

(19)



(11)

**EP 2 540 839 A1**

(12)

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

(43) Date of publication:

**02.01.2013 Bulletin 2013/01**

(51) Int Cl.:

**C12Q 1/68** (2006.01) **C12N 15/09** (2006.01)  
**G01N 33/50** (2006.01) **G01N 33/53** (2006.01)

(21) Application number: **11747526.9**

(86) International application number:

**PCT/JP2011/054363**

(22) Date of filing: **25.02.2011**

(87) International publication number:

**WO 2011/105570 (01.09.2011 Gazette 2011/35)**

(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**

- **TASHIMA, Yoshihiko**

**Kobe-shi  
Hyogo 651-0073 (JP)**

- **KAJITA, Masahiro**

**Kobe-shi  
Hyogo 651-0073 (JP)**

(30) Priority: **26.02.2010 JP 2010042814**

(71) Applicant: **Symex Corporation**  
**Kobe-shi, Hyogo 651-0073 (JP)**

(74) Representative: **Paemen, Liesbet R.J. et al**

**De Clercq & Partners cvba**  
**Edgard Gevaertdreef 10 a**  
**9830 Sint-Martens-Latem (BE)**

(72) Inventors:

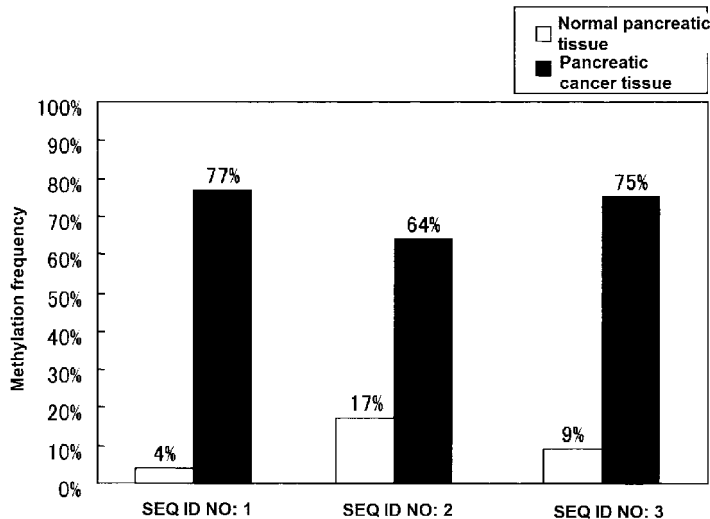
- **SAKAI, Ayako**  
**Kobe-shi**  
**Hyogo 651-0073 (JP)**

(54) **METHOD FOR DETERMINING PRESENCE OR ABSENCE OF CANCER CELL IN BIOLOGICAL SAMPLE, AND MOLECULAR MARKER AND KIT FOR DETERMINATION**

(57) The present invention relates to a method for determining presence or absence of a cancer cell in a biological sample based on the analysis result obtained

by analyzing methylation status of DNA extracted from the biological sample with a novel molecular marker allowing a determination of presence or absence of the cancer cell.

[Fig. 4]



**EP 2 540 839 A1**

**Description****TECHNICAL FIELD**

5 **[0001]** The present invention relates to a method for determination of presence or absence of a cancer cell in a biological sample. More specifically, the present invention relates to the method for determining presence or absence of a cancer cell in a biological sample based on an analysis result obtained by analyzing methylation status of a CpG site in a certain base sequence for a DNA extracted from the biological sample.  
 10 The present invention also relates to a molecular marker and a kit for determination which are used for the above method.

**BACKGROUND ART**

**[0002]** It has been known that chromosome DNAs of higher eukaryotes may undergo methylation at the 5-position of C (cytosine) among other bases constituting DNAs. Such DNA methylation functions as a mechanism for suppression of gene expression. For example, when a region rich in CpG sequences (also referred to as "CpG islands"), which often exists in promoter regions of certain genes, is methylated, transcription of these genes may be suppressed. This phenomenon is referred to as "gene silencing". On the other hand, when a CpG island is not methylated, a transcription factor can bind to the promoter region and the gene can be transcribed.

20 **[0003]** Accordingly, DNA methylation is one of control mechanisms of gene expression. For this reason, DNA methylation plays important roles in various physiological and pathological phenomena such as early embryonic development, expression of tissue specific genes, genomic imprinting and X chromosome inactivation which are characteristic to mammals, stabilization of chromosomes, synchronization of DNA replication and the like.

**[0004]** Recently, it has also been revealed that abnormal DNA methylation, i.e. gene silencing due to DNA methylation is involved in development and progress of diseases such as cancers. Accordingly, it has been attempted to diagnose diseases such as cancers by analyzing methylation status of DNA for various genes. Patent Literature 1, for example, discloses that genes such as NPTX2, SARP2 and CLDN5 and the like are not methylated in normal pancreatic cells while they are methylated in pancreatic cancer cells, and describes a method for detection of pancreatic cancer by analyzing methylation status of these genes. Patent Literature 2 discloses a method for detection of one type of breast cancers, proliferative breast disease, by analyzing methylation status of genes such as BRCA2 and PCDH7 and the like.

30 **[0005]** As described above, many genes have been reported which are abnormally methylated in various cancers. However, only few among them can be used as molecular markers for detection of cancers. Accordingly, there is a need for further development of novel molecular markers useful to be used in a detection method of a cancer cell utilizing gene methylation analysis.

## 35 Citation List

## Patent Literature

**[0006]**

40 Patent Literature 1: Japanese Translation of PCT International Application No. 2007-524369 (corresponding to PCT International Publication WO 2004/083399)

Patent Literature 2: Japanese Translation of PCT International Application No. 2008-506407 (corresponding to PCT International Publication WO 2006/008128)

45

**SUMMARY OF INVENTION**

## Technical Problem

50 **[0007]** An object of the present invention is to provide a method for determination of presence or absence of a cancer cell based on an analysis result obtained by analyzing methylation status of DNA extracted from a biological sample with novel molecular markers allowing determination of presence or absence of a cancer cell.  
 Another object of the present invention is to provide a molecular marker and a kit for determination which can be used in the above method.

55

## Solution to Problem

**[0008]** The present inventors analyzed methylation status of genomic DNA from cancer tissues and cancer cell lines

and identified gene regions in which DNA methylation was found specifically for cancer tissues and cancer cell lines. The present inventors further found that the analysis result obtained by analyzing methylation status for certain base sequences in these gene regions allows highly accurate differentiation between samples containing a cancer cell and the ones without a cancer cell, thereby completing the present invention.

5 [0009] Namely, the present invention provides a method for determination of presence or absence of a cancer cell in a biological sample obtained from a subject comprising the steps of:

- extracting DNA from the biological sample;
- 10 analyzing methylation status, for the DNA obtained in the extracting step, of at least one CpG site located in at least one base sequence selected from base sequences SEQ ID NOs: 1 to 14; and
- determining presence or absence of the cancer cell in the biological sample based on an analysis result obtained in the analyzing step.

15 [0010] The present invention also provides a molecular marker for determination of presence or absence of a cancer cell by analysis of methylation status which is at least one CpG site selected from CpG sites located in base sequences SEQ ID NOs: 1 to 14.

[0011] The present invention further provides a kit for determination of presence or absence of a cancer cell in a biological sample obtained from a subject comprising:

- 20 a non-methylated cytosine conversion agent that converts non-methylated cytosine in DNA extracted from the biological sample to a different base; and
- a primer set for determination of methylation status of at least one CpG site located in at least one base sequence selected from base sequences SEQ ID NOs: 1 to 14 by methylation specific PCR.

#### 25 Advantageous Effects of Invention

[0012] According to the present method for determination of presence or absence of a cancer cell in a biological sample (hereinafter also referred to as "the present method"), presence or absence of a cancer cell in a biological sample obtained from a subject can be determined by analyzing methylation status of the present molecular marker(s) for DNA extracted from the biological sample.

The present invention can further provide the molecular markers and the kit which can be used for the method.

#### BRIEF DESCRIPTION OF DRAWINGS

35 [0013]

Fig. 1 is a table showing methylation status of the 1st to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1 in DNA obtained from 4 pancreatic cancer tissue samples and 5 normal pancreatic tissue samples;

40 Fig. 2 is a table showing methylation status of the 1st to 79th CpG sites from 5' end of the base sequence SEQ ID NO: 2 in DNA obtained from 5 pancreatic cancer tissue samples and 5 normal pancreatic tissue samples;

Fig. 3 is a table showing methylation status of the 1st to 12th CpG sites from 5' end of the base sequence SEQ ID NO: 3 in DNA obtained from 5 pancreatic cancer tissue samples and 4 normal pancreatic tissue samples;

Fig. 4 is a bar graph showing methylation frequency of CpG sites located in the respective base sequences SEQ ID NOs: 1 to 3;

45 Fig. 5 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 15 in DNA obtained from pancreatic cancer tissue and normal tissue by a methylation specific PCR (MSP) method;

Fig. 6 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 16 in DNA obtained from pancreatic cancer tissue and normal tissue by the MSP method;

50 Fig. 7 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 17 in DNA obtained from pancreatic cancer tissue and normal tissue by the MSP method;

Fig. 8 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 18 in DNA obtained from pancreatic cancer tissue and normal tissue by the MSP method;

55 Fig. 9 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 19 in DNA obtained from pancreatic cancer tissue and normal tissue by the MSP method;

Fig. 10 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 15 in DNA obtained from large intestinal tissue and mammary gland tissue by the MSP method;

Fig. 11 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base

sequence SEQ ID NO: 18 in DNA obtained from mammary gland tissue by the MSP method;

Fig. 12 is a table showing methylation status of the 1st to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7 in 6 samples of a breast cancer cell line MCF7 and 5 samples of a normal mammary gland epithelial cell line HMEC;

5 Fig. 13 is a table showing methylation status of the 1st to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8 in 4 MCF7 cell samples and 4 HMEC cell samples;

Fig. 14 is a table showing methylation status of the 1st to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9 in 6 MCF7 cell samples and 6 HMEC cell samples;

10 Fig. 15 is a table showing methylation status of the 1st to 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10 in 5 MCF7 cell samples and 5 HMEC cell samples;

Fig. 16 is a table showing methylation status of the 1st to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 11 in 4 MCF7 cell samples and 4 HMEC cell samples;

15 Fig. 17 is a table showing methylation status of the 1st to 32nd CpG sites from the 5' end of the base sequence SEQ ID NO: 12 in 6 MCF7 cell samples and 6 HMEC cell samples;

Fig. 18 is a table showing methylation status of the 1st to 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13 in 5 MCF7 cell samples and 4 HMEC cell samples;

Fig. 19 is a table showing methylation status of the 1st to 24th CpG sites from the 5' end of the base sequence SEQ ID NO: 14 in 4 MCF7 cell samples and 4 HMEC cell samples;

20 Fig. 20 a bar graph showing methylation frequency of CpG sites located in the respective base sequences SEQ ID NOs: 7 to 14;

Fig. 21 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 20 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

Fig. 22 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 21 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

25 Fig. 23 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 22 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

Fig. 24 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 23 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

30 Fig. 25 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 24 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

Fig. 26 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 25 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

35 Fig. 27 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 26 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

Fig. 28 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 27 in DNA obtained from MCF7 cells and HMEC cells by the MSP method;

40 Fig. 29 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 20 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

Fig. 30 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 21 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

45 Fig. 31 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 22 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

Fig. 32 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 23 in DNAs obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

50 Fig. 33 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 24 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

Fig. 34 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 25 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

55 Fig. 35 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base sequence SEQ ID NO: 26 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

Fig. 36 is a bar graph showing fluorescence intensity of the bands of PCR products obtained by amplifying the base

sequence SEQ ID NO: 27 in DNA obtained from breast cancer tissue and normal mammary gland tissue by the MSP method;

Fig. 37 shows photographs obtained after agarose electrophoresis of reaction solutions resulting from amplification of base sequences SEQ ID NOs: 17 to 19, 21, 22 and 25 to 27 in DNA obtained from colon cancer tissue and normal large intestinal tissue by the MSP method; and

Fig. 38 shows photographs obtained after agarose electrophoresis of reaction solutions resulting from amplification of base sequences SEQ ID NOs: 20 to 27 in DNA obtained from MCF7 cells, normal heart tissue, normal kidney tissue, normal liver tissue, normal lung tissue and normal peripheral blood leukocytes by the MSP method.

## DESCRIPTION OF EMBODIMENTS

**[0014]** As used herein, the term "CpG site" means a site of a sequence in a base sequence in which cytosine (C) and guanine (G) are adjacent in this order from 5' to 3'. The letter "p" in "CpG" represents a phosphodiester bond between cytosine and guanine.

**[0015]** As used herein, to "analyze methylation status" means to analyze presence or absence of methylation of a CpG site located in a base sequence to be analyzed or methylation frequency of a CpG site(s) in the base sequence. In this context, a base sequence to be analyzed is not specifically limited so long as it is a region comprising at least one CpG site among at least one base sequence selected from the base sequences SEQ ID NOs: 1 to 14.

The term "presence or absence of methylation" means whether or not a cytosine residue in a CpG site located in a base sequence to be analyzed is methylated.

The term "methylation frequency" means a ratio of the number of methylated CpG site(s) relative to the total number of CpG site(s) located in a base sequence to be analyzed. According to the present invention, the position and number of CpG site(s) located in each of the base sequences SEQ ID NOs: 1 to 14 are known. Accordingly, if the base sequence to be analyzed is selected among the base sequences SEQ ID NOs: 1 to 14, the number of CpG site(s) located therein can be understood beforehand. Therefore, the number of methylated CpG site(s) itself in the base sequence to be analyzed can also be used as the methylation frequency.

**[0016]** According to the present method, DNA is first extracted from a biological sample obtained from a subject. In this extracting step, the biological sample is not specifically limited so long as it contains DNA of a subject, and is preferably a sample containing genomic DNA, e.g. a clinical specimen. The clinical specimen can specifically include blood, serum, plasma, lymph fluid, urine, nipple discharge, tissue and cells obtained by operations and biopsies and the like.

**[0017]** DNA can be extracted from a biological sample by well-known extraction methods. Extraction of DNA can be carried out, for example, by mixing a biological sample with a treatment solution containing a surfactant for solubilization of cells or tissues (e.g. sodium cholate, sodium dodecyl sulfate etc.), and subjecting the resulting mixture to physical procedure (stirring, homogenization, ultrasonication etc.) to release DNA contained in the biological sample into the mixture. In this case, it is preferable to centrifuge the mixture to precipitate cell debris, collect the supernatant containing DNA and use the supernatant to the next step of analyzing. The obtained supernatant can be purified according to well-known methods. DNA can also be extracted and purified from a biological sample by using commercially available kits.

**[0018]** The extracting step preferably further comprises a step of fragmenting the extracted DNA. According to this step, DNA can be fragmented into appropriate lengths, allowing effective methylated DNA immunoprecipitation (MeDIP) and non-methylated cytosine conversion being carried out.

Fragmentation of DNA can be carried out by ultrasonication, alkaline treatment, restriction enzyme treatment and the like. When the alkaline treatment is carried out with sodium hydroxide, sodium hydroxide is added to a DNA solution to the final concentration of 0.1 to 1.0N and the mixture is incubated at 10 to 40°C for 5 to 15 minutes to fragment the DNA. In case of the restriction enzyme treatment, the restriction enzyme can appropriately be selected based on the base sequence of DNA, which can be *MseI* or *BamHI*, for example.

**[0019]** Next, according to the present method, methylation status of at least one CpG site located in at least one base sequence selected from the base sequences SEQ ID NOs: 1 to 14 is analyzed for the DNA extracted from the biological sample.

**[0020]** All of the above base sequences SEQ ID NOs: 1 to 14 are partial regions of human genomic DNA. More specifically, these base sequences are parts of gene regions described in Table 1 or base sequence regions in the vicinity thereof. The base sequences per se are well-known and can be obtained from well-known databases such as the one provided by the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/gene/>).

**[0021]**

[Table 1]

SEQ ID NO:	Entrez gene ID	Gene symbol	Gene title
1	3205	HOXA9	Homeobox A9
2	55539	KCNQ1DN	KCNQ1 downstream neighbor
3	5913	RAPSN	Receptor-associated protein of the synapse
4 and 5	8326	FZD9	Frizzled homolog 9
6	64405	CDH22	Cadherin 22
7	85474	LBX2	Ladybird homeobox 2
8	5083	PAX9	Paired box 9
9	120	ADD3	Adducin 3
10	232933	CCDC61	Coiled coli domain containing 61
11	84107	ZIC4	Zic family member 4
12	94027	CGB7	Chorionic gonadotropin beta polypeptide 7
13	9760	TOX	Thymocyte selection associated high mobility group box
14	7137	TNNI3	Troponin I type 3

**[0022]** In the analyzing step, the CpG site to be analyzed is preferably selected from the followings:

the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence SEQ ID NO: 2;  
the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;  
the 1st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;  
the 1st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;  
the 1st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;  
the 1st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;  
the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;  
the 1st, 2nd, 4th, 7th to 11th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;  
the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;  
the 1st to 3rd, 5th to 11th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 11;  
the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ ID NO: 12;  
the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and  
the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 14.

**[0023]** The analyzing step can be a step of analyzing presence or absence of methylation of a CpG site located in the base sequences SEQ ID NOs: 1 to 14.

