device of the present invention, and then converted to numerical data. The resulting numerical data are input from the input unit 110 to the analysis device 10. Or, the reading means such as a scanner may be contained in the analysis device 10 of the present invention (not shown in the figure).

**[0126]** The expression level data input from the input unit 110, or the expression level data read and digitized by the reading means incorporated in the analysis device 10, are memorized in the memory unit 140. In this process, the memory unit 140 acts as memory means which memorizes measured values of the expression levels of a plurality of target miRNAs and a measured value(s) of the expression level(s) of one or more correcting endogenous miRNAs that were simultaneously measured in each one of a plurality of samples.

**[0127]** The measured value data on the expression levels of the target miRNAs and the correcting endogenous miRNA(s) in each sample stored in the memory unit 140 are converted to logarithms with base 2 by the conversion unit 160. Subsequently, by the analysis unit 170, a representative value of the expression level(s) of the correcting endogenous miRNA(s) converted to the logarithm(s) is obtained for each sample. As described in the explanation on the comparative analysis method, the representative value may be, for example, the average or the median of the expression level(s) of the one or more correcting endogenous miRNAs (even in cases where only one correcting endogenous miRNA is used for the correction, the representative value can be the average or the median when a plurality of probe-immobilized areas for measurement of the one correcting endogenous miRNA are arranged on the array), or may be a measured value of one specific correcting endogenous miRNA.

**[0128]** After the representative value is obtained, the difference or the ratio between the representative value for the correcting endogenous miRNA(s) in the reference sample and the representative value for the correcting endogenous miRNA(s) in each of the subsequent sample(s) is calculated by the analysis unit 170 to obtain the correction factor for each of the subsequent samples. Details of process of obtaining the correction factor are as described in <Correction-Factor-Obtaining Step> for the comparative analysis method. It should be noted that the program(s) may be constituted such that a correction factor of 0 (in the case where the difference is calculated) or a correction factor of 1 (in the case where the ratio is calculated) is obtained for the first sample, which is selected as the reference sample.

**[0129]** In the device 10, selection of the reference sample may be carried out by an operator of the device 10 by arbitrarily specifying one sample from the input unit 110. Or, the device 10 may automatically select one sample as the

reference sample. For example, after the input of data from the input unit 110, a sample whose data are first memorized in the memory unit 140 may be selected as the reference sample by the device 10. In Fig. 3, for convenience, the step of selection of or input of the reference sample is positioned after the representative-value-obtaining step (S-3). However, the position of this step is not limited thereto, and the step may also be carried out as an earlier step, for example, when the data are stored.

**[0130]** Next, the analysis unit 170 corrects the data on the expression levels of the target miRNAs measured in each of the subsequent sample(s) using each correction factor for each of the subsequent sample(s). Details of the correction operation are as described in <Correction Step> for the comparative analysis method. It should be noted that the program(s) may be constituted such that the correction operation is carried out for the data on the expression levels of the target miRNAs in the first sample, which is selected as the reference sample, using a correction factor of 0 (in the case where the difference is calculated) or a correction factor of 1 (in the case where the ratio is calculated).

**[0131]** Next, by the analysis unit 170, comparison between the expression levels of the target miRNAs in the reference sample and the corrected expression levels of the target miRNAs in the subsequent sample(s) is carried out. The results of the comparison are output by the output unit 130 to the display unit 120 to be displayed. In addition, the result of comparison may be output to an output device such as a printer; recording medium; and/or the like. The output unit 130 can also be constituted such that it outputs the result of the comparative analysis through a network to an external memory device such as a database present in the outside of the device 10.

**[0132]** The memory unit 140 memorizes the measured values of the expression levels of a plurality of target miRNAs and the expression levels of a plurality of correcting endogenous miRNAs, and also memorizes as appropriate the interim results generated in each of the above-described steps.

**[0133]** The above-described operations by the device 10 are controlled by the control unit 150. More specifically, as indicated by dashed arrows in Fig. 2, control information is output from the control unit 150 to the means, i.e., input unit 110, display unit 120, output unit 130, memory unit 140, control unit 150, conversion unit 160, and analysis unit 170, and these means work in concert in accordance with the control information, to allow the operation of the entire device 10.

**[0134]** In the analysis device, instead of using the representative value of the correcting endogenous miRNA(s) in the reference sample as a reference value, a preliminarily-specified fixed value may be registered in the conversion unit 160 or the like in the device 10, and may be used as a reference value that replaces the representative value of the expression level(s) of the correcting endogenous miRNA(s) in the reference sample. Details of such a method are as described in <Correction-Factor-Obtaining Step> and <Correction Step> for the comparative analysis method.

**[0135]** The present invention also provides a program(s) for causing a computer(s) to function as the above-described analysis device. More specifically, the program(s) is/are a program(s) for causing a computer(s) to function as the above-described means (that is, memory means, representative-value-obtaining means, correction-factor-obtaining means, correction means, and output means). The present invention also provides a program(s) for causing a computer(s) to execute the steps of the above-described comparative analysis method of the present invention. The comparative analysis method may comprise the above-described measurement step, representative-value-obtaining step, correction-factor-obtaining step, and correction step, and may also comprise the comparative analysis step of comparing the expression level(s) of a target miRNA(s) among a plurality of body fluid samples based on a corrected expression level(s) of the target miRNA(s). These programs are programs for causing a computer(s) to execute the correction of the expression level(s) of the target miRNA(s) using data on the expression level(s) of the correcting endogenous miRNA(s) which is/are simultaneously measured with the expression level(s) of the target miRNA(s) using a microarray or the like.

**[0136]** The present invention still further provides a computer-readable recording medium in which any of the program(s) described above is recorded.

**[0137]** The "recording medium" may be an arbitrary "portable physical medium" (non-transient recording medium) such as a flexible disk, magnetic optical disk, ROM, EPROM, EEPROM, CD-ROM, MO, or DVD. Or, the "recording medium" may be a "communication medium" which retains the program(s) for a short period, such as a communication line or a carrier wave used in transmitting the program(s) via a network represented by LAN, WAN, or internet.

**[0138]** The "program" is a data processing method written in an arbitrary language or by an arbitrary description method, and may be in any format including source code and binary code. The "program" is not limited to a single configuration, and includes a program having a distributed configuration as a plurality of modules and/or libraries, and a program which implements its function in cooperation with a separate program(s) represented by an OS (Operating System). In each of the devices shown in the embodiment, a well-known constitution and procedure can be used as a specific constitution for reading the recording medium, a reading procedure, an installation procedure after the reading, and the like.

**[0139]** The present invention provides a chip for analysis of miRNA expression, comprising a support on which probes for capturing a plurality of target miRNAs and probes for capturing a plurality of correcting endogenous miRNAs are immobilized. Preferred conditions for this chip are as described for the comparative analysis method of the present invention.

**[0140]** In the present invention, at least one of the miRNAs shown in SEQ ID NOs: 1 to 10 explained below is used

as the correcting endogenous miRNA(s).

**[0141]** SEQ ID NO:1 is the base sequence of hsa-miR-6085 deposited in miRBase under Accession No. MIMAT0023710. In the present description, the term "miR-6085 gene" or "miR-6085" includes hsa-miR-6085 described in SEQ ID NO:1, and its homologues and orthologues in other organism species. The hsa-miR-6085 gene can be  
5 obtained by the method described in Voellenkle C et al. (2012), RNA, vol. 18, pp. 472-484.

**[0142]** SEQ ID NO:2 is the base sequence of hsa-miR-1227-5p deposited in miRBase under Accession No. MIMAT0022941. In the present description, the term "miR-1227-5p gene" or "miR-1227-5p" includes hsa-miR-1227-5p described in SEQ ID NO:2, and its homologues and orthologues in other organism species. The hsa-miR-1227-5p gene  
10 can be obtained by the method described in Berezikov E et al. (2007), Molecular Cell, vol. 28, pp. 328-336.

**[0143]** SEQ ID NO:3 is the base sequence of hsa-miR-2861 deposited in miRBase under Accession No. MIMAT0013802. In the present description, the term "miR-2861 gene" or "miR-2861" includes hsa-miR-2861 described in SEQ ID NO:3, and its homologues and orthologues in other organism species. The hsa-miR-2861 gene can be  
15 obtained by the method described in Li H et al. (2009), Journal of Clinical Investigation, vol. 119, pp. 3666-3677.

**[0144]** SEQ ID NO:4 is the base sequence of hsa-miR-149-3p deposited in miRBase under Accession No. MIMAT0004609. In the present description, the term "miR-149-3p gene" or "miR-149-3p" includes hsa-miR-149-3p described in SEQ ID NO:4, and its homologues and orthologues in other organism species. The hsa-miR-149-3p gene  
20 can be obtained by the method described in Lagos-Quintana M et al. (2002), Current Biology, vol. 12, pp. 735-739.