In this case, one CpG site can be analyzed; however, in order to improve determination accuracy in the subsequent determining step, more than one CpG site is preferably analyzed for presence or absence of methylation. More than one CpG site can be selected from one base sequence or from each of multiple base sequences.

**[0024]** Alternatively, the analyzing step can be a step of analyzing methylation frequency of a CpG site(s) located in the base sequences SEQ ID NOs: 1 to 14.

In this case, the base sequence to be analyzed can contain only one CpG site; however, in order to improve determination accuracy in the subsequent determining step, it is preferable that the base sequence to be analyzed contains more than one CpG site. The base sequence to be analyzed can be any one of the base sequences SEQ ID NOs: 1 to 14; however, more than one base sequence among them is preferably analyzed.

**[0025]** Various methods are well-known as the method for analyzing methylation status. According to the present method, it is not specifically limited as to which analysis method is used; however, the analysis method preferably comprises differentiating between methylated DNA and non-methylated DNA, amplifying DNA and detecting methylated DNA and/or non-methylated DNA.

**[0026]** The step of differentiation between methylated DNA and non-methylated DNA can include a step of carrying out methylation sensitive restriction enzyme treatment, a MeDIP method, non-methylated cytosine converting treatment and the like.

The step of amplifying DNA can include a step of carrying out PCR amplification, quantitative PCR amplification, IVT (in *vitro* transcription) amplification, SPIA™ amplification and the like methods.

The step of detecting methylated DNA and/or non-methylated DNA can include a step of carrying out electrophoresis, sequence analysis, microarray analysis, mass spectrometry, Southern hybridization and the like.

**[0027]** The MeDIP method is the method in which methylated DNA in a biological sample is concentrated by immunoprecipitation using an anti-methylated cytosine antibody or an anti-methylated cytidine antibody, or an antibody which specifically recognizes a methylated DNA-binding protein. According to the analyzing step of the present method, methylated DNA comprised in DNA obtained in the extracting step can be concentrated by the MeDIP method and methylation status of the concentrated methylated DNA can be analyzed.

The methylated DNA concentrated by the MeDIP method can be amplified by e.g. IVT amplification and methylation status of the obtained amplified products can be analyzed by using a microarray. These analysis procedures are referred to as the MeDIP on chip method.

**[0028]** The non-methylated cytosine converting treatment is the one in which DNA extracted from a biological sample is subjected to a reaction with a non-methylated cytosine conversion agent so as to convert non-methylated cytosine(s) in the DNA to a different base (uracil, thymine, adenine or guanine).

In this context, the non-methylated cytosine conversion agent is a substance which can react with DNA and convert non-methylated cytosine in the DNA to a different base (uracil, thymine, adenine or guanine). The non-methylated cytosine conversion agent suitably used can be, for example, bisulfite such as sodium, potassium, calcium or magnesium bisulfite.

In the treatment using bisulfite, non-methylated cytosine(s) in DNA is converted to uracil due to deamination reaction, while a methylated cytosine does not undergo such a base conversion.

Thus, the difference in methylation status in DNA is converted to the difference in a base sequence (C and U) by the non-methylated cytosine converting treatment using bisulfite. The non-methylated cytosine converting treatment using bisulfite is referred to as bisulfite treatment.

**[0029]** When the bisulfite treatment is carried out, the amount added of bisulfite (concentration) is not specifically limited so long as it can sufficiently convert non-methylated cytosine(s) in DNA, and corresponds to, for example, 1M or higher, preferably 1 to 15M, more preferably 3 to 10M as the final concentration in the sample. The incubation conditions (temperature and time) after the addition of bisulfite to a biological sample can be appropriately selected according to the amount added of bisulfite, and for example, when bisulfite is added at the final concentration of 6M, the incubation is carried out at 50 to 80°C for 10 to 90 minutes.

**[0030]** Methylation status of DNA can be analyzed by sequencing DNA after bisulfite treatment and detecting the difference in base sequence from the original sequence. This process is referred to as a bisulfite sequencing method.

**[0031]** Methylation status of DNA can also be analyzed by amplifying DNA after bisulfite treatment by PCR using a primer set described hereinafter and determining presence or absence of a PCR product. This analysis procedure is referred to as a methylation specific PCR (MSP) method.

**[0032]** The MSP method utilizes a primer set which can amplify a base sequence in which cytosine in a CpG site to be analyzed is methylated (i.e. cytosine is not converted to uracil) while which can not amplify a base sequence in which cytosine in a CpG site is not methylated (i.e. cytosine is converted to uracil). According to the MSP method using such a primer set, presence of the PCR product indicates methylation of the CpG site analyzed.

The MSP method can also be carried out by using a primer set which can not amplify a base sequence in which cytosine in a CpG site to be analyzed is not converted to uracil, while which can amplify a base sequence in which cytosine in a CpG site is converted to uracil. In this case, absence of the PCR product indicates methylation of the CpG site analyzed.

**[0033]** Each primer in the above primer sets can be appropriately designed by a person skilled in the art based on the base sequence comprising a CpG site to be analyzed, and it is preferably designed so as to contain cytosine of the CpG site to be analyzed at the 3' end or in the vicinity thereof of the primer.

**[0034]** Methylation status can be analyzed with a microarray in the analyzing step of the present method. In this case, the microarray for analysis can be prepared by immobilizing one or more nucleic probes complementary to the base sequences SEQ ID NOs: 1 to 14 on a substrate. The microarray can be prepared according to well-known methods in the art.

**[0035]** In the analysis using a microarray, DNA extracted from a biological sample is preferably labeled with a labeling substance well-known in the art. Thus, the present method preferably further comprises a step of labeling the extracted DNA. The labeling step is advantageously carried out after the amplifying step because all DNA in the biological sample can be labeled. The labeling substance can include fluorescence substances, haptens such as biotin, radioactive substances and the like. The fluorescence substances can include Cy3, Cy5, Alexa Fluor™, FITC and the like.

Labeling of DNA facilitates measurement of a signal from a probe on a microarray. A method for labeling DNA with the

labeling substance is well-known in the art.

**[0036]** The above signal can be any suitable signal according to the type of microarrays. For example, the signal can be an electric signal generated when a DNA fragment hybridizes to a probe on the microarray, or a fluorescence or luminescence signal generated from a labeling substance when DNA to be analyzed is labeled as described above.

5 Detection of signal can be carried out by using a scanner comprised in a conventional microarray analyzer. The scanner can be, for example, GeneChip® Scanner3000 7G (Affymetrix) and the like.

**[0037]** In the present method, presence or absence of a cancer cell in a biological sample is determined based on the analysis result obtained in the analyzing step.

10 In the determining step, when the result obtained in the analyzing step indicates that there is a methylated CpG site, it can be determined that a cancer cell is present in a biological sample obtained from a subject. On the other hand, when the result obtained indicates that there is no methylated CpG site, it can be determined that no cancer cell is present in the biological sample.

The determination can be carried out based on the analysis result of one CpG site; however, in order to improve determination accuracy, it is preferable that the determination is made based on the analysis result of more than one  
15 CpG site.

**[0038]** In the determining step, when the result obtained in the analyzing step indicates that methylation frequency is higher than a predetermined threshold, it can be determined that a cancer cell is present in a biological sample obtained from a subject. On the other hand, when the result obtained indicates that methylation frequency is lower than a pre-  
20 determined threshold, it can be determined that no cancer cell is present in a biological sample obtained from a subject.

The above threshold can be predetermined as follows. First, methylation frequency is analyzed for DNA extracted from a biological sample which is confirmed to be devoid of cancer cells (normal tissue or normal cells) and a biological sample containing a cancer cell. Next, based on the obtained analysis results, a threshold is determined within a range which is higher than the methylation frequency of the biological sample devoid of cancer cells and lower than that of the biological sample containing a cancer cell. Preferably, a threshold is determined as a value which can highly accurately  
25 differentiate between the biological sample devoid of cancer cells and the biological sample containing a cancer cell.

**[0039]** The present invention also encompasses a molecular marker for determination of presence or absence of a cancer cell which can be used for the present method (hereinafter also referred to as "the present molecular marker"). The present molecular marker is at least one selected from CpG sites located in the base sequences SEQ ID NOs: 1 to 14.

**[0040]** The present molecular marker is preferably selected from the following CpG sites:

30 the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence  
SEQ ID NO: 2;

35 the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;

the 1 st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;

the 1 st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;

the 1 st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;

the 1 st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;

40 the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;

the 1 st, 2nd, 4th, 7th to 11 th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;

the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;

the 1 st to 3rd, 5th to 11 th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence  
SEQ ID NO: 11;

45 the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ ID NO: 12;

the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and

the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 14.

**[0041]** The present invention also encompasses a kit for determination of presence or absence of a cancer cell in a biological sample for implementing the present method (hereinafter also referred to as "the present kit").

The present kit comprises a non-methylated cytosine conversion agent and a primer set for determination of methylation status of at least one CpG site located in the base sequence SEQ ID NOs: 1 to 14 by methylation specific PCR.

**[0042]** The non-methylated cytosine conversion agent comprised in the present kit is not specifically limited so long as it can react with DNA extracted from a biological sample obtained from a subject and convert non-methylated cytosine  
55 in the DNA to a different base (uracil, thymine, adenine or guanine), and is preferably bisulfite. The non-methylated cytosine conversion agent can be in the form of a solution or in the form of a solid which can be dissolved in an appropriate solvent before use to provide a solution.

**[0043]** When the present kit is used, a CpG site selected from the followings is preferably analyzed by the MSP method:

## EP 2 540 839 A1

the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence SEQ ID NO: 2;

- 5 the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;  
the 1 st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;  
the 1 st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;  
the 1 st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;  
the 1 st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;  
10 the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;  
the 1 st, 2nd, 4th, 7th to 11 th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;  
the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;  
the 1 st to 3rd, 5th to 11 th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 11;  
15 the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ ID NO: 12;  
the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and  
the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 14.

[0044] More preferably, a CpG site selected from the followings is analyzed by the MSP method:

- 20 the 9th, 10th and 28th to 30th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
the 4th to 7th CpG sites from the 5' end of the base sequence SEQ ID NO: 2;  
the 6th and 9th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;  
the 3rd, 4th, 11th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 7;  
25 the 8th, 9th and 14th to 16th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;  
the 15th, 16th and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;  
the 2nd and 6th CpG sites from the 5' end of the base sequence SEQ ID NO: 10  
the 1st to 3rd and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 11;  
the 10th, 11 th and 16th to 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 12;  
30 the 7th, 8th and 15th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and  
the 1st, 2nd and 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 14.

[0045] The primer set comprised in the present kit is preferably selected from the followings:

- 35 a primer set of the base sequences SEQ ID NO: 28 and SEQ ID NO: 29;  
a primer set of the base sequences SEQ ID NO: 30 and SEQ ID NO: 31;  
a primer set of the base sequences SEQ ID NO: 32 and SEQ ID NO: 33;  
a primer set of the base sequences SEQ ID NO: 34 and SEQ ID NO: 35;  
a primer set of the base sequences SEQ ID NO: 36 and SEQ ID NO: 37;  
40 a primer set of the base sequences SEQ ID NO: 38 and SEQ ID NO: 39;  
a primer set of the base sequences SEQ ID NO: 40 and SEQ ID NO: 41;  
a primer set of the base sequences SEQ ID NO: 42 and SEQ ID NO: 43;  
a primer set of the base sequences SEQ ID NO: 44 and SEQ ID NO: 45;  
a primer set of the base sequences SEQ ID NO: 46 and SEQ ID NO: 47;  
45 a primer set of the base sequences SEQ ID NO: 48 and SEQ ID NO: 49;  
a primer set of the base sequences SEQ ID NO: 50 and SEQ ID NO: 51; and  
a primer set of the base sequences SEQ ID NO: 52 and SEQ ID NO: 53.

[0046] The above primer set is suitably used when the present method is carried out by analyzing presence or absence of methylation of a CpG site located in the base sequences SEQ ID NOs: 1 to 14 by the MSP method.

[0047] The present invention is hereinafter illustrated in further detail by means of examples, which do not limit the present invention. Examples

Example 1: Identification of novel markers from genomic DNA of pancreatic cancer tissue

- 55 [0048] Genomic DNA from human pancreatic cancer tissue and human normal pancreatic tissue was analyzed by a microarray to search gene regions which is specifically methylated in the genomic DNA from pancreatic cancer tissue. Specific operation procedures carried out in Example 1 followed the instructions attached to the kits and reagents used.

## 1. Preparation of methylated DNA by MeDIP method

[0049] Human pancreatic cancer tissue and human normal pancreatic tissue were used as biological samples from which genomic DNA was extracted. The extracted genomic DNA (4 µg) was incubated with the restriction enzyme *MseI* (NEB Inc.) at 37°C overnight to fragment the DNA into the sizes from 300 to 1000 bp. The reagents were then denatured by heating at 95°C for 10 minutes followed by quenching to 4°C to obtain single-stranded DNA fragments.

The obtained single-stranded DNA fragments were diluted in a dilution buffer (302 µl) included in the Chromatin Immunoprecipitation assay kit (Upstate Biotechnology). The diluted solutions were then added with Protein G Sepharose beads (68 µl: GE Healthcare), rotated at 4°C for an hour and centrifuged to remove proteins and the like which non-specifically bound to the beads (this procedure is hereinafter also called as "pre-clear treatment"). The supernatants were collected and then respectively divided into two aliquots in separate tubes, one of which was added with an anti-methylated cytosine antibody BI-MECY-0500 (10 µg: Eurogentec) and the other with no antibody, serving as a test sample and a control sample (Input), respectively. The respective test samples were rotated overnight at 4°C prior to addition of Protein G Sepharose beads (68 µl: GE Healthcare) and rotation at 4°C for an additional hour. The samples were then centrifuged, the supernatants were removed and the beads were collected. The obtained beads were washed with a washing buffer included in the above kit, and added with an elution buffer (250 µl) to elute single-stranded DNA fragments.

The obtained DNA solutions were incubated with Proteinase K (Sigma) before purification with the Qiaquick PCR purification kit (QIAGEN).

Quantitative PCR allowed confirming that methylated DNA in respective test samples was specifically concentrated.

## 2. Amplification and labeling of nucleic acid in samples

[0050] For the test samples and control samples respectively obtained from the human pancreatic cancer tissue and human normal pancreatic tissue, DNA dephosphorylation was carried out with CIP (Calf intestine phosphatase) (New England Biolab). DNA was then added with dATP at its 3' end using Terminal transferase (ROCHE) and purified using the MinElute purification kit (QIAGEN). The single-stranded DNA fragments in the samples were converted to double-stranded DNA by using the GeneChip® One-Cycle Target Labeling kit (Affymetrix). The resulting double-stranded DNA was used as a template and labeled with biotin by IVT amplification to obtain biotinylated cRNA from each sample. The concentration of the obtained cRNA was determined by measuring the absorbance at 260 nm and 280 nm.

The cRNA from each sample was fragmented with the GeneChip® Sample Cleanup Module (Affymetrix) to obtain test and control samples for microarray analysis.

## 3. Microarray analysis

## (1) Contact of sample and microarray

[0051] The samples for microarray analysis were brought into contact with the microarray GeneChip® Human Promoter 1.0R Array (Affymetrix) to carry out hybridization with probes. One same type of microarray was used for each of the test sample and the control sample. Staining, washing and scanning (measurement of a signal) after contacting with the microarray were carried out according to the instruction provided by Affymetrix.

## (2) Analysis of microarray data

[0052] The resulting microarray data was analyzed according to the following procedures. In this analysis, a conventional computer was used comprising a central processing unit, a memory part, an input part such as a keyboard and an output part such as a display.

(Step 1) The base sequences of all probes on the microarray, the base sequence of the genomic DNA and the signal values measured as above were entered in the computer.

(Step 2) The base sequence of the genomic DNA was sectioned at the recognition sequence "TACC" of the restriction enzyme *MseI* to obtain the base sequences of the resulting DNA fragments. Among these base sequences, the base sequences of DNA fragments were extracted which have 300 bp or more in length and do not comprise the sequence "CG". Accordingly, the base sequences of DNA fragments of 300 bp or more without the "CG" sequence were obtained (hereinafter also referred to as "fragment sequences").

(Step 3) Among all probes on the microarray, probes which do not contain the "CG" sequence were extracted. Probes which were complementary to the above fragment sequences among the extracted probes were selected as corrective probes.

## EP 2 540 839 A1

(Step 4) Signal values obtained with the corrective probes were extracted from the measured signal values entered and their statistically representative value (mode) was calculated. This value was used as the background value. This procedure was carried out respectively for the test samples and the control samples.

(Step 5) The background value was subtracted from the respective measured signal values entered and the obtained values served as corrected values. When a corrected value is negative, the corrected value was regarded as zero. This procedure was also carried out respectively for the test samples and the control samples.

(Step 6) The Wilcoxon's signed rank sum test was carried out for the corrected values of the test and control samples to calculate the significance probability.

**[0053]** The obtained significance probability was visualized with a default browser IGB (Integrated Genome Browser; IGB). The significance probability displayed in IGB is the one which has been converted with the formula: (converted value) =  $-10 \log_{10}(\text{significance probability})$ . When the converted value is 20 or more, it is considered that a gene region containing a CpG site which is specifically methylated in the genomic DNA of pancreatic cancer tissue could be significantly detected.

As a result, the regions represented by the base sequences SEQ ID NOs: 15 to 19 were identified as the gene regions which were specifically methylated in genomic DNA of pancreatic cancer tissue. The base sequences SEQ ID NOs: 15 to 19 are the gene regions or partial base sequence regions in the vicinity thereof of, respectively, *HOXA9*, *KCNQ1DN*, *RAPSN*, *FZD9* and *CDH22* and the regions comprising the base sequences SEQ ID NOs: 1, 2, 3, 4 and 5 as well as 6, respectively.

The correspondence relationship between the base sequences SEQ ID NOs: 15 to 19 and SEQ ID NOs: 1 to 6 and the above genes is summarized in the following Table 2.

**[0054]**

[Table 2]

SEQ ID NO:	Corresponding SEQ ID NO:	Gene
15	1	HOXA9
16	2	KCNQ1DN
17	3	RAPSN
18	4 and 5	FZD9
19	6	CDH22

Example 2: Analysis of methylation status by bisulfite sequencing

**[0055]** By using bisulfite sequencing, it was investigated whether CpG sites in the gene regions represented by the base sequences SEQ ID NOs: 15 to 17 identified in Example 1 were methylated in genomic DNA from pancreatic cancer tissue and human normal pancreatic tissue.

### 1. Bisulfite treatment

**[0056]** To genomic DNA (2  $\mu\text{g}$ ) extracted from human pancreatic cancer tissue and human normal pancreatic tissue used in Example 1 was added 300  $\mu\text{l}$  of a 0.3N sodium hydroxide solution prior to incubation at 37°C for 10 minutes in order to denature the genomic DNA. To each DNA solution was added 300  $\mu\text{l}$  of a 10M sodium bisulfite solution prior to incubation at 80°C for 40 minutes for bisulfite treatment. DNA was purified from the solutions after bisulfite treatment using the Qiaquick PCR purification kit (QIAGEN) to prepare analysis samples.

### 2. Sequence analysis

**[0057]** PCR was carried out with the following primer sets and DNA in the obtained analysis samples as a template.

#### (i) Preparation of PCR reaction solution

**[0058]** A reaction solution (15  $\mu\text{l}$ ) was prepared by mixing the following reagents.

10 x Ex Taq<sup>®</sup> buffer (TAKARA Bio)    1.5  $\mu\text{l}$

## EP 2 540 839 A1

(continued)

	dNTP mix (2.5 mM)	1.2 µl
	F primer (10 µM)	0.6 µl
5	R primer (10 µM)	0.6 µl
	Analysis sample (template)	1.0 µl
	Ex Taq® polymerase (TAKARA Bio)	0.12 µl
	Distilled water	9.98 µl
10	<hr/>	
	Total	15.0 µl

(ii) Primer sequences and reaction conditions

**[0059]** (Primer sequences for amplifying the base sequence SEQ ID NO: 15)

15

F: 5' - TAGTTAGGGATAAAGTGTGAGTGTTA -3' (SEQ ID NO: 54)

R: 5' - CAACTTATTAATAACTATACTTCCCC -3' (SEQ ID NO: 55)

(Reaction condition)

20

**[0060]**

1 cycle of 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 55°C for 15 seconds and 72°C for 30 seconds;

25

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

**[0061]** (Primer sequences for amplifying the base sequence SEQ ID NO: 16)

30

F: 5' - GGGATATTTGTTTTTATATTTAATAAAGT -3' (SEQ ID NO: 56)

R: 5' - ACCTCACAATAAACTACTACAACC -3' (SEQ ID NO: 57)

(Reaction condition)

35

**[0062]**

1 cycle of 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 56°C for 15 seconds and 72°C for 40 seconds;

40

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

**[0063]** (Primer sequences for amplifying the base sequence SEQ ID NO: 17)

45

F: 5' - AGGTATTTATGGGGTAGGAATTATA -3' (SEQ ID NO: 58)

R: 5' - AAATCACTTAATAAAATCCCCTA -3' (SEQ ID NO: 59)

(Reaction condition)

50

**[0064]**

1 cycle at 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 56°C for 15 seconds and 72°C for 40 seconds;

1 cycle of 72°C for 7 minutes; and

55

keep at 4°C.

**[0065]** The obtained PCR products were incorporated into the pCR®2.1 vector included in the TA cloning kit (Invitrogen). Competent cells (TOP10; Invitrogen) were transformed with the obtained plasmid constructs, and the plasmids were extracted and purified from the transformants by using the GenElute Plasmid Miniprep kit (SIGMA). The base sequences

of the PCR products included in the obtained plasmids were analyzed with the Applied Biosystems 3730x1 DNA Analyzer (Applied Biosystems).

After the sequencing, methylation status of CpG sites in the base sequences of the above SEQ ID NOs was analyzed.

**[0066]** Methylation status of CpG sites located in the base sequences SEQ ID NOs: 1 to 3 which are respectively included in the base sequences SEQ ID NOs: 15 to 17 is shown in the tables in Figs. 1 to 3, which is based on the sequencing results described above. In the tables in these Figures, "●", "○" and "-" represent a methylated CpG site, a non-methylated CpG site and an unanalyzable CpG site, respectively. In addition, "PT" and "PN" represent pancreatic cancer tissue and normal pancreatic tissue, respectively.

In the tables in these Figures, the numbers in the rows correspond to the numbers representing the position of the CpG sites from the 5' end of the respective base sequences SEQ ID NOs: 1 to 3. The CpG site with the symbol "\*" or "\*" on the number is the CpG site which tends to be methylated in the pancreatic cancer tissue samples compared to the normal pancreatic tissue samples.

**[0067]** Figs. 1 to 3 show that in all base sequences SEQ ID NOs: 1 to 3, CpG sites tend to be methylated in pancreatic cancer tissue but not in normal pancreatic tissue.

Based on these results, methylation frequency in each of the base sequences SEQ ID NOs: 1 to 3 was analyzed. For example, in the case of the base sequence SEQ ID NO: 1, methylation frequency in pancreatic cancer tissue (PT) is the value calculated by dividing the number of methylated CpG sites "●" in all clones (clones 1 to 4) by the sum of the number of methylated CpG site "●" in all clones and the number of non-methylated CpG site "○" and converting the solution into percentage. Similarly, methylation frequency in normal pancreatic tissue (PN) is the value calculated by dividing the number of methylated CpG sites "●" in all clones (clones 1 to 5) by the sum of the number of methylated CpG site "●" in all clones and the number of non-methylated CpG site "○" and converting the solution into percentage. Methylation frequency in the respective base sequences obtained as above is represented as a bar graph in Fig. 4. This graph shows the average values for the group of the pancreatic cancer tissue samples and the group of the normal pancreatic tissues.

**[0068]** Fig. 4 shows that, when all CpG sites located in the base sequences SEQ ID NOs: 1 to 3 are used as targets for analysis, a threshold for differentiating between the samples containing a cancer cell and those devoid of cancer cells can be selected from the range of 20 to 60%, for example.

Accordingly, it was suggested that, when methylation frequency of a CpG site in the base sequence(s) SEQ ID NO(s): 1, 2 and/or 3 in DNA extracted from a biological sample obtained from a subject is higher than the above threshold, it can be determined that a cancer cell is present in the biological sample.

**[0069]** As the present molecular marker for determination, a CpG site which tends to be methylated in cancer tissue but not in normal tissue is suitable.

Accordingly, when presence or absence of a cancer cell in a sample is determined based on methylation frequency of a CpG site in the base sequences SEQ ID NOs: 1 to 3, more accurate determination can be made when the CpG site(s) as the molecular marker(s) for determination is selected from preferably the CpG site(s) with "\*" or "\*" more preferably the CpG site(s) with "\*" in Figs. 1 to 3.

When presence or absence of a cancer cell in a sample is determined based on the analysis result on presence or absence of methylation of a CpG site(s) in the base sequences SEQ ID NOs: 1 to 3 by methylation specific PCR, more accurate determination can be made when primers are used which can amplify preferably the CpG site(s) with "\*" or "\*", more preferably the CpG site(s) with "\*" in Figs. 1 to 3.

**[0070]** Example 3: Analysis of methylation status by methylation specific PCR (MSP)

Methylation status of CpG sites located in the gene regions of SEQ ID NOs: 15 to 19 identified in Example 1 was studied.

#### 1. Bisulfite treatment

**[0071]** Genomic DNA (2 μg each) obtained from human pancreatic cancer tissue, human normal pancreatic tissue, human normal heart tissue, human normal kidney tissue, human normal liver tissue, human normal lung tissue and human normal peripheral blood leukocytes was subjected to bisulfite treatment in the similar manner as Example 2. DNA was purified from the solutions obtained after bisulfite treatment with the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples.

#### 2. MSP

**[0072]** MSP was carried out with the following primer sets and DNA in the obtained analysis samples as a template. In this MSP, PCR products can be obtained when the template DNA contains a methylated CpG site. In this Example, the CpG sites to be analyzed by MSP are the CpG sites with "\*" in Figs. 1 to 3.

## EP 2 540 839 A1

### (i) Preparation of PCR reaction solution

**[0073]** A reaction solution (25 µl) was prepared by mixing the following reagents.

5	2 x FastStart Universal SYBR Green Master (ROX) (ROCHE)	12.5 µl
	F primer (10 µM)	1.0 µl
	R primer (10 µM)	1.0 µl
	Analysis sample (template)	1.0 µl
10	Distilled water	9.5 µl
	<hr/>	
	Total	25.0 µl

The sample which contains distilled water instead of the template and the sample which did not contain the template were used as negative controls.

### (ii) Primer sequences and reaction conditions

**[0074]** (Primer sequences for amplifying the base sequence SEQ ID NO: 15)

20 F: 5'- CGTGGGTTTTAGTTAGGAGC -3' (SEQ ID NO: 28)  
R: 5'- ATCCAAAACGACGATATTTAACG -3' (SEQ ID NO: 29)

**[0075]** This primer set analyses the 9th, 10th and 28th to 30th CpG sites from the 5' end of the base sequence SEQ ID NO: 15 (the same positions in the base sequence SEQ ID NO: 1).

(Reaction condition)

**[0076]**

30 1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 54°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0077]** (Primer sequences for amplifying the base sequence SEQ ID NO: 16)

35 F: 5'- CGTTTCGTTTCGATTTATAATAGACG -3' (SEQ ID NO: 30)  
R: 5'- AAAACCCATTCTTCTAACTCCG -3' (SEQ ID NO: 31)

This primer set analyses the 7th to 10th and 20th CpG sites from the 5' end of the base sequence SEQ ID NO: 16 (the 4th to 7th and 17th CpG sites from the 5' end of the base sequence SEQ ID NO: 2).

(Reaction condition)

**[0078]**

45 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0079]** (Primer sequences for amplifying the base sequence SEQ ID NO: 17)

50 F: 5'- TGAGGAATGTTAGTAGTAAGGTTACGT -3' (SEQ ID NO: 32)  
R: 5'- CCTATATACTCAAAAAAACCACGTC -3' (SEQ ID NO: 33)

55 This primer set analyses the 9th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 17 (the 6th and 9th CpG sites from the 5' end of the base sequence SEQ ID NO: 3).

## EP 2 540 839 A1

(Reaction condition)

### [0080]

5 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

[0081] (Primer sequences for amplifying the base sequence SEQ ID NO: 18)

10

F: 5'- GTTGGGTTATACGTCGTAGGGC -3' (SEQ ID NO: 34)  
R: 5'- GACAAACGAAAATAAACGTCGAA -3' (SEQ ID NO: 35)

15

This primer set analyses the 21 st to 23rd and 36th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 18 (the 1 st to 3rd from the 5' end of the base sequence SEQ ID NO: 4 and the 1st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5).

(Reaction condition)

20

### [0082]

25

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 57°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

[0083] (Primer sequences for amplifying the base sequence SEQ ID NO: 19)

30

F: 5'- GATAGTTTTAGAGTCGGGGAAGC -3' (SEQ ID NO: 36)  
R: 5'- CCTAATCCTAACAAAATCTACCGAC -3' (SEQ ID NO: 37)

This primer set analyses the 74th and 75th CpG sites from the 5' end of the base sequence SEQ ID NO: 19 (the 1st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6).

(Reaction condition)

35

### [0084]

40

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

45

[0085] After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels, and presence of PCR products was determined based on the bands in the gels. The gels were photographed and the photographs were analyzed with the Quantity One (Bio-Rad) software to quantitatively obtain fluorescence intensity of the bands. The results are shown in Figs. 5 to 9. In these Figures, "-" "PN" and "PT" mean a negative control without the template, normal pancreatic tissue, and pancreatic cancer tissue, respectively.

[0086] In MSP for the base sequences SEQ ID NOs: 15 and 16, almost no band was detected in normal tissues (pancreas, heart, kidney, liver, lung and peripheral blood leukocytes), while a band having significantly high intensity was detected for pancreatic cancer tissue (Figs. 5 and 6).

50

In MSP for the base sequence SEQ ID NO: 17, the intensity of the band detected was significantly in pancreatic cancer tissue than in normal pancreatic tissue (Fig. 7).

In MSP for the base sequences SEQ ID NOs: 18 and 19, almost no band was detected in normal tissues (pancreas, heart, kidney, liver and peripheral blood leukocytes), while a band having significantly high intensity was detected for pancreatic cancer tissue (Figs. 8 and 9).

55

[0087] As described above, in MSP for the base sequences SEQ ID NOs: 15 to 19, a higher amount of PCR products specific for pancreatic cancer tissue was obtained. This indicates that for genomic DNA from each tissue, MSP analysis on presence or absence of a methylated CpG in the base sequences SEQ ID NOs: 15 to 19, i.e. the base sequences SEQ ID NOs: 1, 2, 3, 4 and 5 and 6 allows differentiation between pancreatic cancer tissue and normal pancreatic tissue.

Accordingly, it is suggested that presence or absence of a cancer cell in a biological sample obtained from a subject can be determined by analyzing presence or absence of methylation of a CpG site located in the present molecular markers, the base sequences SEQ ID NOs: 1, 2, 3, 4, 5 and/or 6 by MSP for DNA extracted from a biological sample obtained from a subject.

5

Example 4: Analysis of methylation status in other tissues by MSP

**[0088]** It was studied whether or not a methylated CpG site can be detected for DNA obtained from cancer tissues other than pancreatic cancer tissue.

10

1. Bisulfite treatment

**[0089]** Genomic DNA (2 µg each) obtained from human breast cancer tissue, human normal mammary gland tissue, human colon cancer tissue and human normal large intestinal tissue was subjected to bisulfite treatment in the similar manner as Example 2. DNA was purified from the solutions obtained after bisulfite treatment with the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples.

15

2. MSP

**[0090]** MSP was carried out with the primer sets having the base sequences SEQ ID NOs: 28 and 29 and SEQ ID NOs: 34 and 35 as above and DNA in the obtained analysis samples as a template. MSP was carried out in the same manner as Example 3.

20

After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels in the same manner as Example 3, and fluorescence intensity of the bands appeared in the gels was quantitatively obtained. The results are shown in Figs. 10 and 11. In these Figures, "-" means a negative control without the template, "PN1" and "PN2" mean normal mammary gland tissues, "BT1" and "BT2" mean breast cancer tissues, "CN" means normal large intestinal tissue and "CT1" and "CT2" mean colon cancer tissues.

25

**[0091]** In MSP for the base sequence SEQ ID NO: 15 using the primer set having the base sequences SEQ ID NOs: 28 and 29, almost no band was detected in normal mammary gland tissue and normal large intestinal tissue, while bands having significantly high intensity were detected for breast cancer tissue and colon cancer tissue (Fig. 10).

30

In MSP for the base sequence SEQ ID NO: 18 using the primer set having the base sequences SEQ ID NOs: 34 and 35, the intensity of the band detected was significantly higher in breast cancer tissue than in normal mammary gland tissue (Fig. 11).

**[0092]** As described above, in MSP for the base sequences SEQ ID NOs: 15 and 18, a higher amount of PCR products specific for breast cancer tissue and colon cancer tissue was obtained. This indicates that for genomic DNA from each tissue, MSP analysis on presence or absence of a methylated CpG in the base sequences SEQ ID NOs: 15 and 18, i.e. the base sequences SEQ ID NOs: 1 and 4 and 5 allows differentiation between cancer tissue and normal tissue.

35

Accordingly, it is suggested that presence or absence of a cancer cell can be determined for tissues other than pancreas tissue by analyzing presence or absence of methylation of a CpG site located in the present molecular markers identified from pancreatic cancer tissue by MSP.

40

Example 5: Search for molecular markers for determination from breast cancer cell lines

**[0093]** By analyzing genomic DNA obtained from breast cancer cell lines and normal mammary gland epithelial cells with microarrays, gene regions were searched which were methylated specifically in the genomic DNA of the cell lines derived from breast cancer.

45

The specific procedures in Example 5 followed the instructions attached to the kits and reagents used.

1. Search for methylated genes by MeDIP on chip method

50

(1) Preparation of methylated DNA by MeDIP method

**[0094]** Genomic DNA was extracted and single-stranded DNA fragments were prepared in the similar manner as Example 1 except that the breast cancer cell lines MCF7, MB-MDA231 and SK-BR-3 and the normal mammary gland epithelial cell line HMEC were used as biological samples.