**[0145]** SEQ ID NO:5 is the base sequence of hsa-miR-4463 deposited in miRBase under Accession No. MIMAT0018987. In the present description, the term "miR-4463 gene" or "miR-4463" includes hsa-miR-4463 described in SEQ ID NO:5, and its homologues and orthologues in other organism species. The hsa-miR-4463 gene can be  
25 obtained by the method described in Jima DD et al. (2010), Blood, vol. 116, pp. e118-e127.

**[0146]** SEQ ID NO:6 is the base sequence of hsa-miR-4508 deposited in miRBase under Accession No. MIMAT0019045. In the present description, the term "miR-4508 gene" or "miR-4508" includes hsa-miR-4508 described in SEQ ID NO:6, and its homologues and orthologues in other organism species. The hsa-miR-4508 gene can be  
30 obtained by the method described in Jima DD et al. (2010), Blood, vol. 116, pp. e118-e127.

**[0147]** SEQ ID NO:7 is the base sequence of hsa-miR-6090 deposited in miRBase under Accession No. MIMAT0023715. In the present description, the term "miR-6090 gene" or "miR-6090" includes hsa-miR-6090 described in SEQ ID NO:7, and its homologues and orthologues in other organism species. The hsa-miR-6090 gene can be  
35 obtained by the method described in Yoo JK et al. (2012), Stem Cells and Development, vol. 21, pp. 2049-2057.

**[0148]** SEQ ID NO:8 is the base sequence of hsa-miR-6775-5p deposited in miRBase under Accession No. MIMAT0027450. In the present description, the term "miR-6775-5p gene" or "miR-6775-5p" includes hsa-miR-6775-5p described in SEQ ID NO:8, and its homologues and orthologues in other organism species. The hsa-miR-6775-5p gene  
40 can be obtained by the method described in Ladewig E et al. (2012), Genome Research, vol. 22, pp.1634-1645.

**[0149]** SEQ ID NO:9 is the base sequence of hsa-miR-6803-5p deposited in miRBase under Accession No. MIMAT0027506. In the present description, the term "miR-6803-5p gene" or "miR-6803-5p" includes hsa-miR-6803-5p described in SEQ ID NO:9, and its homologues and orthologues in other organism species. The hsa-miR-6803-5p gene  
45 can be obtained by the method described in Ladewig E et al. (2012), Genome Research, vol. 22, pp.1634-1645.

**[0150]** SEQ ID NO: 10 is the base sequence of hsa-miR-5787 deposited in miRBase under Accession No. MIMAT0023252. In the present description, the term "miR-5787 gene" or "miR-5787" includes hsa-miR-5787 described in SEQ ID NO:10, and its homologues and orthologues in other organism species. The hsa-miR-5787 gene can be  
50 obtained by the method described in Yoo H et al. (2011), Biochem Biophys Res Commun, vol. 415, pp. 567-572.

## EXAMPLES

**[0151]** The present invention is described below more concretely by way of Examples. However, the present invention is not limited to the following Examples.

<Example 1>

**[0152]** By the following processes, the correcting endogenous miRNAs to be used in the present invention were selected.

(DNA Microarray)

**[0153]** Using a "3D-Gene" human miRNA oligo chip (based on miRBase release 20), manufactured by Toray Industries, Inc., the following experiment was carried out.

(Preparation of Sample RNA)

**[0154]** As sample RNAs, RNAs extracted from 157 specimens of healthy human serum using the "3D-Gene" RNA extraction reagent from liquid sample kit were used. The obtained sample RNAs were labeled using the "3D-Gene" miRNA labeling kit (Toray Industries, Inc.). The labeled sample RNAs were each subjected to hybridization and washing according to the standard protocol for the "3D-Gene" miRNA chip (Toray Industries, Inc.). The DNA microarray after the reaction was subjected to detection of fluorescence signals using a microarray scanner (Toray Industries, Inc.). Settings of the scanner were as follows: laser output, 100%; photomultiplier voltage, AUTO.

(Obtaining miRNA Signal Values)

**[0155]** The miRNA signal values obtained from the DNA microarray were converted to logarithms with base 2, and experimental error correction was carried out using an external standard nucleic acid added upon the RNA extraction and the labeling.

(Obtaining Correcting Endogenous miRNAs)

**[0156]** Using the miRNA signal values obtained from a total of 157 samples after the experimental error correction, miRNAs showing minimal inter-sample variation were extracted by the geNorm method. The GeNorm method is a search method for endogenous housekeeping genes (reference genes) in the quantitative RT-PCR method, proposed by Vandesompele et al. (Genome Biology 2002, 3: research 0034 - research 0034.11). This method is used for selecting mRNAs for endogenous housekeeping genes specific to the samples to be investigated. More specifically, the signal value ratio between two genes is calculated for all possible two-gene combinations generated from all genes to be detected in each individual sample. Subsequently, the standard errors among the samples are calculated for the values (A) of the ratios obtained for all combinations of the genes. Subsequently, the total (M) of the standard errors (V) obtained between a particular gene and other genes is calculated. Genes having small M values are considered to be excellent endogenous housekeeping genes.

**[0157]** On the other hand, it is said that the lower the signal values, the lower the values of A and V obtained by the above-described calculation method, so that genes with low signal values are found as apparently excellent endogenous housekeeping genes. Since the DNA microarray used in the present experiment comprehensively detects all miRNAs deposited in miRBase release 20, miRNAs having low signal values are preferentially selected when the geNorm method is carried out using these data. Because of such a phenomenon, use of the geNorm method for selection of endogenous housekeeping genes in DNA microarray detection has been avoided so far.

**[0158]** In the present Example, among the miRNAs detected with the DNA microarray, only miRNAs that were stably detected showing a signal value of over 64 in more than half of the samples were selected as subjects to which the geNorm method should be applied. Furthermore, the geNorm method is characterized in that endogenous housekeeping genes are selected based on the gene expression ratios calculated within one sample, and therefore a gene showing a stable expression level can be hardly selected among samples which show largely different expression patterns from each other. In view of this, we obtained CV (standard error / average) in advance for each of the miRNA signal values from a total of 157 samples whose experimental error had been corrected, and only miRNAs having a CV of not more than 0.1 were selected as subjects to which the geNorm method should be applied. As a result, among the signal values obtained from probes for detection of 2555 kinds of miRNAs, the signal values from probes for detection of 281 kinds of miRNAs were considered to be appropriate as the subjects to which the geNorm method should be applied.

**[0159]** From the 281 kinds of miRNAs to which the geNorm method was actually applied, we selected 10 kinds of endogenous housekeeping genes which had low M values, that is, which can be used as correcting endogenous miRNAs. The SEQ ID NOs. of these 10 kinds of correcting endogenous miRNAs, the names of the miRNAs, the MIMAT numbers deposited in miRBase, the M values calculated by geNorm, the base sequences, and the ratio of guanine/cytosine (GC%) in each sequence, are shown in Table 1.

[Table 1]

SEQ ID NO.	miRNA name	MIMAT ID	geNorm M value	Base sequence	GC%
1	hsa-miR-6085	MIMAT0023710	0.058	AAGGGGCUGGGGGAGCACA	68
2	hsa-miR-1227-5p	MIMAT0022941	0.058	GUGGGGCCAGGCGGUGG	82
3	hsa-miR-2861	MIMAT0013802	0.059	GGGGCCUGGCGGUGGGCGG	89

(continued)

SEQ ID NO.	miRNA name	MIMAT ID	geNorm M value	Base sequence	GC%
4	hsa-miR-149-3p	MIMAT0004609	0.059	AGGGAGGGACGGGGCUGUGC	76
5	hsa-miR-4463	MIMAT0018987	0.060	GAGACUGGGUGGGGCC	76
6	hsa-miR-4508	MIMAT0019045	0.060	GCGGGGCUGGGCGCGCG	94
7	hsa-miR-6090	MIMAT0023715	0.061	GGGGAGCGAGGGGCGGGGC	89
8	hsa-miR-6775-5p	MIMAT0027450	0.061	UCGGGGCAUGGGGGAGGGAGGCUGG	76
9	hsa-miR-6803-5p	MIMAT0027506	0.062	CUGGGGGUGGGGGGCUGGGCGU	82
10	hsa-miR-5787	MIMAT0023252	0.063	GGGCUGGGGCGCGGGAGGU	85

**[0160]** Detection of miRNAs using a DNA microarray is carried out by the hybridization method, which is dependent on formation of nucleic acid double strands between target miRNAs and probes for capturing the target miRNAs. Therefore, it is expected that sequences having high contents of guanine and cytosine, which stabilize nucleic acid double strands, may be especially stably detected by the DNA microarray. That is, it can be generally assumed that sequences having high contents of guanine and cytosine are preferentially selected as endogenous housekeeping genes. However, as shown in Table 1, the endogenous housekeeping genes (correcting endogenous miRNAs) we selected by the geNorm method did not necessarily have high contents of guanine and cytosine. Furthermore, the phenomenon that miRNAs having high contents of guanine and cytosine preferentially become endogenous housekeeping genes (correcting endogenous miRNAs) was not also observed.