55

The obtained single-stranded DNA fragments were diluted in the similar manner as Example 1 followed by addition of Protein G Sepharose beads (68 µl: GE Healthcare). Pre-clear treatment was then carried out in the similar manner as Example 1. The supernatants were collected and then respectively divided into two aliquots in separate tubes, one of

## EP 2 540 839 A1

which was added with an anti-methylated cytosine antibody BI-MECY-0500 (10 µg: Eurogentec) and the other with a normal mouse IgG antibody (4 µg: Santa Cruz), serving as a test sample and a control sample, respectively. These test and control samples were subjected to immunoprecipitation in the similar manner as Example 1 to elute single-stranded DNA fragments.

The obtained DNA solutions were incubated with Proteinase K (Sigma) before purification with the Qiaquick PCR purification kit (QIAGEN).

Quantitative PCR allowed confirming that methylated DNA in the test samples was specifically concentrated.

### 2. Amplification and labeling of nucleic acid in samples

**[0095]** DNA in the test and control samples respectively obtained from MCF-7, MB-MDA231 and SK-BR-3 and HMEC cells was amplified on WT-Ovation™ Pico RNA Amplification System Version 1.0 (NuGEN). The concentration of the amplified nucleic acid was determined by measuring the absorbance (260 nm and 280 nm) of the samples.

The nucleic acid contained in the above amplified test and controls samples was fragmented and biotinylated on FL-Ovation™ cDNA Biotin Module V2 (NuGEN). The nucleic acid obtained from the samples served as the test and control samples for microarray analysis.

### 3. Microarray analysis

#### (1) Contact of sample and microarray

**[0096]** The samples for microarray analysis were brought into contact with the GeneChip® Human Promoter 1.0R Array (Affymetrix) to carry out hybridization with probes in the similar manner as Example 1. One same type of microarray was used for each of the test sample and the control sample.

#### (2) Analysis of microarray data

**[0097]** The resulting microarray data was analyzed according to the similar procedures as in Example 1. As a result, the regions represented by the base sequences SEQ ID NOs: 20 to 27 were identified as the gene regions which were specifically methylated in genomic DNA of breast cancer cell lines. The base sequences SEQ ID NOs: 20 to 27 are the gene regions or partial base sequence regions in the vicinity thereof of, respectively, *LBX2*, *PAX9*, *ADD3*, *CCDC61*, *ZIC4*, *CGB7*, *TOX* and *TNNI3* and the regions comprising the base sequences SEQ ID NOs: 7, 8, 9, 10, 11, 12, 13 and 14, respectively.

The correspondence relationship between the base sequences SEQ ID NOs: 20 to 27 and SEQ ID NOs: 7 to 14 and the above genes is summarized in the following Table 3.

**[0098]**

[Table 3]

SEQ ID NO:	Corresponding SEQ ID NO:	Gene
20	7	LBX2
21	8	PAX9
22	9	ADD3
23	10	CCDC61
24	11	ZIC4
25	12	CGB7
26	13	TOX
27	14	TNNI3

#### Example 6: Analysis of methylation status by bisulfite sequencing

**[0099]** By using bisulfite sequencing, it was investigated whether CpG sites in the gene regions represented by SEQ ID NOs: 20 to 27 identified in Example 5 were methylated in genomic DNA of a breast cancer cell line and a normal mammary gland epithelial cell line.

## EP 2 540 839 A1

### 1. Bisulfite treatment

5 **[0100]** Genomic DNA (2 µg) extracted from the breast cancer cell line MCF7 and the normal mammary gland epithelial cell line HMEC used in Example 5 was subjected to bisulfite treatment in the same manner as Example 2. DNA was purified from the solutions after bisulfite treatment using the Qiaquick PCR purification kit (QIAGEN) to prepare analysis samples.

### 2. Sequence analysis

10 **[0101]** PCR was carried out with the following primer sets and DNA in the obtained analysis samples as a template.

(i) Preparation of PCR reaction solution

15 **[0102]** A reaction solution (15 µl) was prepared by mixing the same reagents as Example 2.

(ii) Primer sequences and reaction conditions

**[0103]** (Primer sequences for amplifying the base sequence SEQ ID NO: 20)

20 F: 5'- GGAAGAGGTTTAAGTGGATTTTTTT -3' (SEQ ID NO: 60)  
R: 5'- TTTTCTTTCCAAACCCAACCTA -3' (SEQ ID NO: 61)

(Reaction condition)

25 **[0104]**

30 1 cycle of 95°C for 4.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds;  
1 cycle of 72°C for 7 minutes; and  
keep at 4°C.

**[0105]** (Primer sequences for amplifying the base sequence SEQ ID NO: 21)

35 F: 5'- AGTTAGGATTGTGTAATATTAGTTTT -3' (SEQ ID NO: 62)  
R: 5'- CACTATACAACCATCAACTACAAC -3' (SEQ ID NO: 63)

(Reaction condition)

40 **[0106]**

45 1 cycle of 95°C for 4.5 minutes;  
40 cycles of 95°C for 30 seconds, 55.5°C for 15 seconds and 72°C for 40 seconds;  
1 cycle of 72°C for 7 minutes; and  
keep at 4°C.

**[0107]** (Primer sequences for amplifying the base sequence SEQ ID NO: 22)

50 F: 5'- GTATATTTTTAGGGAGGAGGGGGG -3' (SEQ ID NO: 64)  
R: 5'- CAACCCCTACTTCACCTCCACATA -3' (SEQ ID NO: 65)

(Reaction condition)

**[0108]**

55 1 cycle of 95°C for 4.5 minutes;  
40 cycles of 95°C for 30 seconds, 65.1°C for 15 seconds and 72°C for 30 seconds;  
1 cycle of 72°C for 7 minutes; and  
keep at 4°C.

## EP 2 540 839 A1

**[0109]** (Primer sequences for amplifying the base sequence SEQ ID NO: 23)

F: 5'- TTATTAGGATGAGTTATTGGTTATTT -3' (SEQ ID NO: 66)

R: 5'- ACCTCCCTAACCCCAACTAC -3' (SEQ ID NO: 67)

5

(Reaction condition)

**[0110]**

10

1 cycle of 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 56.8°C for 15 seconds and 72°C for 40 seconds;

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

15

**[0111]** (Primer sequences for amplifying the base sequence SEQ ID NO: 24)

F: 5'- GTTTGGGTAGTTTATTGGTT -3' (SEQ ID NO: 68)

R: 5'- CTAAAAATTTCTCAACTCCTAACTC -3' (SEQ ID NO: 69)

20

(Reaction condition)

**[0112]**

25

1 cycle of 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 55.7°C for 15 seconds and 72°C for 30 seconds;

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

30

**[0113]** (Primer sequences for amplifying the base sequence SEQ ID NO: 25)

F: 5'- GAGGTGATATTAAGGATTTTTGGGT -3' (SEQ ID NO: 70)

R: 5'- TACAACTCAACTCCAATAACCCACAC -3' (SEQ ID NO: 71)

(Reaction condition)

35

**[0114]**

1 cycle of 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 61.5°C for 15 seconds and 72°C for 40 seconds;

40

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

**[0115]** (Primer sequences for amplifying the base sequence SEQ ID NO: 26)

45

F: 5'- AATAAGATTTGTTTAGTTTTATTGTTAAAA -3' (SEQ ID NO: 72)

R: 5'- TCTACCTAATACTCACAACCCCTACA -3' (SEQ ID NO: 73)

(Reaction condition)

50

**[0116]**

1 cycle at 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds;

1 cycle of 72°C for 7 minutes; and

55

keep at 4°C.

**[0117]** (Primer sequences for amplifying the base sequence SEQ ID NO: 27)

## EP 2 540 839 A1

F: 5'- GTAGTAGTAGAGTTTGTAGAGGGGTG -3' (SEQ ID NO: 74)

R: 5'- ATAAAAAATACCTATCCAAAAA -3' (SEQ ID NO: 75)

(Reaction condition)

### [0118]

1 cycle at 95°C for 4.5 minutes;

40 cycles of 95°C for 30 seconds, 50.9°C for 15 seconds and 72°C for 30 seconds;

1 cycle of 72°C for 7 minutes; and

keep at 4°C.

[0119] After the completion of PCR, plasmid constructs containing each of the PCR products were prepared in the similar manner as Example 2 and the constructs were sequenced.

After sequencing, methylation status of CpG sites in the base sequences of the above SEQ ID NOs was analyzed.

[0120] Methylation status of CpG sites located in the base sequences SEQ ID NOs: 7 to 14 which are respectively included in the base sequences SEQ ID NOs: 20 to 27 is shown in the tables in Figs. 12 to 19 based on the results of sequencing described above. In the tables in these Figures, "●", "○" and "-" represent a methylated CpG site, a non-methylated CpG site and an unanalyzable CpG site, respectively.

In the tables in these Figures, the numbers in the rows correspond to the numbers representing the position of the CpG sites from the 5' end of the respective base sequences SEQ ID NOs: 7 to 14. The CpG site with the symbol "\*" or "\*" on the number is the CpG site which tends to be methylated in the MCF7 cell samples compared to the HMEC cell samples.

[0121] Figs. 12 to 19 show that in all base sequences SEQ ID NOs: 7 to 14, CpG sites tend to be methylated in MCF7 cells but not in HMEC cells.

Based on these results, methylation frequency in each of the base sequences SEQ ID NOs: 7 to 14 was analyzed. Methylation frequency was calculated in the similar manner as Example 2.

The obtained methylation frequency is shown in the bar graph of Fig. 20. This graph shows the average values for the group of MCF7 cell samples and the group of HMEC cell samples.

[0122] Fig. 20 shows that, when all CpG sites located in the base sequences SEQ ID NOs: 7 to 14 are used as targets for analysis, a threshold for differentiating between the samples containing a cancer cell and those devoid of cancer cells can be selected from the range of 15 to 35%, for example.

Accordingly, it was suggested that, when methylation frequency of a CpG site in the base sequence(s) SEQ ID NO(s): 7, 8, 9, 10, 11, 12, 13 and/or 14 in DNA extracted from a biological sample obtained from a subject is higher than the above threshold, it can be determined that a cancer cell is present in the biological sample.

[0123] As the present molecular marker for determination, a CpG site which tends to be methylated in cancer cell lines but not in normal cell lines is suitable.

Accordingly, when presence or absence of a cancer cell in a sample is determined based on methylation frequency of a CpG site in the base sequences SEQ ID NOs: 7 to 14, more accurate determination can be made when the CpG site (s) as the molecular marker(s) for determination is selected from preferably the CpG site(s) with "\*" or "\*\*", more preferably the CpG site(s) with "\*" in Figs. 12 to 19.

When presence or absence of a cancer cell in a sample is determined based on the analysis result on presence or absence of methylation of a CpG site(s) in the base sequences SEQ ID NOs: 7 to 14 by methylation specific PCR, more accurate determination can be made when primers are used which can amplify preferably the CpG site(s) with "\*" or "\*\*", more preferably the CpG site(s) with "\*" in Figs. 12 to 19.

Example 7: Analysis of methylation status by MSP

[0124] Methylation status of CpG sites located in the gene regions of SEQ ID NOs: 20 to 27 identified in Example 5 was analyzed by MSP for genomic DNA of a breast cancer cell line and a normal mammary gland epithelial cell line.

### 1. Bisulfite treatment

[0125] Genomic DNA (2 μg each) obtained from the breast cancer cell line MCF7 and the normal mammary gland epithelial cell line HMEC was subjected to bisulfite treatment in the similar manner as Example 2. DNA was purified from the solutions obtained after bisulfite treatment with the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples.

## EP 2 540 839 A1

### 2. MSP

**[0126]** MSP was carried out with the following primer sets and DNA in the obtained analysis samples as a template. In this Example, the CpG sites analyzed by MSP are the CpG sites with "\*" in Figs. 12 to 19. The preparation of the PCR reaction solution was carried out in the similar manner as Example 3. Primer sequences and reaction conditions are as described hereinbelow.

(Primer sequences for amplifying the base sequence SEQ ID NO: 20)

#### **[0127]**

F: 5'- TTTTAGAGTTTAGGATTGGCGGC -3' (SEQ ID NO: 38)  
R: 5'- TACAACCTAACACTACCCGAAAACG -3' (SEQ ID NO: 39)

**[0128]** This primer set analyses the 3rd, 4th, 11th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 20 (the same positions in the base sequence SEQ ID NO: 7).

(Reaction condition)

#### **[0129]**

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0130]** (Primer sequences for amplifying the base sequence SEQ ID NO: 21)

F: 5'- ATAGGTTGGAAACGTAGTTTTTCG -3' (SEQ ID NO: 40)  
R: 5'- CTATAACGTCTAACGAATCCTCGC -3' (SEQ ID NO: 41)

This primer set analyses the 8th, 9th and 14th to 16th CpG sites from the 5' end of the base sequence SEQ ID NO: 21 (the same positions in the base sequence SEQ ID NO: 8).

(Reaction condition)

#### **[0131]**

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 30 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0132]** (Primer sequences for amplifying the base sequence SEQ ID NO: 22)

F: 5'- TGTTGTAAAGTTTGTTCGGTTTCGT -3' (SEQ ID NO: 42)  
R: 5'- TTCTACTTCATTTAAACCCCTCGAA -3' (SEQ ID NO: 43)

This primer set analyses the 15th, 16th and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 22 (the same positions in the base sequence SEQ ID NO: 9).

(Reaction condition)

#### **[0133]**

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0134]** (Primer sequences for amplifying the base sequence SEQ ID NO: 23)

## EP 2 540 839 A1

F: 5'- TGTGTGGAGTAGAATTTTGAGTAAATATGC -3' (SEQ ID NO: 44)

R: 5'- AATTTAAAAACAAAAACAACCGCA -3' (SEQ ID NO: 45)

This primer set analyses the 7th and 11th CpG sites from the 5' end of the base sequence SEQ ID NO: 23 (the 2nd and 6th CpG sites from the 5' end of the base sequence SEQ ID NO: 10).

(Reaction condition)

### [0135]

1 cycle of 95°C for 9.5 minutes;

35 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

[0136] (Primer sequences for amplifying the base sequence SEQ ID NO: 24)

F: 5'- GTAGTTTATTGGTTCGCGGTC -3' (SEQ ID NO: 46)

R: 5'- AAAAAAATATATAAAAAATAACGAT -3' (SEQ ID NO: 47)

This primer set analyses the 1st to 3rd and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 24 (the same positions in the base sequence SEQ ID NO: 11).

(Reaction condition)

### [0137]

1 cycle of 95°C for 9.5 minutes;

40 cycles of 95°C for 30 seconds, 51°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

[0138] (Primer sequences for amplifying the base sequence SEQ ID NO: 25)

F: 5'- AGAGTTCGGTTTATTTGGGATAGAATC -3' (SEQ ID NO: 48)

R: 5'- GACCGAAACGTCCTAAACCG -3' (SEQ ID NO: 49)

This primer set analyses the 10th, 11th and 16th to 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 25 (the same positions in the base sequence SEQ ID NO: 12).

(Reaction condition)

### [0139]

1 cycle of 95°C for 9.5 minutes;

35 cycles of 95°C for 30 seconds, 62°C for 30 seconds and 72°C for 30 seconds; and keep at 4°C.

[0140] (Primer sequences for amplifying the base sequence SEQ ID NO: 26)

F: 5'- TATTGTTTAAGATTCGGAGTTGCGA -3' (SEQ ID NO: 50)

R: 5'- CTCCCAACATTTACCTAATAACGAA -3' (SEQ ID NO: 51)

This primer set analyses the 9th, 10th and 17th CpG sites from the 5' end of the base sequence SEQ ID NO: 26 (the 7th, 8th and 15th CpG sites from the 5' end of the base sequence SEQ ID NO: 13).

(Reaction condition)

### [0141]

## EP 2 540 839 A1

1 cycle of 95°C for 9.5 minutes;  
30 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

5 **[0142]** (Primer sequences for amplifying the base sequence SEQ ID NO: 27)

F: 5'- GGGAGGGAAGCGTAGTTTATTC -3' (SEQ ID NO: 52)  
R: 5'- CTAAAAAATTTAAAAAACAAAAACGAT -3' (SEQ ID NO: 53)

10 This primer set analyses the 1st, 2nd and 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 27 (the same positions in the base sequence SEQ ID NO: 14).

(Reaction condition)

15 **[0143]**

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 54°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

20 **[0144]** After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels in the same manner as Example 3, and fluorescence intensity of the bands appeared in the gels was quantitatively obtained. The results are shown in Figs. 21 to 28. In these Figures, "-" means a negative control without the template. In MSP for each of the base sequences SEQ ID NOs: 20 to 27, bands having significantly higher intensity were detected for MCF7 cells than for HMEC cells (see Figs. 21 to 28).

25 **[0145]** As described above, in MSP for the base sequences SEQ ID NOs: 20 to 27, a higher amount of PCR products specific for breast cancer cells was obtained. This indicates that for genomic DNA from each cell line, MSP analysis on presence or absence of a methylated CpG in the base sequences SEQ ID NOs: 20 to 27, i.e. the base sequences SEQ ID NOs: 7 to 14 allows differentiation between breast cancer cells and normal mammary gland epithelial cells. Accordingly, it is suggested that presence or absence of a cancer cell can be determined by analyzing presence or absence of methylation of a CpG site located in the present molecular marker(s) for determination, i.e. the base sequence (s) SEQ ID NO(s): 7, 8, 9, 10, 11, 12, 13 and/or 14 by MSP for DNA extracted from a biological sample obtained from a subject.

35 Example 8: Analysis of methylation status of genomic DNA from breast cancer tissue and normal mammary gland tissue by MSP

**[0146]** By using MSP, it was investigated whether a methylated CpG site could be detected located in the gene regions represented by the base sequences SEQ ID NOs: 20 to 27 for not only a breast cancer cell line but also for breast cancer tissue.

### 1. Bisulfite treatment

45 **[0147]** Genomic DNA (2 µg each) obtained from human breast cancer tissue and human normal mammary gland tissue was subjected to bisulfite treatment in the same manner as Example 2. DNA was purified from the solutions after bisulfite treatment using the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples.

### 2. MSP

50 **[0148]** MSP was carried out with the primer sets having the base sequences SEQ ID NOs: 38 to 53 as above and DNA in the obtained analysis samples as a template. The PCR reaction solution was prepared in the similar manner as Example 3, and the reaction conditions of MSP using the above primer sets are the same as Example 7. After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels in the same manner as Example 3, and fluorescence intensity of the bands appeared in the gels was quantitatively obtained. The results are shown in Figs. 29 to 36. In these Figures, "-" means a negative control without the template.

55 **[0149]** In MSP for the base sequences SEQ ID NOs: 20 to 27 using the primer sets as above, almost no band was detected in normal mammary gland tissue, while bands having significantly high intensity were detected for breast cancer tissue (see Figs. 29 to 36).

## EP 2 540 839 A1

As described above, in MSP for the base sequences SEQ ID NOs: 20 to 27, a higher amount of PCR products specific for breast cancer tissue was obtained. This indicates that for genomic DNA from each tissue, MSP analysis on presence or absence of a methylated CpG in the base sequences SEQ ID NOs: 20 to 27 i.e., the base sequences SEQ ID NOs: 7 to 14 allows differentiation between breast cancer tissue and normal mammary gland tissue.

Accordingly, it is suggested that presence or absence of a cancer cell in tissue obtained from a subject can be determined by analyzing presence or absence of methylation of a CpG site located in the present molecular markers for determination.

Example 9: Analysis of methylation status of genomic DNA from colon cancer tissue and normal large intestinal tissue by MSP

**[0150]** Methylation status of CpG sites located in the gene regions represented by the base sequences SEQ ID NOs: 17 to 19, 21, 22 and 25 to 27 was studied for genomic DNA of colon cancer tissue and normal large intestinal tissue by MSP.

### 1. Bisulfite treatment

**[0151]** Genomic DNA of human colon cancer tissue and human normal large intestinal tissue (2 µg each: BioChain) was subjected to bisulfite treatment in the similar manner as Example 2. DNA was purified from the solutions obtained after bisulfite treatment with the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples.

### 2. MSP

**[0152]** MSP was carried out with the primer sets represented by the above SEQ ID NOs: 32 to 37, 40 to 43 and 48th to 53rd and DNA in the obtained analysis samples as a template. PCR reaction solutions were prepared as follows.

#### (i) Preparation of PCR reaction solution

**[0153]** A reaction solution (25 µl) was prepared by mixing the following reagents.

2 x FastStart Universal SYBR Green Master (ROX) (ROCHE)	12.5 µl
F primer (10 µM)	1.0 µl
R primer (10 µM)	1.0 µl
Analysis sample (template)	1.0 µl
Distilled water	9.5 µl
<hr/>	
Total	25.0 µl

**[0154]** The reaction conditions of MSP using the above respective primer sets are as follows.

(Reaction condition of MSP using the primer set (SEQ ID NOs: 32 and 33) amplifying the base sequence SEQ ID NO: 17)

#### **[0155]**

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0156]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 34 and 35) amplifying the base sequence SEQ ID NO: 18)

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 58°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0157]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 36 and 37) amplifying the base sequence SEQ ID NO: 19)

1 cycle of 95°C for 9.5 minutes;

## EP 2 540 839 A1

35 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

5 **[0158]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 40 and 41) amplifying the base sequence SEQ ID NO: 21)

1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

10 **[0159]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 42 and 43) amplifying the base sequence SEQ ID NO: 22)

15 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

20 **[0160]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 48 and 49) amplifying the base sequence SEQ ID NO: 25)

1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

25 **[0161]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 50 and 51) amplifying the base sequence SEQ ID NO: 26)

30 1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

**[0162]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 52 and 53) amplifying the base sequence SEQ ID NO: 27)

35 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 53°C for 15 seconds and 72°C for 30 seconds; and keep at 4°C.

40 **[0163]** After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels, and the presence or absence of the bands for the PCR products was determined. The result is shown in Fig. 37. In this Figure, "CN" and "CT" mean normal large intestinal tissue and colon cancer tissue, respectively.

**[0164]** Fig. 37 shows that in MSP for the base sequences SEQ ID NOs: 17 to 19, 21, 22 and 25 to 27 using the above primer sets, no band was detected for normal large intestinal tissue, and the bands were detected only for colon cancer tissue.

45 As described above, in MSP for the base sequences SEQ ID NOs: 17 to 19, 21, 22 and 25 to 27, PCR products specific for colon cancer tissue were obtained. This indicates that for genomic DNA from each tissue, MSP analysis on presence or absence of a methylated CpG in the base sequences SEQ ID NOs: 17 to 19, 21, 22 and 25 to 27, i.e. the base sequences SEQ ID NOs: 3 to 6, 8, 9 and 12 to 14 allows differentiation between colon cancer tissue and normal large intestinal tissue.

50 Accordingly, it is suggested that presence or absence of a cancer cell in tissue other than pancreas tissue and mammary gland tissue can be determined by analyzing presence or absence of methylation of a CpG site located in the present molecular markers identified from pancreatic cancer tissue and breast cancer cell lines by MSP.

55 Example 10: Analysis of methylation status of genomic DNA from various types of normal tissue by MSP

**[0165]** Methylation status of CpG sites located in the gene regions represented by the base sequences SEQ ID NOs: 20 to 27 was studied for genomic DNA from various types of normal tissue.

## EP 2 540 839 A1

### 1. Bisulfite treatment

5 [0166] Genomic DNA (2 µg each: BioChain) from human normal heart tissue, human normal kidney tissue, human normal liver tissue, human normal lung tissue and human normal peripheral blood leukocytes was subjected to bisulfite treatment in the similar manner as Example 2. DNA was purified from the solutions obtained after bisulfite treatment with the Qiaquick PCR purification kit (QIAGEN) to obtain analysis samples. Genomic DNA from the breast cancer cell line MCF7 was treated in the similar manner as described for the above normal tissues to obtain a positive control sample.

### 2. MSP

10 [0167] MSP was carried out with the primer sets represented by SEQ ID NOs: 38 to 53 as above and DNA in the obtained analysis samples and the positive control sample as a template. PCR reaction solutions were prepared in the similar manner as Example 9.

15 The reaction conditions of MSP using the above respective primer sets are as follows.

(Reaction condition of MSP using the primer set (SEQ ID NOs: 38 and 39) amplifying the base sequence SEQ ID NO: 20)

#### [0168]

20 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

25 [0169] (Reaction condition of MSP using the primer set (SEQ ID NOs: 40 and 41) amplifying the base sequence SEQ ID NO: 21)

30 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 62°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

[0170] (Reaction condition of MSP using the primer set (SEQ ID NOs: 42 and 43) amplifying the base sequence SEQ ID NO: 22)

35 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

40 [0171] (Reaction condition of MSP using the primer set (SEQ ID NOs: 44 and 45) amplifying the base sequence SEQ ID NO: 23)

1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

45 [0172] (Reaction condition of MSP using the primer set (SEQ ID NOs: 46 and 47) amplifying the base sequence SEQ ID NO: 24)

50 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 50°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

[0173] (Reaction condition of MSP using the primer set (SEQ ID NOs: 48 and 49) amplifying the base sequence SEQ ID NO: 25)

55 1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

## EP 2 540 839 A1

**[0174]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 50 and 51) amplifying the base sequence SEQ ID NO: 26)

5 1 cycle of 95°C for 9.5 minutes;  
35 cycles of 95°C for 30 seconds, 60°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

**[0175]** (Reaction condition of MSP using the primer set (SEQ ID NOs: 52 and 53) amplifying the base sequence SEQ ID NO: 27)

10 1 cycle of 95°C for 9.5 minutes;  
40 cycles of 95°C for 30 seconds, 53°C for 15 seconds and 72°C for 30 seconds; and  
keep at 4°C.

15 **[0176]** After completion of PCR, the reaction solutions were subjected to electrophoresis on 2% agarose gels, and the presence or absence of the PCR products was determined. The result is shown in Fig. 38. In this Figure, "MCF7" means the breast cancer cell line MCF7, "H" means the normal heart tissue, "K" means the normal kidney tissue, "Li" means the normal liver tissue, "Lu" means the normal lung tissue and "Phe" means the normal peripheral blood leukocytes.

20 **[0177]** Fig. 38 shows that in MSP for the base sequences SEQ ID NOs: 20 to 27 using the above primer sets, no band was detected for all normal tissues. The above MSP reaction conditions are the ones which sufficiently allow the detection of a band when the positive control sample prepared from genomic DNA of the MCF7 cells was used as a template.

As described above, in MSP for the base sequences SEQ ID NOs: 20 to 27, no PCR product was obtained for various types of normal tissue. This indicates that by MSP analysis of genomic DNA of the normal tissues, methylation of CpG sites located in the base sequences SEQ ID NOs: 20 to 27, i.e. SEQ ID NOs: 7 to 14 is not detected.

25 Accordingly, it is suggested that analysis of presence or absence of methylation of the CpG site(s) located in the present molecular marker for determination allows determination of absence of a cancer cell in various types of normal tissue.

30

35

40

45

50

55

EP 2 540 839 A1

SEQUENCE LISTING

<110> Sysmex Corporation

5 <120> Method for determining existence of cancer cells in specimen, markers and kit for detection

<130> TM5530PC

<150> JP2010-042814

10 <151> 2010-02-26

<160> 75

<170> PatentIn version 3.5

15 <210> 1

<211> 334

<212> DNA

<213> Homo sapiens

20 <400> 1

tcgggtggga gccggaatgg ggcctggatg gagagcaggc gccaccctgg gaacaaagct 60

ggttcagccc gtctgcagct gtcgctgccca gttggggcca catggcctct gaggaatgtc 120

agcagcaagg tcacgtgggc ctctgctgca cccaccgcg agctgggacg tggcttcctt 180

25 gagcatacag gcaccttgcc acctgggcac ccccttgggc gtccttagg cacctctgac 240

actgaccctg accgccctgg cagcacgggc cctcccagca gccttgccaa ggggccagct 300

ccagcaagtg ggcaggcctt gggctagagg tgct 334

30

<210> 2

<211> 548

<212> DNA

<213> Homo sapiens

35 <400> 2

cagtcagggga caaagtgtga gtgtcaagcg tgggacagtc accccttctg gccgacagcg 60

gttcaggttt aatgccataa ggccggctgg agggcaagcc cgcgaaggag agcgcaccgg 120

40 gcgtgggctc cagccaggag cgcagtacc tgccgtccgg cgccgccgcc gccacgggcg 180

cctgggggtg cacgtagggg tgggtggtgat ggtggtggta caccgcagcg ggtacagcgt 240

tggcgcccgc cgcgtgcaact gggttccacg aggcgccaaa caccgtcgcc ttggactgga 300

45 agctgcaagg gctgaagtgc ggggtgctcgg ccagcgtcgc cgctgcccgg ggaggctggc 360

ccaggggtccc cggcgcatag cggccaacgc tcagctcatc cgcggcgtcg gcgcccagca 420

ggaacgagtc cacgtagtag ttgccaggg ccccagtggt ggccatcacc gtgcccagcg 480

50 cctggcccgc ccggcccgcac ccacggaaat tatgaaactg cagatttcat gtaacaactt 540

ggtggcac 548

55 <210> 3

<211> 577

<212> DNA

EP 2 540 839 A1

<213> Homo sapiens

<400> 3  
 5 cagaatgaac tcctcggcat cgctttgaat ccgggctcac atatggttgc taattagctg 60  
 gctatctctt ggagttaaat gatttttttt ctcgccccgt tcgcatccat aatagacgcg 120  
 aacagagtcg tgattgttcc ctggtgcggg acattctgcg cctgaatgag cgggggctgc 180  
 10 tgtgtcgacg cggatttgcc atatggttgc gcggagccag gaagaatggg cccctgtgtg 240  
 tgcgaccggg acgcagcggc gcgggccgtg ccagccgggc ggagccgggt gcccgcggc 300  
 gcaggacgag cgggcgcagg acgagcgcgc gcagcggggc cgggcgcacc gtcggggcgc 360  
 15 ggggaccccg agggccagcg gagccgtgct gcgcgcaggg atggacgctg ggcgccaag 420  
 gcccgttccc gtggacaggc gcgcggaggg cagcaggacc gatgcaccga cggacgcacc 480  
 gacggacgcg cggacagatg cgccgaccgc aggctctccc gccgtcccgg ggccgtcgtg 540  
 20 cccgtgcagg ccgcgagcag cggcggggcc ggctctg 577

<210> 4  
 <211> 38  
 <212> DNA  
 25 <213> Homo sapiens

<400> 4  
 catgccaac ctgctgggcc acacgtcgca gggcgagg 38

30 <210> 5  
 <211> 40  
 <212> DNA  
 <213> Homo sapiens

35 <400> 5  
 accaggtctc gacgccatt cccgcctgcc ggcccatgtg 40

40 <210> 6  
 <211> 55  
 <212> DNA  
 <213> Homo sapiens

<400> 6  
 gatcaccagg cagcacctgg acagctccag agtcggggaa gcgccatggt tcttg 55

45 <210> 7  
 <211> 508  
 <212> DNA  
 <213> Homo sapiens

50 <400> 7  
 gtggtggtgg ggtcagtgga gaggaggggt tgttcagaga gggctctcta aaggaagag 60  
 aggtggaaaa tatgtgtgcg ggtggagtgg gggttggggg tggggttaca agcaatttg 120  
 55 gagggaacag aatccaggtg aaatatggga gtgatggaat acaggctgaa agtgggatgt 180  
 gtggagcaga accttgagca aatatgcccg aagaatgcag ttggaataac ggggtgcatt 240

EP 2 540 839 A1

gaagttgaat atgcgggagc cctaggagac cgtgggtggc tgggttaggg aggagagtgc 300  
 ggctgcctcc ctgcccccaa atctgggaga gcaggatgga atgaaggggc tttggacggg 360  
 5 agagaggggg ctggtttccg gggaatgtag catgaggggtt gggctggaat atgtggagtg 420  
 ggaggttctg cagagaggag gggaagggga gtccagaatt ccgtggccat ctgggggaca 480  
 agtgacgggc agctggggtc agggaggc 508  
 10  
 <210> 8  
 <211> 508  
 <212> DNA  
 <213> Homo sapiens  
 15  
 <400> 8  
 gcctgggtag ccattggcc cgcggccggg tccgacgtca gtgatgacag gtcctgccgc 60  
 acgcggtcgg tctctctctc tctctctcac acttccctgc tccaaccgcc accccctcac 120  
 20 acacctctcc cccaaacata taccctcttt ccctccctca cccaccccc acagcatctc 180  
 aagcttccac tgctcatgcg ctctgaccg ggcacaaatc cttttctctc cccttgagcc 240  
 tcctccccct cctctccgga aagctgagt gatcagcttt gactcagggc tcctggaagt 300  
 25 actcggggct gggttatgcc accattcacc acaactttgg ggaggcagca ttgcgggacc 360  
 tcagtccggg ctgatgggca gcatcgtgtg cgcagagatg gcggtctcgt tgggcagttt 420  
 gtcctgcag cctgcaagag gagggggtcg ggatcgaagg accgaaggag gttgctgact 480  
 30 caggagccag gagctgagaa actcctag 508  
 35  
 <210> 9  
 <211> 655  
 <212> DNA  
 <213> Homo sapiens  
 40  
 <400> 9  
 gagtgacac taaggactcc tgggtgctaa aagagggagg ggctggggtc tgaactgtgg 60  
 ttgttgctgc tataggagtc gagactgaga ccaccgggtca gggaatggga gccagcccag 120  
 gggctcggcg tgcaccggcc atggcggggg ctgtgcagga ggcggtgcc agcgagagcc 180  
 45 cggcttactt gggatagaac cgaagtccc gcgccgaggt ggtcagatag agctggcggc 240  
 ggttcaggac gccccggctg gatactgtga aggggtgggccc agacagcgcg gggttcttcg 300  
 tgcaggggat tagagcctga gcaagattgg gggaggaggc atacccaaga catagatttc 360  
 ataccatc ccaagtctg ggtcccagat tcagaatccc tctccttcgg cctcgacagc 420  
 50 catgacggcc acggccctca cccccgcc cagggtcaga gcccttctg ccaccacgc 480  
 caggaactt cagcttctt tctcataccc gaacccttc cagcctcacc tcctaaatc 540  
 caagccctcc tgctggtaaa ggaactcaaa tgcaggcccc cagccaccac aaaatcccc 600  
 55 tggttctgcc cccgggtctca agccttctg gtgtggccac tggagctgag ctgca 655