<Example 2>

**[0161]** Correction of the expression levels of target RNAs in a plurality of body fluid samples was carried out using the selected correcting endogenous miRNAs individually.

**[0162]** Due to the physical instability of RNA, measurement of the RNA expression level often requires comparison of the absolute expression level among samples with different qualities, that is, with different degrees of RNA degradation. In particular, when body fluids are used as samples, there may be a situation where an attempt is made to obtain a certain result by measuring the expression levels of target miRNAs in samples that have been stored at room temperature for a long time.

**[0163]** In such a situation, by carrying out correction depending on the degree of degradation of RNA using as a reference the amount(s) of a particular correcting endogenous miRNA(s) contained in serum, absolute comparison of the miRNA expression profile (relative miRNA expression level ratio) among the samples may become possible irrespective of the RNA quality.

**[0164]** In the present Example, serum collected from human was stored under various conditions to prepare serum samples containing miRNAs at various levels of quality, and the correction method of the present invention was applied to the measured values of extracted RNA samples derived from these serum samples.

(DNA Microarray)

**[0165]** Using a "3D-Gene" human miRNA oligo chip (based on miRBase release 20; Toray Industries, Inc.), manufactured by Toray Industries, Inc., the following experiment was carried out.

(Preparation of Sample RNA)

**[0166]** Human blood was collected, and subjected to serum separation. The separated serum was then left to stand under the following two different conditions to allow changes in the quality of the RNA contained in the serum. The two different conditions were as follows. 1: After the serum separation, the serum was left to stand at room temperature for 6 hours, and RNA extraction was then carried out. 2: After the serum separation, the serum was left to stand at 25°C for 24 hours, and RNA extraction was then carried out. From the serum prepared under each condition, total RNA was extracted using the "3D-Gene" RNA extraction reagent from liquid sample kit RNA, to provide sample RNAs.

**[0167]** Each obtained sample RNA was labeled using the "3D-Gene" miRNA labeling kit (Toray Industries, Inc.). The

labeled sample RNAs were each subjected to hybridization and washing according to the standard protocol for the "3D-Gene" human miRNA oligo chip (Toray Industries, Inc.). The DNA microarray after the reaction was subjected to detection of fluorescence signals using a microarray scanner (Toray Industries, Inc.). Settings of the scanner were as follows: laser output, 100%; photomultiplier voltage, AUTO.

5

(Correction of miRNA Signal Values)

[0168] Fig. 4A shows the distributions of the signal values obtained by converting the signal values of miRNAs obtained from the DNA microarray to logarithms with base 2 (that is, uncorrected signal values of the expression levels). On the other hand, Figs. 4B to K show the distributions of the corrected signal values obtained by correction using the signal values of the correcting endogenous miRNAs shown in SEQ ID Nos:1 to 10 and reference values which were fixed values separately set in advance. In the figures, Day-0 (dashed line) indicates the sample RNA through Condition 1, and Day-1\_25°C (solid line) indicates the sample RNA through Condition 2.

10

15

[0169] Here, the correction using the signal values of the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 was carried out by the method described below.

15

[0170] For each of the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, a fixed value to be used as a reference value was set. Further, the signal values of the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 in each sample were converted to logarithms with base 2, and the converted signal values were used as representative values. The ratio between each representative value and each reference value (reference value / representative value) was calculated to provide a correction factor. Thus, the correction factor was set for each of the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10. All miRNA signal values in each sample were multiplied by the correction factor calculated from each correcting endogenous miRNA, to carry out correction.

20

[0171] As a result, as shown in Figs. 4B to K, the expression profiles of the two samples prepared under the different conditions could be agreed with each other by using any one of the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10.

25

<Example 3>

[0172] As shown in Example 2, the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 have a sufficient effect to correct the expression levels of the target miRNAs even when they are used individually. Moreover, even if partial loss of data from endogenous miRNAs for correction occurred due to failure in the experiment or the like, complementation of the lost data may be possible by using them in combination. In view of this, a plurality of correcting endogenous miRNAs were used in combination to perform correction of the expression levels of target miRNAs, and the effect of the correction was checked. The samples used herein were the same as those in Example 2.

30

[0173] The signal values of the correcting endogenous miRNAs shown in SEQ ID NOs:2, 4, and 5 were converted to logarithms with base 2, and the average or the median of the converted signal values was obtained to provide a representative value. As a reference value, a fixed value was used. The average of the expression levels (logarithmically converted signal values) of all of the three correcting endogenous miRNAs in all of a plurality of serum RNA samples including the two kinds of samples prepared in Example 2 was calculated, and the calculated average was used as the fixed value (reference value). In each sample, the ratio between the representative value of the correcting endogenous miRNAs and the reference value (reference value / representative value) was calculated to provide a correction factor for the sample. All target miRNA signal values in each corresponding sample were multiplied by each corresponding correction factor to perform correction of the target miRNA signal values.

35

40

[0174] Fig. 5 shows the distributions of the corrected signal values of the target miRNAs obtained by using the average of the logarithmically converted signal values as a representative value of the expression levels of the correcting endogenous miRNAs. Similarly to Example 2 in which the correction was carried out by using the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 individually, the target miRNA expression profiles of the samples could be agreed with each other by performing the correction using a combination of a plurality of correcting endogenous miRNAs.

45

50

<Example 4>

[0175] As shown in Example 3, the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 are also useful for the correction when they are used in combination. Whether any combination of SEQ ID NOs:1 to 10 can exhibit a similar correction effect or not was then investigated. The samples used herein were the same as those in Example 2.

55

[0176] The signal values of the correcting endogenous miRNAs shown in SEQ ID NOs:2, 3, 4, 5, and 6 were converted to logarithms with base 2, and the average of the converted signal values of arbitrarily selected three kinds of correcting endogenous miRNAs was obtained for each sample to provide a representative value in the sample. A fixed value was used as a reference value. The average of the logarithmically converted signal values of the three kinds of correcting

endogenous miRNAs, which were used to obtain the representative value, in a plurality of serum RNA samples including the two kinds of samples prepared in Example 2 was used as the fixed value (reference value). In each sample, the ratio between the representative value and the reference value (reference value / representative value) was calculated to provide a correction factor for the sample. All target miRNA signal values in each corresponding sample were multiplied by each corresponding correction factor, thereby performing correction of the expression levels of the miRNAs in the samples.

[0177] The distributions of the corrected signal values are shown in Fig. 8. Fig. 8A shows the distribution of the uncorrected expression level signal values, and Figs. 8B to I show the distributions of the corrected expression level signal values. Similarly to Example 2 in which the correction was carried out by using the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 individually, the target miRNA expression profiles of the samples could be agreed with each other.

<Comparative Example 1>

[0178] The study in Example 1 was carried out in the same manner except that, in the step of labeling each sample RNA obtained from serum RNA using the "3D-Gene" miRNA labeling kit (Toray Industries, Inc.), two kinds of synthetic RNA sequences (each of which is a 20-mer sequence; referred to as spike control 1 and spike control 2) were added as reference substances that were externally added. The sequences of these synthetic RNAs were fluorescently labeled similarly to the serum-derived miRNAs, and their signal levels were detected by the "3D-Gene" human miRNA oligo chip (Toray Industries, Inc.). The signal values obtained from a total of 16 spots, that is, 8 spots for each of the spike control 1 and the spike control 2, were converted to logarithms with base 2, and the median of the converted signal values was obtained to provide a representative value. The ratio between this representative value and a fixed value (reference value) which was set in advance (reference value / representative value of each sample) was calculated to provide a correction factor for each sample. All miRNA signal values in each sample were multiplied by this correction factor to carry out correction. The distributions of the uncorrected signal values are shown in Fig. 6A, and the distributions of the corrected signal values are shown in Fig. 6B. It was shown that the correction by the spike controls externally added upon the labeling failed to sufficiently correct the signal values of the target miRNAs in the samples.

<Comparative Example 2>

[0179] Table 2 shows the expression levels of let-7d-5p, let-7g-5p, let-7i-5p, miR-16, miR-31, and miR-223 measured in the same manner as in Example 2, whose use in correction of the expression levels of miRNAs has been shown in prior art examples. Those except has-miR-16-5p only showed signals within the range of 0 to 5 in terms of the values calculated by conversion of the obtained signal values to logarithms with base 2. Thus, they were found to be not suitable as references for correcting the expression levels of all miRNAs because of their insufficient expression levels. Fig. 7 shows the distributions of the corrected signal values obtained by correction with hsa-miR-16-5p. It was shown that the correction by the expression level of hsa-miR-16-5p failed to sufficiently correct the signal values of the target miRNAs in the samples.