EP 2 540 839 A1

<210> 10  
 <211> 692  
 <212> DNA  
 <213> Homo sapiens

5  
 <400> 10  
 agctaggact gtgcaacacc agtctccgct ctccggagtcg gcacgggggc tgggtgggtgg 60  
 cgggatgaac tcggggaagc cgacaggctg gaaacgcagc tcctcgactt tgcagcgcac 120  
 10 caaccccccta tcgccccgtc cgctgctcag gtgcgaggat tcgccagacg ccacaggact 180  
 gggggcgact cgcaagggca tcatgccacc tagaggaaac gcgcagtgag tgagccctgg 240  
 gtcactgaaa atgggagaac gttcagagtcg gtcaagatga ctccggcctg cgagttgatt 300  
 15 cagactcctc ggaagccaga aattgggtct agagtgggtt ggtccctccg tcagcagttc 360  
 gggcccagcc gaagcttggg gcctagaggg cagacgtccg gctctgatca gagtacagcg 420  
 ggctgggcct ggcagcctgc aggcaacctg gaccttgcaa taaagctgca ttcaacataa 480  
 20 ctcagaccac agtgggagcc cccgaagggc ttcattgtgag tcccgaacc agagtttaga 540  
 ggcgcggcct ggttctcgtt tatcttttcc agacagcatt taggccaata aggtccttga 600  
 aaggaggcag ccccaaagcc ctcaaggcca cacctgaggg cagggtctca aacagctgtg 660  
 25 cttgaagaac cggcaatctt gcttttccca ca 692

<210> 11  
 <211> 530  
 <212> DNA  
 <213> Homo sapiens

30  
 <400> 11  
 taacacaata cttgagtgta cagacaaaac agatttttaga ctaaaagaac tatgtcacca 60  
 35 gcaaataatg aaacttcatac acacaaactc ctatgcgctt gccaaagttc cccgtacctt 120  
 tcaacttgcc cgtagctgaa gggctccctg tacaatggaa gcccttctgc cttgctcaag 180  
 actggatgtc cctcttttct aacttgccata aacaccatcg gcacacagca cagtcctttg 240  
 40 agtgggtctc aacttgaca cagcatcgag gagcttctcc acggggaagt gtgtgtgtga 300  
 gagcctgcct gtgtgtcaac actgtccaag actcggagct gcgactcggc cgcgtggtgc 360  
 ggagtccagg gctgagaagc cagcgacta aatatctctg tatcaggtaa aaaagccgctt 420  
 45 tggatgatgac tcgccaccag gcaaatgttg ggagtcccaa cgatttttcc cgtggaatgc 480  
 accgagggtc gccatggatg tgccctgcagg ggctgtgagt atcaggcaga 530

<210> 12  
 <211> 586  
 <212> DNA  
 <213> Homo sapiens

50  
 <400> 12  
 55 gcacatTTTT agggaggagg ggggggatga agatgaggcg gatggcgtga ggtatttgag 60  
 ggggacgagt ctgcagatga ggaaaggaaa tgaagtggat tgcgccagc gccttgaaa 120

EP 2 540 839 A1

cgaaggggga taaggggtggg gaacgaagtt tgttgtgtgt gttaggggtgg ggttgGCCCA 180  
 agggaccatg aatcattaga atggcgctcg gaggggtgtg cttgcaagtg tactggtgca 240  
 5 ggagggacgt ggaagtcgtg ggtggcagag aaaagtgtgc accgctggga ggtctatttc 300  
 aagaataggt tgggatgaga aacatgtctg ggcgaggggg gtgggaggcc cgggtgcacc 360  
 tgtggatgga tccacagccc tgcaattgcg tgttgcaaag tttgcccgcc cccgcattca 420  
 10 ttttgaaggt tcttggcgat gcctcgactt ttcgcctcaa agaagtgagc tctcataagg 480  
 ttcagaggcc gggagccaaa gcgggagtga agactctcga ggggtccaaa tgaagcagaa 540  
 gcgagagccc ctgtcctccc cctatgtgga ggtgaagcag gggttg 586  
 15  
 <210> 13  
 <211> 436  
 <212> DNA  
 <213> Homo sapiens  
 20  
 <400> 13  
 gcagcagcag agtctgcaga ggggtgggag ggaagcgcag cccaccgagg gcttcaggat 60  
 aaagaccagg cgtggggaac cgcctctgcc cttctaaacc ctccagtttg gtctccactg 120  
 25 ttccaaggcc ccgtcccacc ccgagcagta ctccccgcta aagccacgcc ccgagcggcc 180  
 aaaccccgcc cacttccgcc cacctacccc gaaagcccca cccattctca agctccgcc 240  
 cctgagcacc tgctgtctct ttcccagtcg cgcccgtcct caccttcagc tgcaattttc 300  
 30 tcgaggcgga gatcttagat tttttctgcc aggggtgagat ggagcaagga aggatcatgg 360  
 agggggattc ggagacgacg gtggagggga cctcaagaca cccccagcaa acccagccgg 420  
 tccagatttg ggccca 436  
 35  
 <210> 14  
 <211> 567  
 <212> DNA  
 <213> Homo sapiens  
 40  
 <400> 14  
 ggaagaggcc caagtggacc tccttcctcc accctggaac tctggccaga ggagagcgc 60  
 gaggtgagga gtccagggcc ccagagccca ggattggcgg cggtttgtc ttcaatcgtc 120  
 45 cacctgtatc tcctcgtctg acaggtgggg ccgggagtca gggccggcag ggccgaggca 180  
 gaggccggga tctggagcgc cttcgggcag tgctaagctg cacaggactt ccggggacaa 240  
 cgcgcgtagc gaggcgacgt cggcgcgcat ctctccaca tcgcgcttga gcttggtctg 300  
 50 ccggttctgg aaccaagtga ccacctgcgc gttggccagg ccgagtcgcg tagctagccc 360  
 gtctcgtctg gacggcgcca ggtacttctg gaagacgaag cgccgctcca gctccagcac 420  
 ctggtgcgcg gtgaacgcag tgcgtgactt gcgcccgtttg cggccgaagg gaccagggcc 480  
 55 cagcgcgtcc ggacctgcc gcctgtagg gatagggagg gggtcagttt ccagcctccg 540

EP 2 540 839 A1

ggaatcaggc tgggcctgga aagaaaa 567

5  
 <210> 15  
 <211> 673  
 <212> DNA  
 <213> Homo sapiens

<400> 15  
 10  
 aggcactcat ggggcaggaa tcacagtgcc agctgcccc cgaaacgtgg gaacaaaagc 60  
 agcgtcgggt gggagccgga atggggcctg gatggagagc aggcgccacc ctgggaacaa 120  
 agctggttca gcccgctctgc agctgtcgct gccagttggg gccacatggc ctctgaggaa 180  
 15  
 tgtcagcagc aaggtcacgt gggcctctgc tgcaccacc cgcgagctgg gacgtggctt 240  
 ccctgagcat acaggcacct tgccacctgg gcaccccctt gggcgtccct taggcacctc 300  
 tgacactgac cctgaccgcc ctggcagcac gggccctccc agcagccttg ccaaggggcc 360  
 20  
 agctccagca agtgggcagg ccttgggcta gaggtgctcg ggggaaggag agcagagctg 420  
 gacctgcagg tccggggaag ccagggtgagc gcccagccct gcttcttccc catcaagaca 480  
 cccctgaccg cagcatcccc tggctctgtt cctgaaattc actgagcacc gtgggtcact 540  
 25  
 catccaggcc caggctgaca ccattcacca actctcacct tagagagtct gggaaagacc 600  
 tccagctggt gagaggcctg agctctagcc tggcttcctt cgaactctca gtgggaccct 660  
 agccaagtga cct 673

30  
 <210> 16  
 <211> 579  
 <212> DNA  
 <213> Homo sapiens

<400> 16  
 35  
 cagtccagga caaagtgtga gtgtcaagcg tgggacagtc accccttctg gccgacagcg 60  
 gttcaggttt aatgccataa ggccggctgg agggcaagcc cgcgaaggag agcgcaccgg 120  
 40  
 gcgtgggctc cagccaggag cgcattgtacc tgccgtccgg cgcgcccgcc gccacgggcg 180  
 cctgggggtg cacgtagggg tgggtggtgat ggtggtggta caccgcagcg ggtacagcgt 240  
 tggcggccgc cgcgtgcact gggttccacg aggcgccaaa caccgtcgcc ttggactgga 300  
 45  
 agctgcacgg gctgaagtcg gggtgctcgg ccagcgtcgc cgcctgccgg ggaggctggc 360  
 ccagggtccc cggcgcatag cggccaacgc tcagctcatc cgcggcgtcg gcgccagca 420  
 ggaacgagtc cacgtagtag ttgcccaggg ccccagtggt ggccatcacc gtgccagcg 480  
 50  
 cctgcccgc cggcccgcac ccacggaaat tatgaaactg cagatttcat gtaacaactt 540  
 ggtggcaccg ggggggaagt acagtcacct aataagttg 579

55  
 <210> 17  
 <211> 642  
 <212> DNA  
 <213> Homo sapiens

EP 2 540 839 A1

<400> 17  
 gggacatctg tcttctatac ctaataaagt cgcccgacgc agaatgaact cctcggcatc 60  
 gctttgaatc cgggctcaca tatggttgct aattagctgg ctatctcttg gagttaaagt 120  
 5 attttttttc tcgccccgtt cgcatccata atagacgcga acagagtcgt gattgttccc 180  
 tgttgcggga cattctgcgc ctgaatgagc gggggctgct gtgtcgacgc ggatttgcca 240  
 tatggttgcg cggagccagg aagaatgggc ccctgtgtgt gcgaccggga cgcagcggcg 300  
 10 cgggccgtgc cagccgggcg gagccgggtg cccgacggcg caggacgagc gggcgcagga 360  
 cgagcgcgagc cagcgggccc gggcgcaccg tcggggcgcg gggaccccga gggccagcgg 420  
 agccgtgctg cgcgcagggg tggacgctgg gcgcccgaagg cccgttcccg tggacagggc 480  
 15 cgcggagggc agcaggaccg atgcaccgac ggacgcaccg acggacgcgc ggacagatgc 540  
 gccgaccgca ggctctcccg ccgtcccggg gccgtcgtgc ccgtgcaggc cgcgagcagc 600  
 20 ggcggggccg gctctgcggc tgcagcagcc ccattgtgag gc 642

<210> 18  
 <211> 492  
 <212> DNA  
 <213> Homo sapiens

<400> 18  
 gctgctgctg tggcagctgc tggcggcggg cggcgcggca ctggagatcg gccgcttcga 60  
 30 cccggagcgc gggcgcgggg ctgcgccgtg ccaggcgggtg gagatcccca tgtgccgcgg 120  
 catcggctac aacctgaccg gcatgcccaa cctgctgggc cacacgtcgc agggcgaggc 180  
 ggctgccgag ctagcggagt tcgcgccgtt ggtgcagtac ggctgccaca gccacctgcg 240  
 35 cttcttctctg tgctcgctct acgcgcccat gtgcaccgac caggtctcga cgcccattcc 300  
 cgcctgccgg cccatgtgcg agcaggcgcg cctgcgctgc gcgcccatca tggagcagtt 360  
 caacttcggc tggccggact cgctcgactg cgcccggctg cccacgcgca acgaccgca 420  
 40 cgcgctgtgc atggaggcgc ccgagaacgc cacggccggc cccgcgaggc cccacaaggg 480  
 cctgggcatg ct 492

<210> 19  
 <211> 654  
 <212> DNA  
 <213> Homo sapiens

<400> 19  
 tggagtgggg ccccaacttta ccttgcccac atacaggggc tccgtgcccg tgtactcctc 60  
 taccacgaag aactggttcc acaccagcc gcgtttgacg cggccggctc ccagcgcgcc 120  
 55 gtccctgccga gctccgggcg ccgacggcga ggggtgtgcc gctgcccaaca ggcgccccag 180  
 cagcgtcggc ggcggcggca gcagcagcag cagcagtagc gcgggggaca gcgcgactcc 240

EP 2 540 839 A1

cgccccggagc cccctacctt cgggcctcgg cctcatcctt ggccctgcgcg gggctggggc 300  
ccaggagcat ggacgagagg caccagggcg ccgctgcttg gtcgcacaac gatgcggcgc 360  
5 cgtgtcacat ggtggcctca gcgcggcccg cgggatgtcg cccccgacgg ggcacccgga 420  
cgggcccgcg gccctgaacg ccgcccagcc gcgggggtac ccgctggag ggggagggg 480  
cgcgcccgca tccggggcat ggacagcccc cggggtgccc gcccgcccc ccgtcgccgc 540  
10 gtcgcgtgcg gatcaccagg cagcacctgg acagctccag agtcggggaa gcgccatggt 600  
tcctgcgcag aaaggatgcg ggttggggcc ggcagatcct gccaggacta gggg 654

<210> 20  
<211> 657  
<212> DNA  
<213> Homo sapiens

<400> 20  
20 ccaccaggat gagccactgg tcattccgaa gcaaggatgg aggatgcagg tagagggaca 60  
gggtcagcgg cgggagggga gagggggctc aggggtccggg gaatcctaga agtggcggtg 120  
gtggtgggggt cagtggagag gaggggttgt tcagagaggg ctctctaaag ggaagagagg 180  
25 tggaaaatat gtgtgcgggt ggagtggggg ttgggggtgg ggttacaagc aatttgggag 240  
ggaacagaat ccaggtgaaa tatgggagtg atggaataca ggctgaaagt gggatgtgtg 300  
gagcagaacc ttgagcaaat atgcgggaag aatgcagttg gaataacggg gtgcattgaa 360  
30 gttgaatatg cgggagccct aggagaccgt ggggtggctgg ttagggagg agagtgcggc 420  
tgcctccctg cccccaaatc tgggagagca ggatggaatg aaggggcttt ggacgggaga 480  
gagggggctg gtttccgggg aatgtagcat gagggttggg ctggaatatg tggagtggga 540  
35 ggttctgcag agaggagggg aaggggagtc cagaattccg tggccatctg ggggacaagt 600  
gacgggcagc tggggtcagg gaggccgtgt ggagctggca cgggtgtggg atttatt 657

<210> 21  
40 <211> 508  
<212> DNA  
<213> Homo sapiens

<400> 21  
45 gcctgggtag cccattggcc cgcggccggg tccgacgtca gtgatgacag gtcctgccgc 60  
acgcggtcgg tctctctctc tctctctcac acttccctgc tccaaccgcc acccctcac 120  
acacctctcc cccaaacata taccctcttt ccctccctca ccccacccc acagcatctc 180  
50 aagcttccac tgctcatgcg ctctgaccgg ggcacaaatc cttttctctc cccttgagcc 240  
tcctccccct cctctccgga aagctgagtg gatcagcttt gactcagggc tcctggaagt 300  
actcggggct gggttatgcc accattcacc acaactttgg ggaggcagca ttgcgggacc 360  
55 tcagtcgggg ctgatgggca gcatcgtgtg gcgagagatg gcggtctcgt tgggcagttt 420  
gtccctgcag cctgcaagag gagggggtcg ggatcgaagg accgaaggag gttgctgact 480

EP 2 540 839 A1

caggagccag gagctgagaa actcctag 508

5 <210> 22  
 <211> 655  
 <212> DNA  
 <213> Homo sapiens

<400> 22  
 10 gaggtgacac taaggactcc tgggtgctaa aagagggagg ggctggggtc tgaactgtgg 60  
 ttgttgctgc tataggagtc gagactgaga ccaccgggtca gggaatggga gccagcccag 120  
 gggctcggcg tgcaccggcc atggcggggg ctgtgcagga ggcgggtgcc agcgagagcc 180  
 15 cggcttactt gggatagaac cgaaagtccc gcgccgaggt ggtcagatag agctggcggc 240  
 ggttcaggac gccccggctg gatactgtga aggggtgggccc agacagcgcg gggttcttcg 300  
 tgcaggcgat tagagcctga gcaagattgg gggaggaggc atacccaaga catagatttc 360  
 20 ataccatc ccaagtctctg ggtcccaggt tcagaatccc tctccttcgg cctcgcacagc 420  
 catgacggcc acggccctca caccocgccc cagggtcaga gcccttcctg ccaccacgc 480  
 cagggaaactt cagcttcctt tctcataccc gaacccttcc cagcctcacc tccctaaatc 540  
 25 caagccctcc tgctggtcaa ggaactcaaa tgcaggcccc cagccaccac aaaatcccc 600  
 tggttctgcc ccgggtctca agccttcttg gtgtggccac tggagctgag ctgca 655

30 <210> 23  
 <211> 738  
 <212> DNA  
 <213> Homo sapiens

<400> 23  
 35 agctaggact gtgcaacacc agtctccgct ctccggagtcg gcacgggggc tgggtgggtgg 60  
 cgggatgaac tccgggaagg cgacaggctg gaaacgcagc tcctcgactt tgcagcgcac 120  
 caacccccta tgcgccgctc cgctgctcag gtgcgaggat tgcgcagacg ccacaggact 180  
 40 gggggcgact cgcaagggca tcatgccacc tagaggaaac gcgcagtgag tgagccctgg 240  
 gtcactgaaa atgggagaac gttcagtcg gtcaagatga ctccggcctg cgagttgatt 300  
 cagactcctc ggaagccaga aattgggtct agagtgggtt ggtccctccg tcagcagttc 360  
 45 gggcccagcc gaagcttggg gcctagaggg cagacgtccg gctctgatca gactacagcg 420  
 ggctgggcct ggcagcctgc aggcaacctg gaccttgcaa taaagctgca ttcaacataa 480  
 ctcagaccac agtgggagcc cccgaaggc ttcatgtgag tcccgaacc agagtttaga 540  
 50 ggcgcggcct ggttctcgtt tatctttcc agacagcatt taggccaata aggtctttga 600  
 aaggaggcag ccccaaagcc ctcaggccga cacctgaggg cagggctctca aacagctgtg 660  
 55 cttgaagaac cggcaatctt gcttttcca cacgcagctg cacctctcct gagcgtgca 720  
 gctgatggct gtatagtg 738

EP 2 540 839 A1

<210> 24  
 <211> 680  
 <212> DNA  
 <213> Homo sapiens

5  
 <400> 24  
 aataagattt gtccagtctc attgccaaaa tagttctgag cttttggttc cggttaactac 60  
 agtgaataa gtctttactc tccaagatag ggctgcaact ttattatfff ttattagcaa 120  
 10 tagtattggg gtcataagca aaataaaacg taacacaata cttgagtgtg cagacaaaac 180  
 agattttaga ctaaaagaac tatgtcacca gcaataatg aaacttcatc acacaaactc 240  
 ctatgcgctt gccaaagttc cccgtacctt tcaacttgcc cgtagctgaa gggctccctg 300  
 15 tacaatggaa gcccttctgc cttgctcaag actggatgtc cctcttttct aacttgctta 360  
 aacaccatcg gcacacagca cagtcctttg agtgggtctc aactggaca cagcatcgag 420  
 gagcttctcc acggggaagt gtgtgtgtga gagcctgcct gtgtgtcaac actgtccaag 480  
 20 actcggagct gcgactcggg cgcgtggtgc ggagtccagg gctgagaagg cagcgcacta 540  
 aatatctctg tatcaggtaa aaaagccgtt tggatgatgac tcgccaccag gcaaattgtg 600  
 ggagtcccaa cgatttttcc cgtggaatgc accgagggtc gccatggatg tgctgcagg 660  
 25 ggctgtgagt atcaggcaga 680

<210> 25  
 <211> 586  
 <212> DNA  
 <213> Homo sapiens

30  
 <400> 25  
 gcacattttt agggaggagg ggggggatga agatgaggcg gatggcgtga ggtatttgag 60  
 35 ggggacgagt ctgcagatga ggaaaggaaa tgaagtggat tgtcgccagt gccttgghaa 120  
 cgaaggggga taagggtggg gaacgaagtt tgttgtgtgt gttagggtgg ggttgghcca 180  
 agggaccatg aatcattaga atggcgctcg gaggggtgtg cttgcaagtg tactgggtgca 240  
 40 ggagggacgt ggaagtcgtg ggtggcagag aaaagtgtgc accgctggga ggtctatttc 300  
 aagaataggt tgggatgaga aacatgtctg ggcgaggggg gtgggaggcc cgggtgcacc 360  
 tgtggatgga tccacagccc tgcaattgag tgttgcaaag tttgcccggc ccgcattca 420  
 45 ttttgaaggt tcttggcgat gcctcgactt ttcgcctcaa agaagtgagc tctcataagg 480  
 ttcagaggcc gggagccaaa gcgggagtga agactctcga ggggtccaaa tgaagcagaa 540  
 50 gcgagagccc ctgtcctccc cctatgtgga ggtgaagcag gggttg 586

<210> 26  
 <211> 527  
 <212> DNA  
 <213> Homo sapiens

<400> 26

EP 2 540 839 A1

gcagcagcag agtctgcaga ggggtgggag ggaagcgcag cccaccggg gcttcaggat 60  
aaagaccagg cgtggggaac cgcctctgcc cttctaaacc ctccagtttg gtctccactg 120  
5 ttccaaggcc ccgtcccacc ccgagcagta ctccccgcta aagccacgcc ccgagcggcc 180  
aaaccccgcc cacttccgcc cacctacccc gaaagcccca ccatttctca agctccgccc 240  
cctgagcacc tgcctgtctt ttcccagtc cgcctgtcct caccttcagc tgcaattttc 300  
10 tcgaggcggg gatccttagat tttttctgcc aggggtgagat ggagcaagga aggatcatgg 360  
agggggattc ggagacgacg gtggagggga cctcaagaca ccccagcaa acccagccgg 420  
tccagatttg ggcccacgtc caacttgagc cctgagtcta cgggaggcca cggccctcat 480  
15 tcccatccac caatctgggt ctcttccttt ggataggcac ttcccat 527

<210> 27  
<211> 567  
20 <212> DNA  
<213> Homo sapiens

<400> 27  
ggaagaggcc caagtggacc tccttctccc accctggaac tctggccaga ggcagagcgc 60  
25 gaggtgagga gtccagggcc ccagagccca ggattggcgg cggctttgtc ttcaatcgtc 120  
cacctgtatc tcctcgtctg acaggtgggg ccgggagtca gggccggcag ggccgaggca 180  
gaggccggga tctggagcgc cttcggggcag tgctaagctg cacaggactt ccggggacaa 240  
30 cgcgcgtagc gaggcgacgt cggcgcgcat ctctccaca tcgcgcttga gcttggctcg 300  
ccggttctgg aaccaagtga ccacctgcgc gttggccagg ccgagtcgcg tagctagccc 360  
gtctcgctcg gacggcgcca ggtacttctg gaagacgaag cgcctctcca gctccagcac 420  
35 ctgttgcgcg gtgaacgcag tgcgtgactt gcgccgtttg cggccgaagg gaccagggcc 480  
cagcgcgtcc ggacctgccc gccctgtagg gatagggagg gggtcagttt ccagcctccg 540  
ggaatcaggc tgggcctgga aagaaaa 567

40 <210> 28  
<211> 27  
<212> DNA  
<213> Artificial Sequence

45 <220>  
<223> primer

<400> 28  
tgaggaatgt tagtagtaag gttacgt 27

50 <210> 29  
<211> 25  
<212> DNA  
<213> Artificial Sequence

55 <220>  
<223> primer

	<400> 29		
	cctatatact caaaaaaacc acgtc		25
5	<210> 30		
	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> primer		
	<400> 30		
	cgtggggtttt agttaggagc		20
15	<210> 31		
	<211> 23		
	<212> DNA		
	<213> Artificial Sequence		
20	<220>		
	<223> primer		
	<400> 31		
	atccaaaacg acgatattta acg		23
25	<210> 32		
	<211> 26		
	<212> DNA		
	<213> Artificial Sequence		
30	<220>		
	<223> primer		
	<400> 32		
35	cgtttcgttc gtatttataa tagacg		26
	<210> 33		
	<211> 23		
	<212> DNA		
40	<213> Artificial Sequence		
	<220>		
	<223> primer		
45	<400> 33		
	aaaaccatt cttcctaact ccg		23
	<210> 34		
	<211> 22		
50	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> primer		
55	<400> 34		
	gttgggttat acgtcgtagg gc		22

<210> 35  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 5  
 <220>  
 <223> primer  
 <400> 35  
 gacaaacgaa aataaacgtc gaa 23  
 10  
 <210> 36  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 15  
 <220>  
 <223> primer  
 <400> 36  
 gatagtttta gagtcgggga agc 23  
 20  
 <210> 37  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 25  
 <220>  
 <223> primer  
 <400> 37  
 cctaataccta acaaaatcta ccgac 25  
 30  
 <210> 38  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
 35  
 <220>  
 <223> primer  
 <400> 38  
 tgtgtggagt agaatdddga gtaaatatgc 30  
 40  
 <210> 39  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 45  
 <220>  
 <223> primer  
 <400> 39  
 aatttaaaaa caaaaaaaca accgca 26  
 50  
 <210> 40  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 55

<220>  
 <223> primer  
  
 <400> 40  
 5 gtagtttatt ggttcgcggt c 21

<210> 41  
 <211> 27  
 <212> DNA  
 10 <213> Artificial Sequence

<220>  
 <223> primer  
  
 <400> 41  
 15 aaaaaaata tataaaaaa taacgat 27

<210> 42  
 <211> 27  
 <212> DNA  
 20 <213> Artificial Sequence

<220>  
 <223> primer  
  
 <400> 42  
 25 agagttcggg ttatttgga tagaatc 27

<210> 43  
 <211> 20  
 <212> DNA  
 30 <213> Artificial Sequence

<220>  
 <223> primer  
  
 <400> 43  
 35 gaccgaaacg tcctaaaccg 20

<210> 44  
 <211> 24  
 <212> DNA  
 40 <213> Artificial Sequence

<220>  
 <223> primer  
  
 <400> 44  
 45 ataggttgga aacgtagttt ttcg 24

<210> 45  
 <211> 24  
 <212> DNA  
 50 <213> Artificial Sequence

<220>  
 <223> primer  
  
 <400> 45  
 55 ctataacgtc taacgaatcc tcgc 24

<210> 46  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 5  
 <220>  
 <223> primer  
 <400> 46  
 10 tattggttaa gattcggagt tgcga 25  
 <210> 47  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 15  
 <220>  
 <223> primer  
 <400> 47  
 20 ctcccaacat ttacctaata acgaa 25  
 <210> 48  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 25  
 <220>  
 <223> primer  
 <400> 48  
 30 tgttgtaaag tttgttcggt ttcgt 25  
 <210> 49  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 35  
 <220>  
 <223> primer  
 <400> 49  
 40 ttctacttca tttaaacccc tcgaa 25  
 <210> 50  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 45  
 <220>  
 <223> primer  
 <400> 50  
 50 gggaggggaag cgtagtttat tc 22  
 <210> 51  
 <211> 28  
 <212> DNA  
 55

<213> Artificial Sequence  
 <220>  
 <223> primer  
 5  
 <400> 51  
 ctaaaaaatt taaaaaaca aaaacgat 28  
 <210> 52  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> primer  
 15  
 <400> 52  
 ttttagagtt taggattggc ggc 23  
 <210> 53  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> primer  
 25  
 <400> 53  
 tacaacttaa cactaccga aaacg 25  
 <210> 54  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> primer  
 35  
 <400> 54  
 aggtatttat gggtaggaa ttata 25  
 <210> 55  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> primer  
 45  
 <400> 55  
 aaatcactta actaaaatcc cacta 25  
 <210> 56  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> primer  
 55

	<400> 56		
	tagttaggga taaagtgtga gtgtta		26
5	<210> 57		
	<211> 27		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
10	<223> primer		
	<400> 57		
	caacttatta aataactata cttcccc		27
15	<210> 58		
	<211> 30		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
20	<223> primer		
	<400> 58		
	gggatatttg ttttttatat ttaataaagt		30
25	<210> 59		
	<211> 25		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
30	<223> primer		
	<400> 59		
	acctcacaat aaaactacta caacc		25
35	<210> 60		
	<211> 26		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
40	<223> primer		
	<400> 60		
	ttattaggat gagttattgg ttattt		26
45	<210> 61		
	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
50	<223> primer		
	<400> 61		
	acctccctaa ccccaactac		20
55	<210> 62		

<211> 20  
 <212> DNA  
 <213> Artificial Sequence

5

<220>  
 <223> primer

<400> 62  
 gtttgggtag tttattggtt 20

10

<210> 63  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

15

<220>  
 <223> primer

<400> 63  
 ctaaaaattt ctcaactcct aactc 25

20

<210> 64  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

25

<220>  
 <223> primer

<400> 64  
 gagtgatat taaggatddd tgggt 25

30

<210> 65  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

35

<220>  
 <223> primer

<400> 65  
 tacaactcaa ctccaataac cacac 25

40

<210> 66  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

45

<220>  
 <223> primer

<400> 66  
 agttaggatt gtgtaatatt agtdddd 26

50

<210> 67  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

55

<220>

<223> primer  
 <400> 67  
 cactatacaa ccatcaacta caac 24  
 5

<210> 68  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
 10

<220>  
 <223> primer  
 <400> 68  
 aataagattt gtttagtttt attgttaaaa 30  
 15

<210> 69  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 20

<220>  
 <223> primer  
 <400> 69  
 tctacctaact actcacaacc cctaca 26  
 25

<210> 70  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 30

<220>  
 <223> primer  
 <400> 70  
 gtatatatttt agggaggagg gggg 24  
 35

<210> 71  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 40

<220>  
 <223> primer  
 <400> 71  
 caacccttac ttcacctcca cata 24  
 45

<210> 72  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 50

<220>  
 <223> primer  
 <400> 72  
 gtagtagtag agttttaga ggggtg 26  
 55

<210> 73  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 5  
 <220>  
 <223> primer  
 <400> 73  
 10 ataaaaaata cctatccaaa aaaaa 25  
 <210> 74  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 15  
 <220>  
 <223> primer  
 20 <400> 74  
 ggaagagggt taagtggatt ttttt 25  
 <210> 75  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 25  
 <220>  
 <223> primer  
 30 <400> 75  
 ttttctttcc aaaccaacc ta 22

35 **Claims**

1. A method for determining presence or absence of a cancer cell in a biological sample obtained from a subject comprising the steps of:

40 extracting DNA from the biological sample;  
 analyzing methylation status, for the DNA obtained in the extracting step, of at least one CpG site located in at least one base sequence selected from base sequences SEQ ID NOs: 1 to 14; and  
 determining presence or absence of the cancer cell in the biological sample based on an analysis result obtained in the analyzing step.

45 2. The method according to claim 1, wherein the CpG site is selected from:  
 the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
 the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st  
 50 to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence  
 SEQ ID NO: 2;  
 the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;  
 the 1 st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;  
 the 1 st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;  
 55 the 1 st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;  
 the 1 st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;  
 the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;  
 the 1 st, 2nd, 4th, 7th to 11 th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;

## EP 2 540 839 A1

the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;  
the 1 st to 3rd, 5th to 11 th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence  
SEQ ID NO: 11;  
the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ  
ID NO: 12;  
the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and  
the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID  
NO: 14.

- 5
- 10
- 15
- 20
- 25
- 30
3. The method according to claim 1 or 2, wherein the analyzing step is a step of analyzing methylation status of more than one CpG site.
  4. The method according to any one of claims 1 to 3, wherein the analyzing step is a step of analyzing presence or absence of methylation of the CpG site.
  5. The method according to claim 4, wherein in the determining step, it is determined that the cancer cell is present in the biological sample when the result obtained in the analyzing step indicates that there is a methylated CpG site.
  6. The method according to any one of claims 1 to 3, wherein the analyzing step is a step of analyzing methylation frequency of the CpG site.
  7. The method according to claim 6, wherein in the determining step, it is determined that the cancer cell is present in the biological sample when the result obtained in the analyzing step indicates that the methylation frequency is higher than a predetermined threshold.
  8. The method according to any one of claims 1 to 7, wherein the analyzing step comprises concentrating methylated DNA contained in the DNA obtained from the extracting step by immunoprecipitation and analyzing methylation status of the concentrated methylated DNA.
  9. A molecular marker for determination of presence or absence of a cancer cell by analysis of methylation status which is at least one CpG site selected from CpG sites located in base sequences SEQ ID NOs: 1 to 14.
  10. The molecular marker for determination according to claim 9, wherein the CpG site is selected from:

35

40

45

50

55

the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1;  
the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence  
SEQ ID NO: 2;  
the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;  
the 1 st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;  
the 1 st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;  
the 1 st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;  
the 1 st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;  
the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;  
the 1 st, 2nd, 4th, 7th to 11 th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;  
the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;  
the 1 st to 3rd, 5th to 11 th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence  
SEQ ID NO: 11;  
the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ  
ID NO: 12;  
the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and  
the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID  
NO: 14.

11. A kit for determination of presence or absence of a cancer cell in a biological sample obtained from a subject comprising:

a non-methylated cytosine conversion agent that converts non-methylated cytosine in DNA extracted from the

## EP 2 540 839 A1

biological sample to a different base; and

a primer set for determination of methylation status of at least one CpG site located in base sequences SEQ ID NOs: 1 to 14 by methylation specific PCR.

5 **12.** The kit for determination according to claim 11, wherein the CpG site is selected from:

the 1st, 3rd to 7th, 9th to 26th and 28th to 54th CpG sites from the 5' end of the base sequence SEQ ID NO: 1; the 1st to 11th, 13th to 23rd, 25th, 26th, 28th, 29th, 31st, 32nd, 34th, 35th, 38th, 40th to 44th, 46th to 49th, 51st to 57th, 59th to 66th, 68th, 70th to 73rd, 75th, 76th, 78th and 79th CpG sites from the 5' end of the base sequence SEQ ID NO: 2;

10 the 1st to 10th and 12th CpG sites from the 5' end of the base sequence SEQ ID NO: 3;

the 1 st to 3rd CpG sites from the 5' end of the base sequence SEQ ID NO: 4;

the 1 st to 4th CpG sites from the 5' end of the base sequence SEQ ID NO: 5;

the 1 st and 2nd CpG sites from the 5' end of the base sequence SEQ ID NO: 6;

15 the 1 st to 7th and 9th to 52nd CpG sites from the 5' end of the base sequence SEQ ID NO: 7;

the 1st to 16th, 18th to 25th and 27th to 39th CpG sites from the 5' end of the base sequence SEQ ID NO: 8;

the 1 st, 2nd, 4th, 7th to 11 th and 13th to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 9;

the 1st to 6th, 8th and 10th CpG sites from the 5' end of the base sequence SEQ ID NO: 10;

20 the 1 st to 3rd, 5th to 11 th, 13th, 15th to 19th and 21st to 23rd CpG sites from the 5' end of the base sequence SEQ ID NO: 11;

the 1st to 6th, 8th, 10th to 22nd, 25th to 28th and 32nd CpG sites from the 5' end of the base sequence SEQ ID NO: 12;

the 1st to 3rd, 7th to 13th, 15th and 19th CpG sites from the 5' end of the base sequence SEQ ID NO: 13; and

25 the 1st to 4th, 14th to 16th, 18th, 19th, 21st and 22nd CpG sites from the 5' end of the base sequence SEQ ID NO: 14.