[Table 2]

miRNA name	ID	Expression level Day0-1	Expression level Day-1_25°C
hsa-let-7e-5p	MIMAT0000066	0.0	0.0
hsa-let-7e-3p	MIMAT0004485	4.7	4.1
hsa-let-7g-5p	MIMAT0000414	4.0	0.0
hsa-let-7g-3p	MIMAT0004584	4.0	0.0
hsa-let-7i-5p	MIMAT0000415	4.4	0.0
hsa-let-7i-3p	MIMAT0004585	3.9	0.0
hsa-miR-16-1-3p	MIMAT0004489	3.6	0.0
hsa-miR-16-2-3p	MIMAT0004518	3.8	4.5
hsa-miR-16-5p	MIMAT0000069	7.1	7.5
hsa-miR-31-5p	MIMAT0000089	0.0	0.0
hsa-miR-31-3p	MIMAT0004504	3.7	4.8

(continued)

miRNA name	ID	Expression level Day0-1	Expression level Day-1_25°C
hsa-miR-223-5p	MIMAT0004570	3.7	4.4
hsa-miR-223-3p	MIMAT0000280	6.0	0.0

## DESCRIPTION OF SYMBOLS

**[0180]**

- 10 Device
- 110 Input unit
- 120 Display unit
- 130 Output unit
- 140 Memory unit
- 150 Control unit
- 160 Conversion unit
- 170 Analysis unit

**Claims**

1. A method for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, said method comprising:

a measurement step of simultaneously measuring the expression level(s) of a target miRNA(s) and the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 in the respective body fluid samples;

a representative-value-obtaining step of obtaining a representative value for each of said plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

a correction-factor-obtaining step of obtaining, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained in the representative-value-obtaining step; and

a correction step of correcting the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample.

2. The method according to claim 1, wherein the correction in said correction step is carried out as follows:

(a) in the case where a value calculated by subtracting said reference value from said representative value is obtained as a correction factor in said correction-factor-obtaining step, the correction factor is subtracted from the measured value(s) of the expression level(s) of the target miRNA(s);

(b) in the case where a value calculated by subtracting said representative value from said reference value is obtained as a correction factor in said correction-factor-obtaining step, the correction factor is added to the measured value(s) of the expression level(s) of the target miRNA(s);

(c) in the case where a value calculated by dividing said representative value by said reference value is obtained as a correction factor in said correction-factor-obtaining step, the measured value(s) of the expression level(s) of the target miRNA(s) is/are divided by the correction factor; or

(d) in the case where a value calculated by dividing said reference value by said representative value is obtained as a correction factor in said correction-factor-obtaining step, the measured value(s) of the expression level(s) of the target miRNA(s) is/are multiplied by the correction factor.

3. The method according to claim 1 or 2, wherein said reference value is a fixed value arbitrarily defined in connection with the expression level(s) of the correcting endogenous miRNA(s), or a representative value of the expression level(s) of the correcting endogenous miRNA(s) obtained for a first body fluid sample, said first body fluid sample

being arbitrarily selected from said plurality of body fluid samples.

- 5
4. The method according to any one of claims 1 to 3, wherein said measurement step comprises carrying out hybridization by bringing nucleic acid probes for capturing a plurality of target miRNAs and a probe(s) for capturing the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, said probes being immobilized on a support, into contact with a nucleic acid sample derived from each of the body fluid samples, said nucleic acid sample being labeled with a labeling substance, and obtaining the expression levels of the target miRNAs and said one or more correcting endogenous miRNAs as signal intensity measurement values.
- 10
5. The method according to any one of claims 1 to 4, wherein said body fluid sample is blood, serum, or plasma.
- 15
6. The method according to any one of claims 1 to 5, wherein said representative value is an average or a median expressed as a logarithmic value calculated from the measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10.
- 20
7. A miRNA expression analysis device for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, said device comprising:

memory means which memorizes measured values of the expression level(s) of a target miRNA(s) and measured values of the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, both of the measured values being obtained by measurements on a plurality of body fluid samples;

25 representative-value-obtaining means which obtains a representative value for each of said plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

30 correction-factor-obtaining means which obtains, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained by said representative-value-obtaining means;

correction means which corrects the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample; and

35 output means which outputs a result(s) of comparison of the expression level(s) of the target miRNA(s) among at least two body fluid samples based on the corrected expression level(s) of the target miRNA(s).

8. The device according to claim 7, wherein said correction means carries out said correction as follows:

40 (a) in the case where a value calculated by subtracting said reference value from said representative value is obtained as a correction factor in said correction-factor-obtaining means, the correction factor is subtracted from the measured value(s) of the expression level(s) of the target miRNA(s);

(b) in the case where a value calculated by subtracting said representative value from said reference value is obtained as a correction factor in said correction-factor-obtaining means, the correction factor is added to the measured value(s) of the expression level(s) of the target miRNA(s);

45 (c) in the case where a value calculated by dividing said representative value by said reference value is obtained as a correction factor in said correction-factor-obtaining means, the measured value(s) of the expression level(s) of the target miRNA(s) is/are divided by the correction factor; or

50 (d) in the case where a value calculated by dividing said reference value by said representative value is obtained as a correction factor in said correction-factor-obtaining means, the measured value(s) of the expression level(s) of the target miRNA(s) is/are multiplied by the correction factor.

- 55
9. The device according to claim 7 or 8, wherein the measured values of the expression level(s) of the target miRNA(s) contained in the plurality of body fluid samples and the measured values of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 that are memorized in said memory means are values obtained by carrying out hybridization by bringing probes for capturing a plurality of target miRNAs and a probe(s) for capturing said one or more correcting endogenous miRNAs, said probes being immobilized on a support, into contact with a nucleic acid sample derived from each of the body fluid samples, said nucleic acid sample being labeled with a labeling substance, and obtaining the

expression levels of the target miRNAs and said one or more correcting endogenous miRNAs as signal intensity measurement values.

- 5 **10.** A program(s) for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, said program(s) causing one or more computers to execute:

a measurement step of simultaneously measuring the expression level(s) of a target miRNA(s) and the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 in the respective body fluid samples;

10 a representative-value-obtaining step of obtaining a representative value for each of said plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

a correction-factor-obtaining step of obtaining, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained in the representative-value-obtaining step; and

15 a correction step of correcting the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample.

- 20 **11.** A program(s) for comparative analysis of the expression level(s) of a target miRNA(s) among a plurality of body fluid samples, said program(s) causing one or more computers to function as:

memory means which memorizes measured values of the expression level(s) of a target miRNA(s) and measured values of the expression level(s) of one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10, said target miRNA(s) and said one or more correcting endogenous miRNAs being contained in each of said plurality of body fluid samples;

representative-value-obtaining means which obtains a representative value for each of said plurality of body fluid samples from a measured value(s) of the expression level(s) of the one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10;

30 correction-factor-obtaining means which obtains, as a correction factor for correction of the expression level(s) of the target miRNA(s) in each body fluid sample, the difference or the ratio between a reference value that is arbitrarily set in connection with the expression level(s) of the correcting endogenous miRNA(s) and the representative value for said each body fluid sample obtained by said representative-value-obtaining means;

correction means which corrects the expression level(s) of the target miRNA(s) measured in each body fluid sample using the correction factor obtained for said each body fluid sample; and

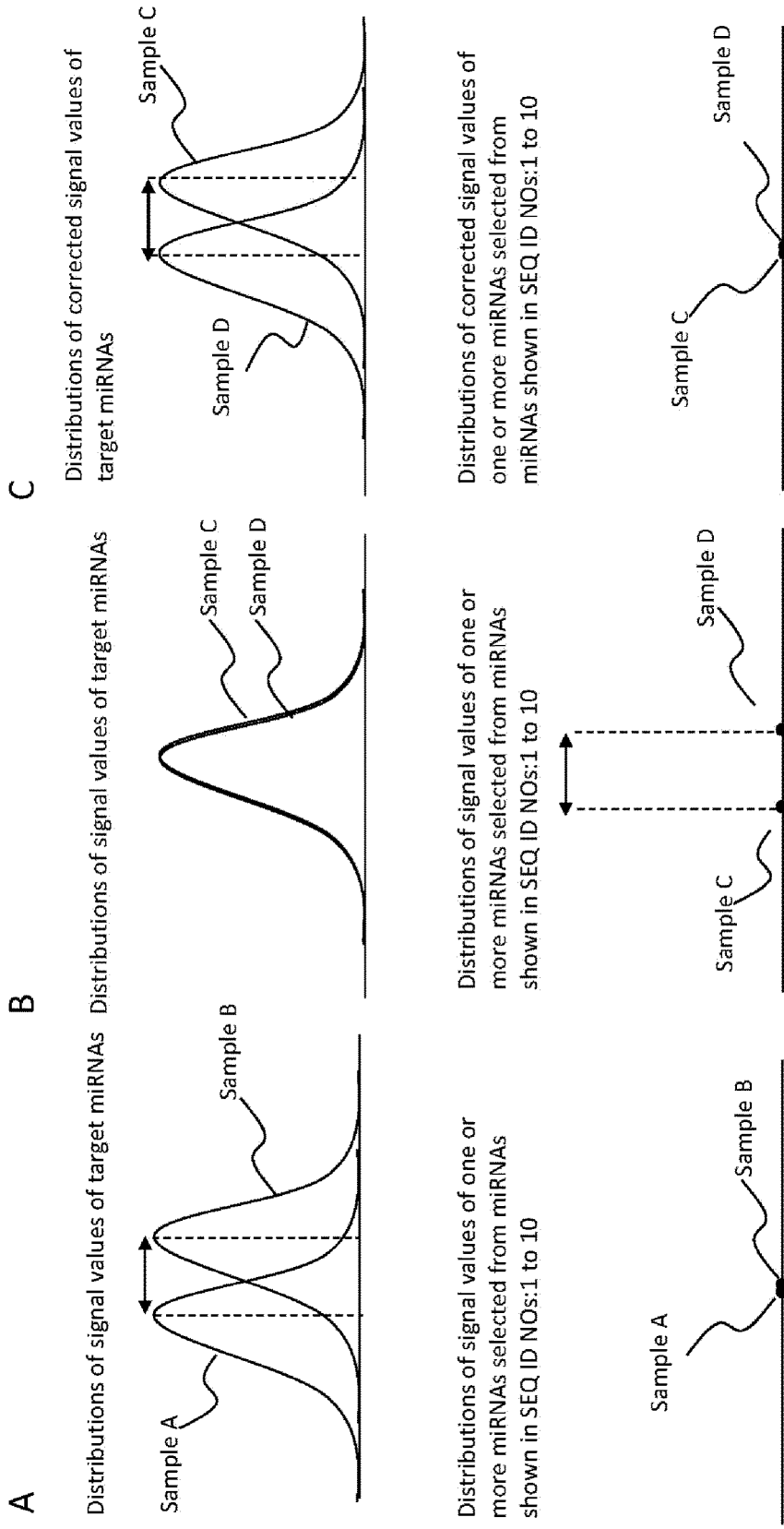
35 output means which outputs a result(s) of comparison of the expression level(s) of the target miRNA(s) among at least two body fluid samples based on the corrected expression level(s) of the target miRNA(s).

- 40 **12.** A computer-readable recording medium in which the program(s) according to claim 10 or 11 is/are recorded.

- 45 **13.** A chip for analysis of miRNA expression, comprising a support on which probes for capturing a plurality of target miRNAs and a probe(s) for capturing one or more correcting endogenous miRNAs selected from the correcting endogenous miRNAs shown in SEQ ID NOs:1 to 10 are immobilized.

50

55



**Fig. 1**

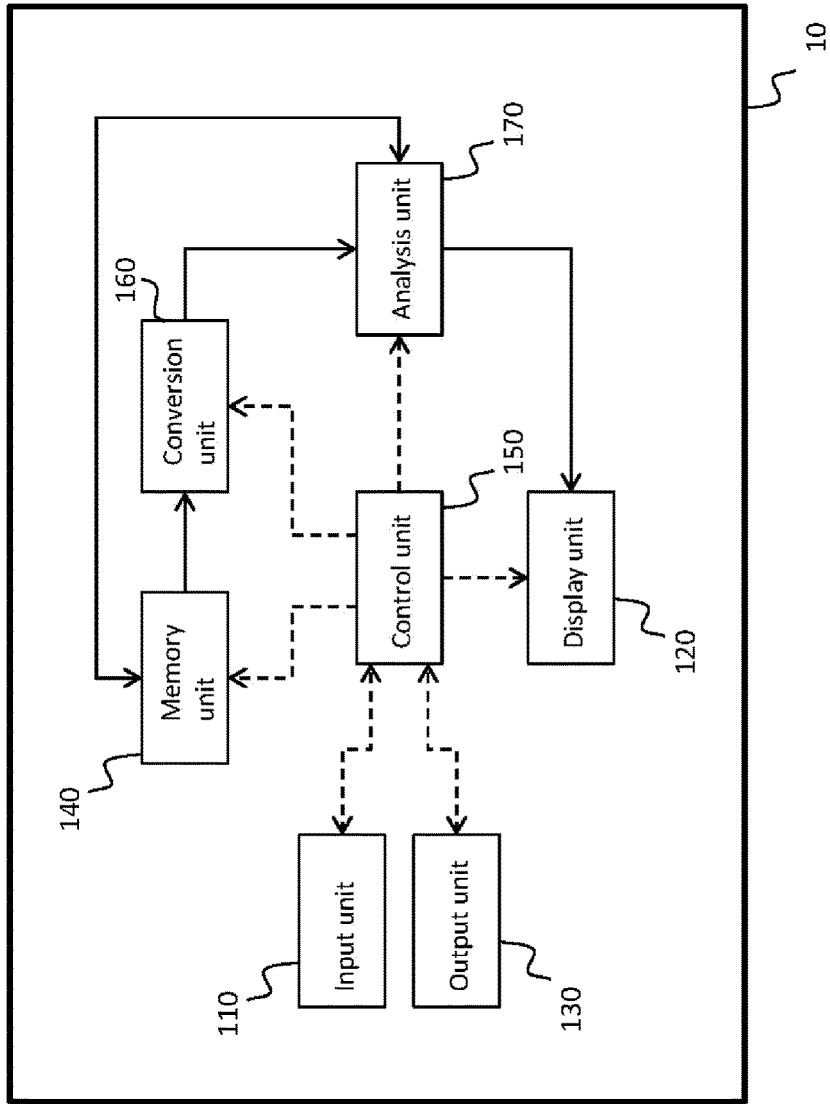
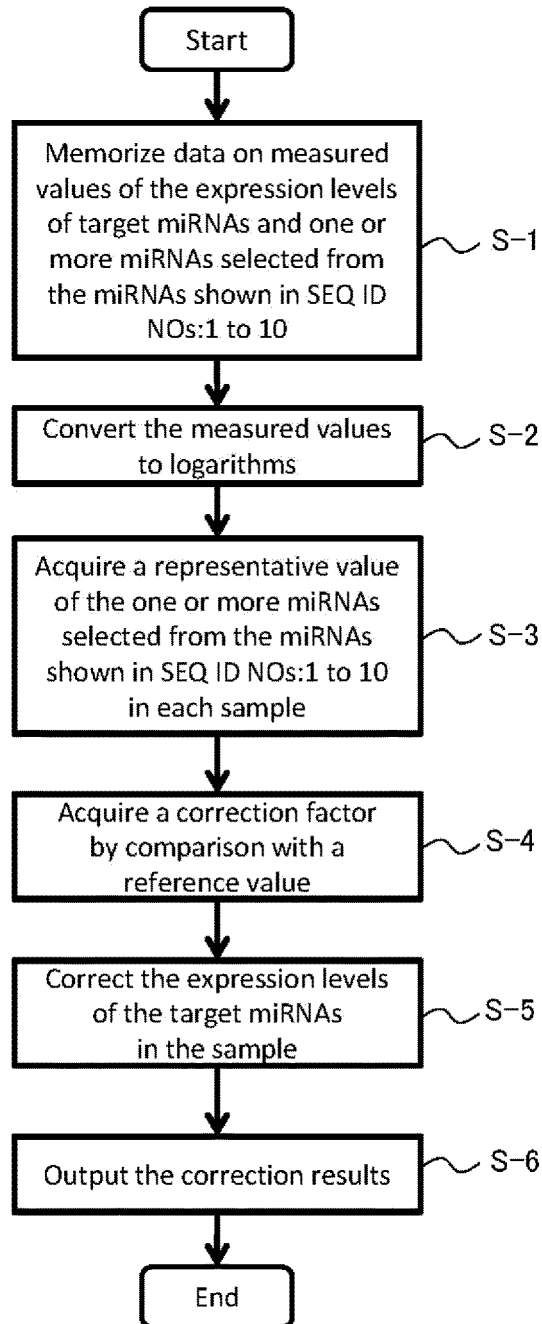


Fig. 2

**Fig. 3**

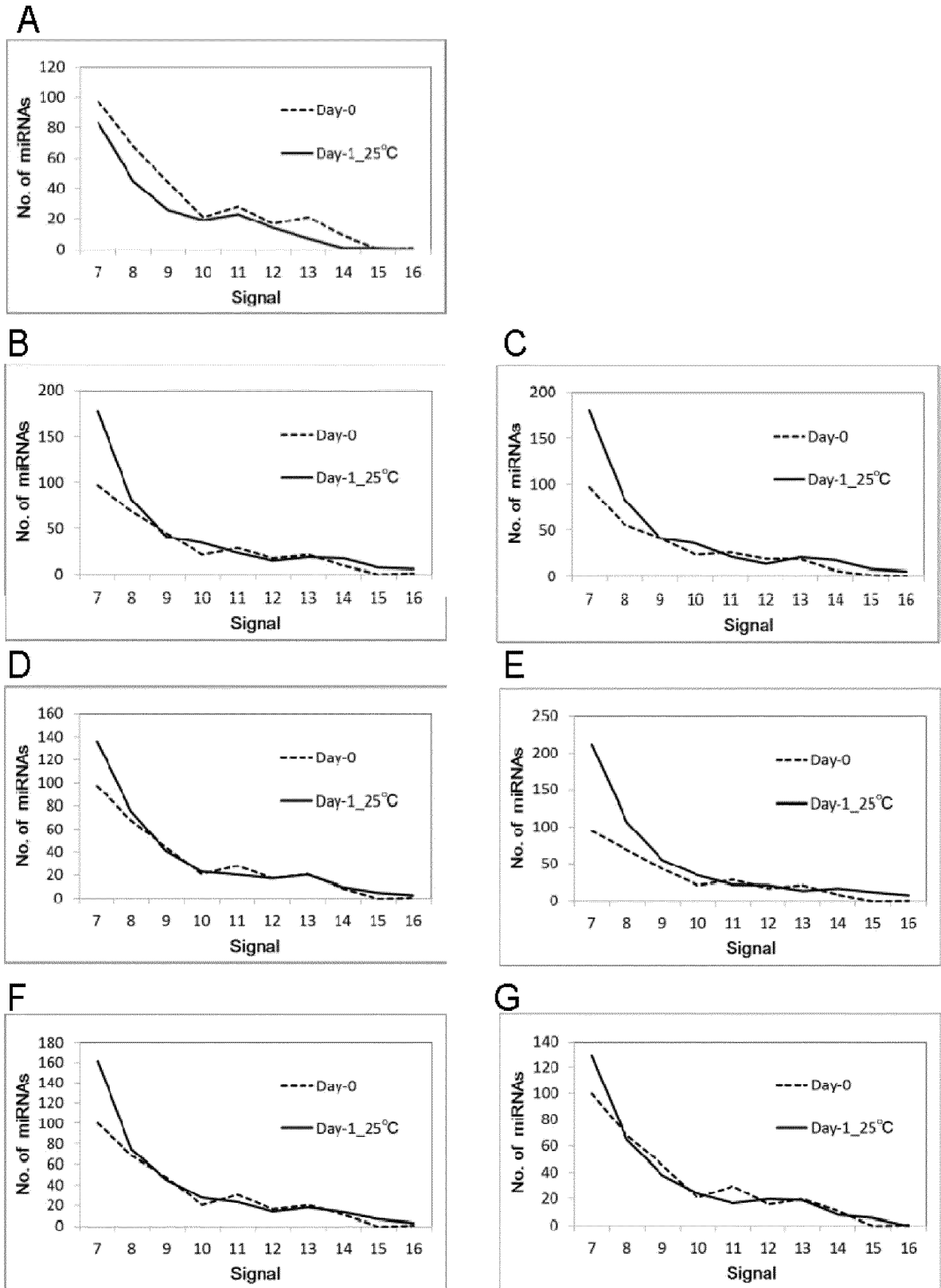
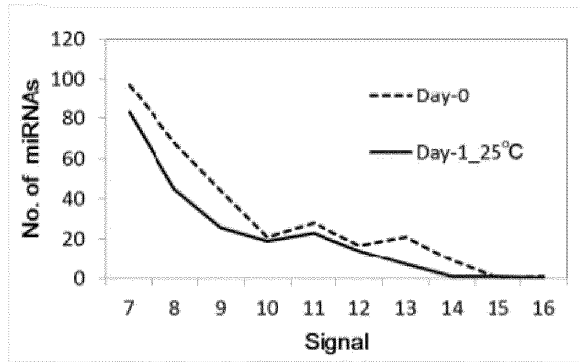
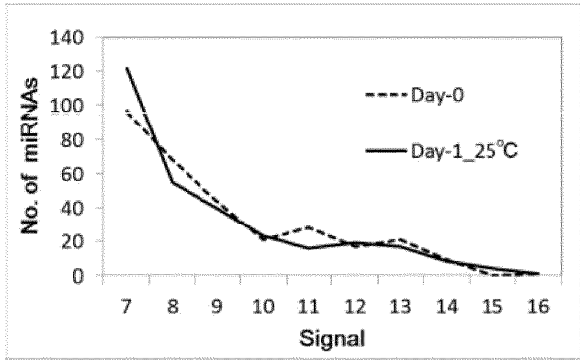


Fig. 4-1

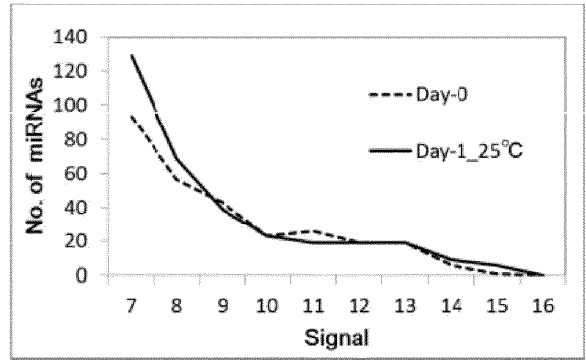
A



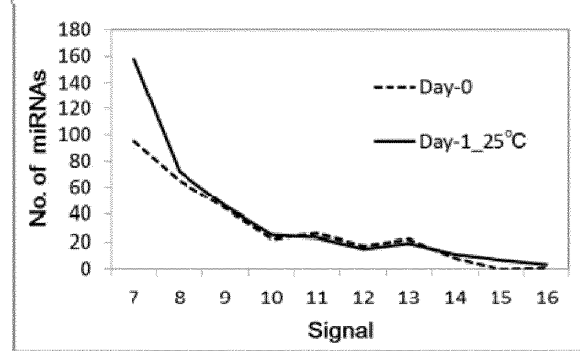
H



I



J



K

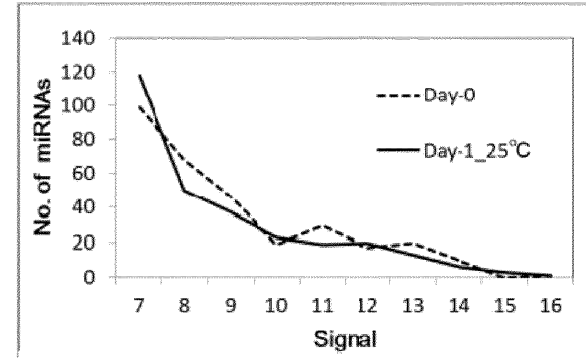


Fig. 4-2

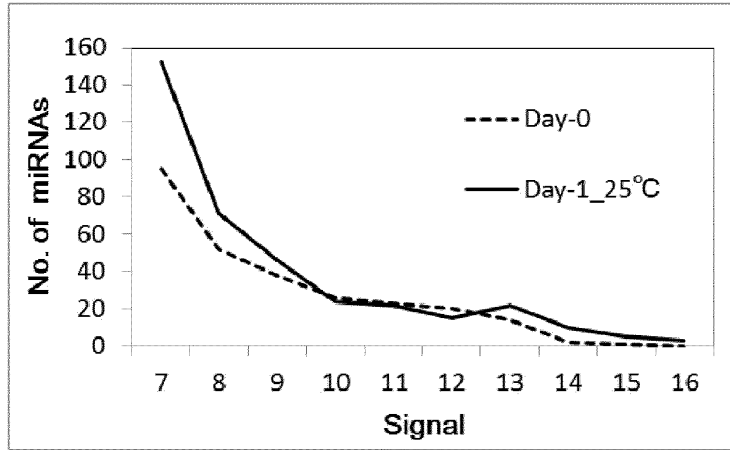


Fig. 5

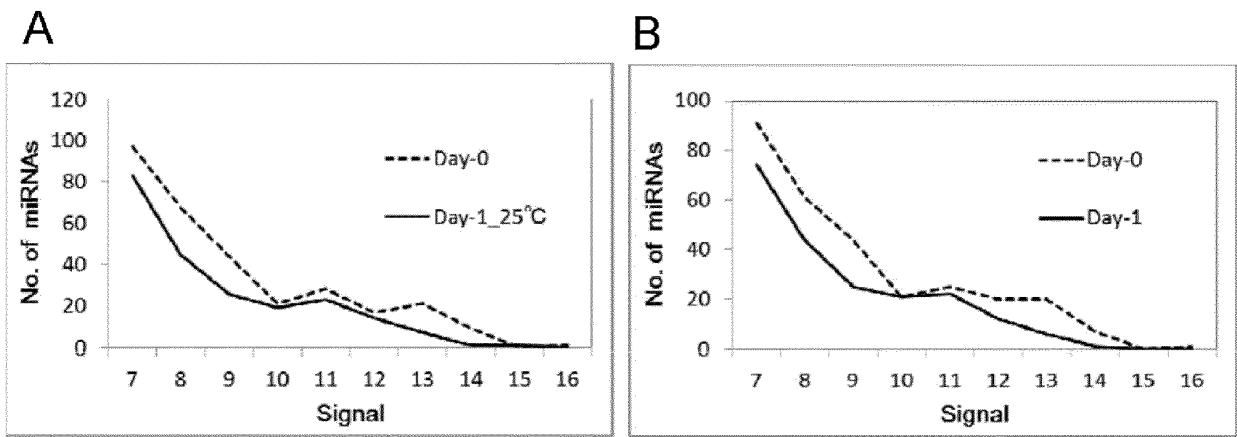


Fig. 6

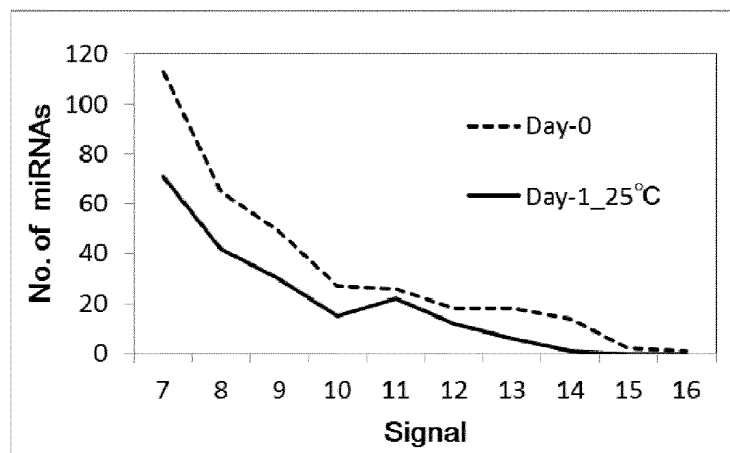
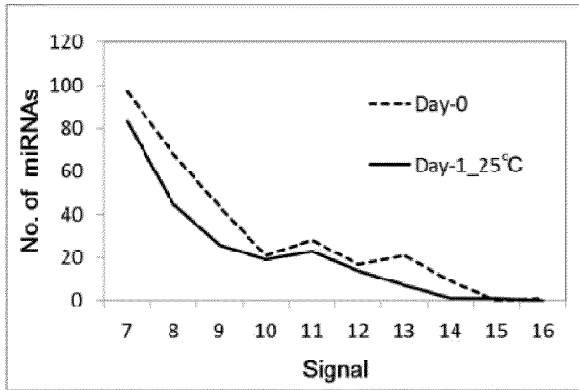
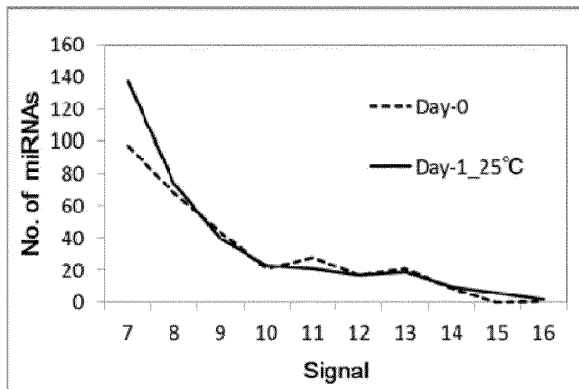


Fig. 7

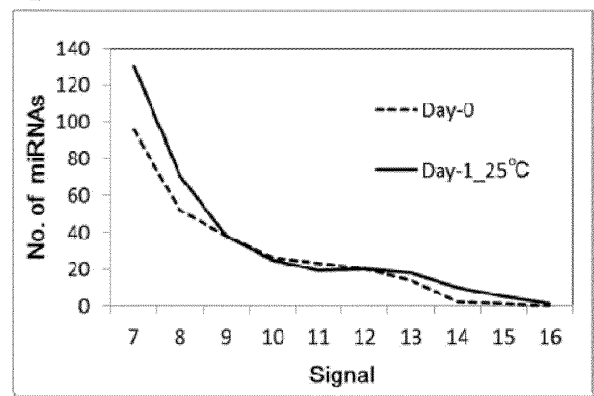
A



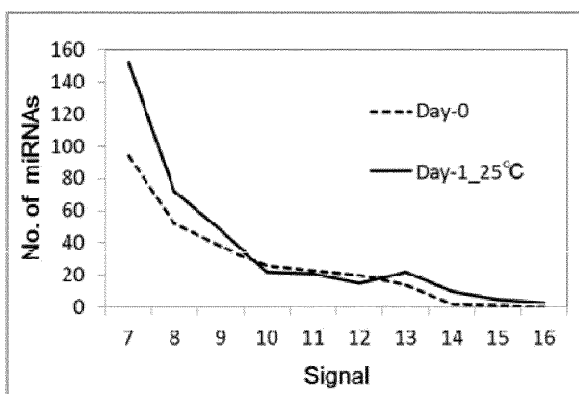
B



C



D



E

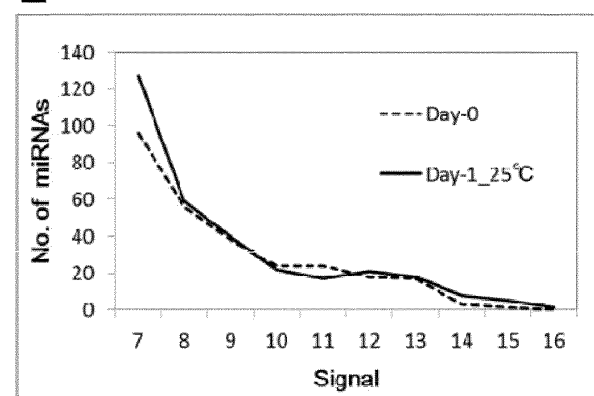
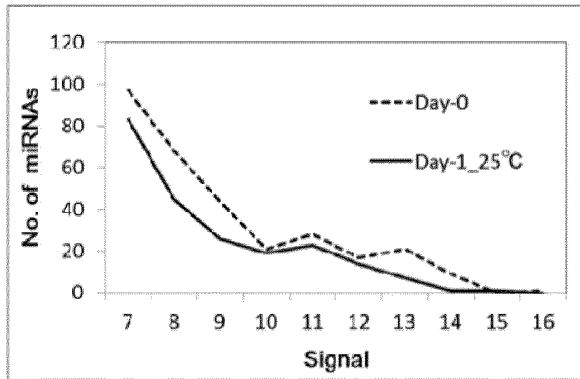
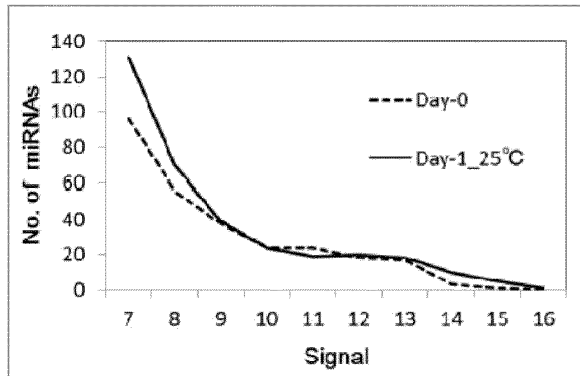


Fig. 8-1

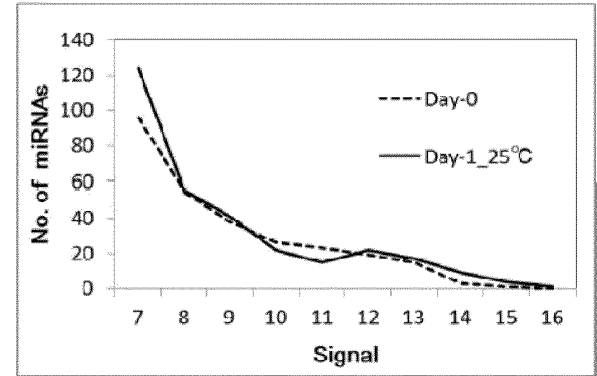
A



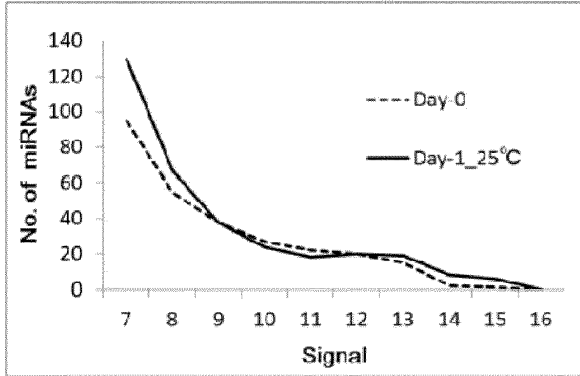
F



G



H



I

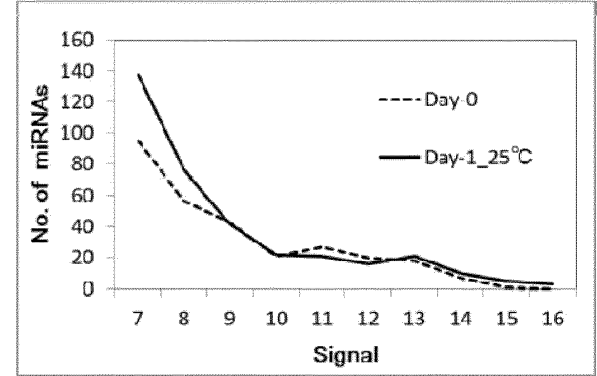


Fig. 8-2

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/076085

5	A. CLASSIFICATION OF SUBJECT MATTER C12Q1/68(2006.01)i, C12M1/00(2006.01)i, C12M1/34(2006.01)i, C12N15/09(2006.01)i, G01N33/53(2006.01)i, G01N37/00(2006.01)i	
	According to International Patent Classification (IPC) or to both national classification and IPC	
10	B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12Q1/68, C12M1/00, C12M1/34, C12N15/09, G01N33/53, G01N37/00	
15	Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2015 Kokai Jitsuyo Shinan Koho 1971-2015 Toroku Jitsuyo Shinan Koho 1994-2015	
20	Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) JSTPlus/JMEDPlus/JST7580 (JDreamIII), CAPLUS/REGISTRY/MEDLINE/BIOSIS (STN), GenBank/EMBL/DDBJ/GeneSeq, DWPI (Thomson Innovation)	
	C. DOCUMENTS CONSIDERED TO BE RELEVANT	
	Category*	Citation of document, with indication, where appropriate, of the relevant passages
25	Y	JP 2014-007995 A (Toray Industries, Inc.), 20 January 2014 (20.01.2014), claims (Family: none)
30	Y	WO 2014/034685 A1 (Kyoto University), 06 March 2014 (06.03.2014), claims; paragraph [0020] (Family: none)
35	Y	US 2009/0075258 A1 (LATHAM J. Gary et al.), 19 March 2009 (19.03.2009), claims; paragraphs [0056] to [0059] (Family: none)
40	<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.	
45	* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
50	Date of the actual completion of the international search 01 December 2015 (01.12.15)	Date of mailing of the international search report 08 December 2015 (08.12.15)
55	Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer  Telephone No.

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP2015/076085

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2009/133915 A1 (NEC Corp.), 05 November 2009 (05.11.2009), claims; paragraphs [0068] to [0079] & US 2011/0143360 A1 claims; paragraphs [0084] to [0101] & US 2012/0270746 A1 & US 2014/0170648 A1 & EP 2281903 A1	1-13
Y	WO 2014/048441 A1 (MIRRX THERAPEUTICS), 03 April 2014 (03.04.2014), SEQ ID NO:142, 439, 552, 576, 713, 1326, 1556 & EP 2900822 A1 & CA 2886116 A & CN 104955950 A	1-13
Y	LADEWIG Erik et al., "Discovery of hundreds of mirtrons in mouse and human small RNA data", Genome Research, 2012, Vol.22, No.9, p.1634-1645	1-13

Form PCT/ISA/210 (continuation of second sheet) (July 2009)

## REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

## Patent documents cited in the description

- JP 2007075095 A [0011]
- JP 2007097429 A [0011]
- JP 2014007995 A [0011]
- JP 4244788 B [0042]
- US 5705610 B, Ronald [0044]
- US 6142266 B, Michel [0044]
- US 7037659 B, Francesco [0044]
- JP 3922454 B, Hirota [0045]

## Non-patent literature cited in the description

- **ROBERTS, T. C.** *PLoS ONE*, 2014, vol. 9 (2), e89237 [0012]
- **CHEN, X. et al.** *PLoS ONE*, 2013, vol. 8 (11), e79652 [0012]
- **FAVALORO.** *Methods Enzymol.*, 1980, vol. 65, 718-749 [0033]
- **SAMBROOK, J. et al.** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 1998 [0039]
- *miRNA Experimental Protocols*. Yodosha Co., Ltd, 2008 [0047]
- *Microarray data statistical analysis protocols*. Yodosha Co., Ltd, 2008 [0054]
- **VOELLENKLE C et al.** *RNA*, 2012, vol. 18, 472-484 [0141]
- **BEREZIKOV E et al.** *Molecular Cell*, 2007, vol. 28, 328-336 [0142]
- **LI H et al.** *Journal of Clinical Investigation*, 2009, vol. 119, 3666-3677 [0143]
- **LAGOS-QUINTANA M et al.** *Current Biology*, 2002, vol. 12, 735-739 [0144]
- **JIMA DD et al.** *Blood*, 2010, vol. 116, e118-e127 [0145] [0146]
- **YOO JK et al.** *Stem Cells and Development*, 2012, vol. 21, 2049-2057 [0147]
- **LADEWIG E et al.** *Genome Research*, 2012, vol. 22, 1634-1645 [0148] [0149]
- **YOO H et al.** *Biochem Biophys Res Commun*, 2011, vol. 415, 567-572 [0150]
- **VANDESOMPELE et al.** *Genome Biology*, 2002 [0156]

专利名称(译)	用于比较分析miRNA表达水平的方法和装置		
公开(公告)号	<a href="#">EP3196316A1</a>	公开(公告)日	2017-07-26
申请号	EP2015841913	申请日	2015-09-15
[标]申请(专利权)人(译)	东丽株式会社		
申请(专利权)人(译)	TORAY INDUSTRIES, INC.		
当前申请(专利权)人(译)	TORAY INDUSTRIES, INC.		
[标]发明人	KOZONO SATOKO KONDOU SATOSHI		
发明人	KOZONO, SATOKO KONDOU, SATOSHI		
IPC分类号	C12Q1/68 C12M1/00 C12M1/34 C12N15/09 G01N33/53 G01N37/00		
CPC分类号	G16B40/00 C12M1/00 C12M1/34 C12N15/09 C12Q1/68 C12Q1/6816 C12Q1/6876 C12Q2600/158 C12Q2600/178 G01N33/53 G01N37/00 G16B25/00 C12Q2525/207 C12Q2537/165 C12Q2545/101		
优先权	2014187685 2014-09-16 JP		
其他公开文献	EP3196316A4 EP3196316A9		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

公开了一种用于比较分析多个体液样品中靶miRNA的表达水平的方法。在用于本发明的比较分析的方法中，使用校正内源miRNA的表达水平校正每个体液样品中靶miRNA的表达水平，所述校正内源miRNA与表达水平同时测量。样品中的靶miRNA。作为校正内源miRNA，使用选自特定10种校正内源miRNA的一种或多种miRNA。通过本发明，可以比常规技术更准确地进行体液样品中靶miRNA的比较分析。

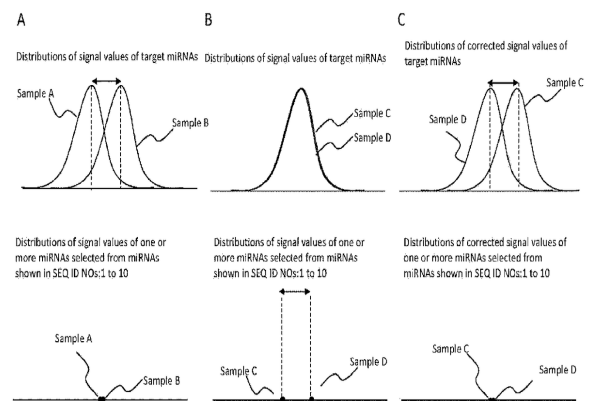


Fig. 1