**13.** The kit for determination according to claim 11 or 12, wherein the primer set is at least one selected from:

a primer set of the base sequences SEQ ID NO: 28 and SEQ ID NO: 29;

30 a primer set of the base sequences SEQ ID NO: 30 and SEQ ID NO: 31;

a primer set of the base sequences SEQ ID NO: 32 and SEQ ID NO: 33;

a primer set of the base sequences SEQ ID NO: 34 and SEQ ID NO: 35;

a primer set of the base sequences SEQ ID NO: 36 and SEQ ID NO: 37;

35 a primer set of the base sequences SEQ ID NO: 38 and SEQ ID NO: 39;

a primer set of the base sequences SEQ ID NO: 40 and SEQ ID NO: 41;

a primer set of the base sequences SEQ ID NO: 42 and SEQ ID NO: 43;

a primer set of the base sequences SEQ ID NO: 44 and SEQ ID NO: 45;

a primer set of the base sequences SEQ ID NO: 46 and SEQ ID NO: 47;

a primer set of the base sequences SEQ ID NO: 48 and SEQ ID NO: 49;

40 a primer set of the base sequences SEQ ID NO: 50 and SEQ ID NO: 51; and

a primer set of the base sequences SEQ ID NO: 52 and SEQ ID NO: 53.

45

50

55

[Fig. 1]

SEQ ID NO: 1

CpG site		1*	2	3*	4*	5*	6*	7*	8	9**	10**	11*	12*
PT	clone1	●	●	●	●	●	●	●	-	●	●	●	●
	clone2	●	●	○	○	○	○	○	-	●	○	●	○
	clone3	●	●	●	-	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	-	○	●	●	●
PN	clone1	○	●	○	○	○	○	○	○	○	○	○	○
	clone2	○	●	○	○	○	○	○	○	○	○	○	○
	clone3	○	●	○	○	○	○	○	○	○	○	○	○
	clone4	○	●	○	○	○	○	○	○	○	○	○	○
	clone5	○	●	○	○	○	○	○	○	○	○	○	○

CpG site		13*	14*	15*	16*	17*	18*	19*	20*	21*	22*	23*	24*
PT	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	●	○	○	○	○	○	○	○	○	○	○	○
	clone3	●	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		25*	26*	27	28**	29**	30**	31*	32*	33*	34*	35*	36*
PT	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	○	○	○	○	●	○	○	●	●	○	○	●
	clone3	○	○	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	●	○	○	○	○	○	○	○	○	○
	clone2	○	○	●	○	○	○	○	○	○	○	○	○
	clone3	○	○	●	○	○	○	○	○	○	○	○	○
	clone4	○	○	●	○	○	○	○	○	○	○	○	○
	clone5	○	○	●	○	○	○	○	○	○	○	○	○

CpG site		37*	38*	39*	40*	41*	42*	43*	44*	45*	46*	47*	48*
PT	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	●	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	○
PN	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		49*	50*	51*	52*	53*	54*
PT	clone1	●	●	●	●	●	●
	clone2	○	○	○	○	○	●
	clone3	●	●	●	●	●	●
	clone4	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	○
	clone2	○	○	○	○	○	○
	clone3	○	○	○	○	○	○
	clone4	○	○	○	○	○	○
	clone5	○	○	○	○	○	○

[Fig. 2-1]

SEQ ID NO: 2

CpG site		1*	2*	3*	4**	5**	6**	7**	8*	9*	10*	11*	12
PT	clone1	○	●	○	●	●	●	○	●	○	○	○	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	●	○	○	○	○	●	○	○	○	○
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	○	○	○	○	○	○	●
	clone2	○	○	○	○	○	○	○	○	○	○	○	●
	clone3	○	○	○	○	○	○	○	○	○	○	○	●
	clone4	○	○	○	○	○	○	○	○	○	○	○	●
	clone5	○	○	○	○	○	●	○	○	○	○	○	●

CpG site		13*	14*	15*	16*	17*	18*	19*	20*	21*	22*	23*	24
PT	clone1	●	●	○	○	●	●	○	○	○	●	○	○
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	○	●	●	○	●	○	○	○	○	○	○	○
	clone4	○	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	○	○	○	○	○	○	●
	clone2	○	○	○	○	○	○	○	○	○	○	○	●
	clone3	○	○	○	○	○	○	○	○	○	○	○	●
	clone4	○	○	○	○	○	○	○	○	○	○	○	●
	clone5	○	○	○	○	○	○	○	○	○	○	○	●

[Fig. 2-2]

CpG site		25*	26*	27	28*	29*	30	31*	32*	33	34*	35*	36
PT	clone1	○	●	○	○	○	○	○	○	○	○	○	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	○	●	○	●	○	●	○	○	●	●	○	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	●	○	○	●	○	○	●
	clone2	○	○	○	○	○	●	○	○	●	○	○	●
	clone3	○	○	○	○	○	●	○	○	●	○	○	●
	clone4	○	○	○	○	○	●	○	-	●	○	○	●
	clone5	○	○	●	○	○	●	○	○	●	○	○	●

CpG site		37	38*	39	40*	41*	42*	43*	44*	45	46*	47*	48*
PT	clone1	○	○	○	○	○	○	○	○	●	○	○	○
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	○	●	●	○	●	○	●	●	○	○	○	●
	clone4	●	●	○	●	●	●	●	●	●	●	●	●
PN	clone1	●	○	●	○	○	○	○	○	●	○	○	○
	clone2	●	○	●	○	○	○	○	○	●	○	○	○
	clone3	●	○	●	○	○	○	○	○	●	○	○	○
	clone4	●	○	●	○	○	○	○	○	●	○	○	○
	clone5	●	○	●	○	○	○	○	○	●	○	○	○

CpG site		49*	50	51*	52*	53*	54*	55*	56*	57*	58	59*	60*
PT	clone1	○	○	○	○	○	○	●	○	○	○	○	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	○	○	○	○	○	○	○	●	○	○	○	○
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	●	○	○	○	○	○	○	○	○	○	○
	clone2	○	●	○	○	○	○	○	○	○	●	○	○
	clone3	○	●	○	○	○	○	○	○	○	○	○	○
	clone4	○	●	○	○	○	○	○	○	○	○	○	○
	clone5	○	●	○	○	○	○	○	○	○	○	○	○

CpG site		61*	62*	63*	64*	65*	66*	67	68*	69	70*	71*	72*
PT	clone1	○	●	○	○	●	●	●	○	○	○	○	○
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	○	●	○	●	○	●	●	○	●	○	○	○
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
PN	clone1	○	○	○	○	○	●	●	○	●	○	○	○
	clone2	○	○	○	○	○	○	●	○	○	○	○	○
	clone3	○	○	○	○	○	○	●	○	●	○	○	○
	clone4	○	○	○	○	○	○	●	○	○	○	○	○
	clone5	○	○	○	○	○	○	●	○	○	○	○	○

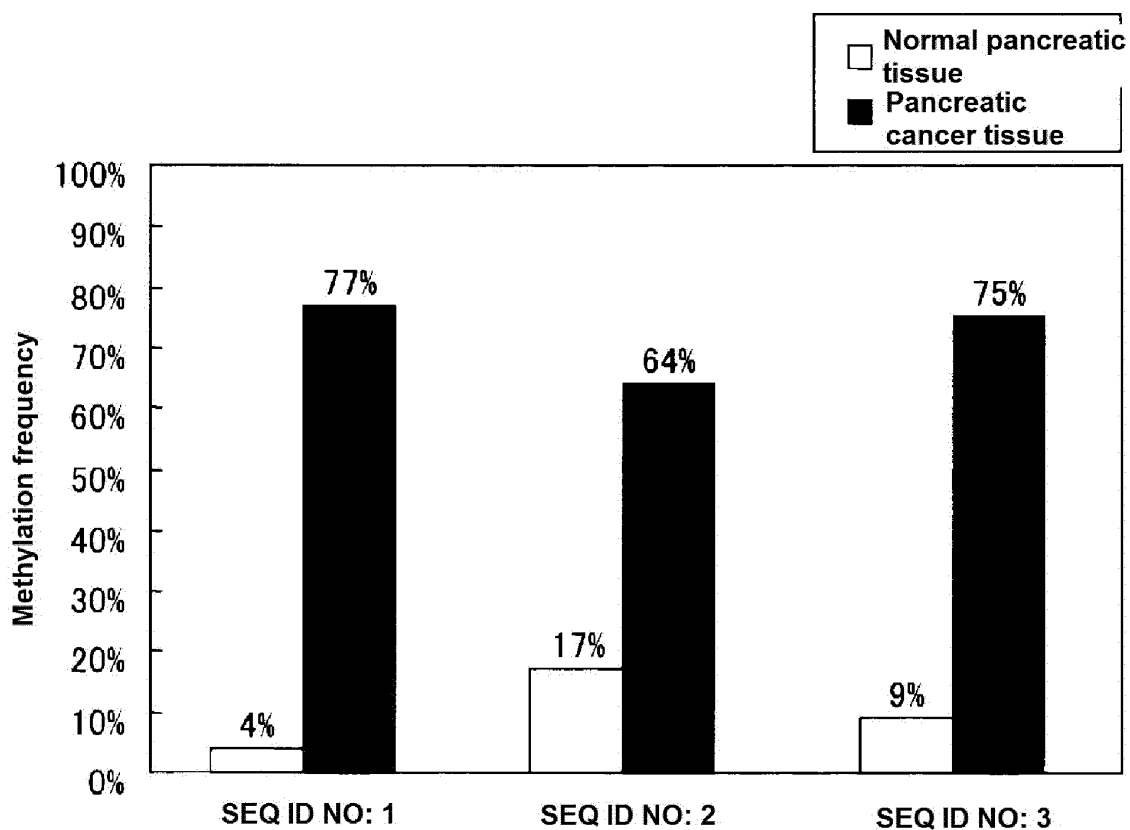
CpG site		73*	74	75*	76*	77	78*	79*
PT	clone1	○	○	○	●	○	●	○
	clone2	●	●	●	●	●	●	●
	clone3	○	○	○	○	○	○	○
	clone4	●	●	●	●	●	●	●
PN	clone1	○	●	○	○	●	○	○
	clone2	○	●	○	○	●	○	○
	clone3	○	●	○	○	●	○	○
	clone4	○	●	○	○	●	○	○
	clone5	○	●	○	○	●	○	○

[Fig. 3]

SEQ ID NO: 3

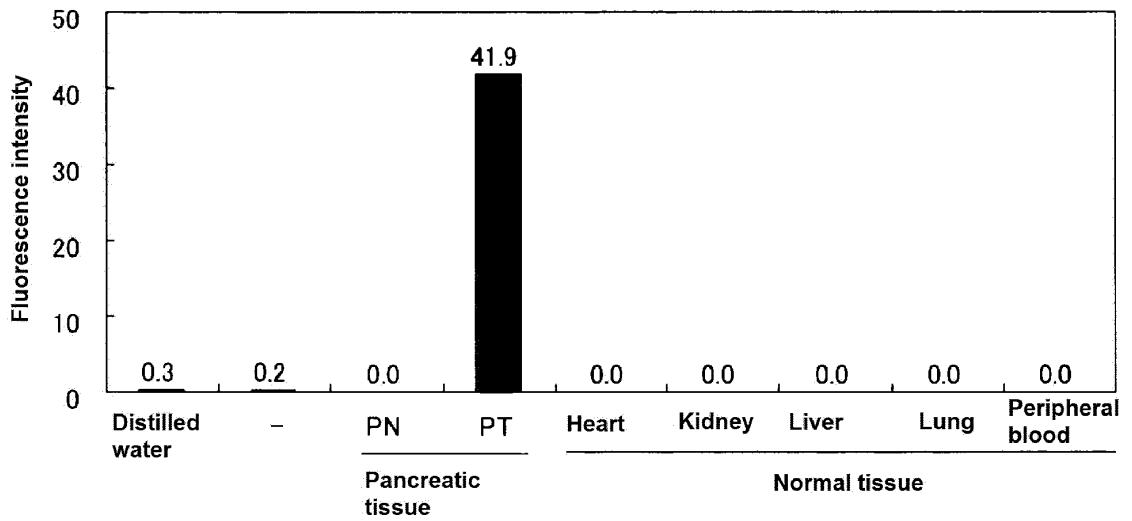
CpG site		1*	2*	3*	4*	5*	6**	7*	8*	9**	10*	11	12*
PT	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	●	○	●	○	○	○	○	○	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
	clone5	○	●	●	○	○	○	○	○	○	○	●	○
PN	clone1	-	-	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	●	○
	clone3	○	○	●	○	○	○	○	○	○	●	●	○
	clone4	-	○	○	○	○	○	○	○	○	○	○	○

[Fig. 4]



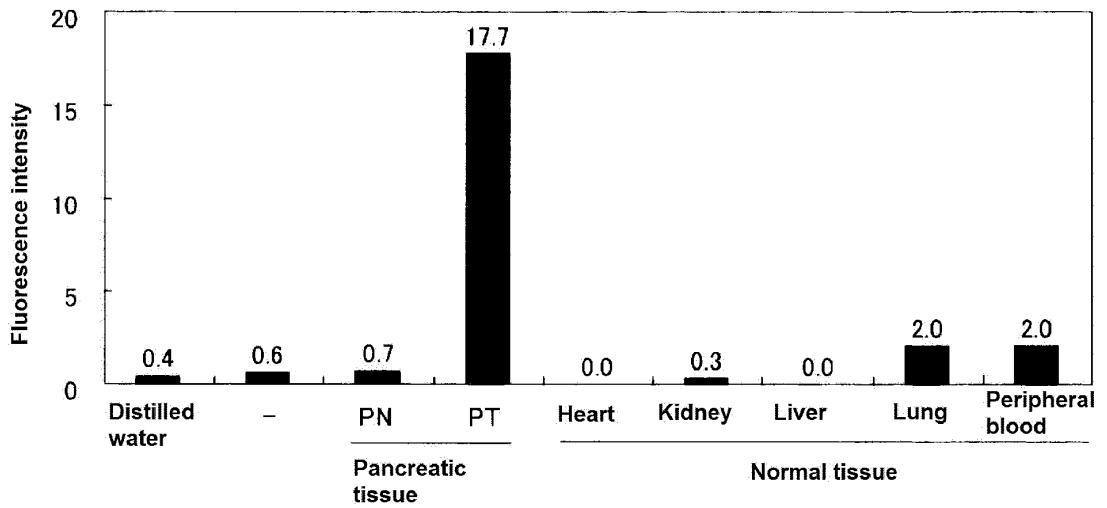
[Fig. 5]

SEQ ID NO: 15



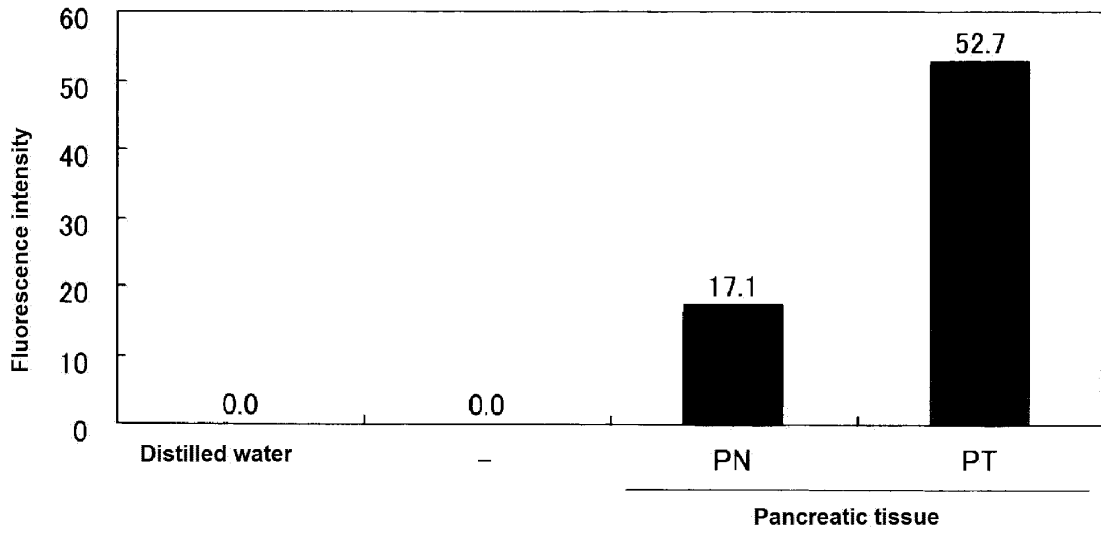
[Fig. 6]

SEQ ID NO: 16



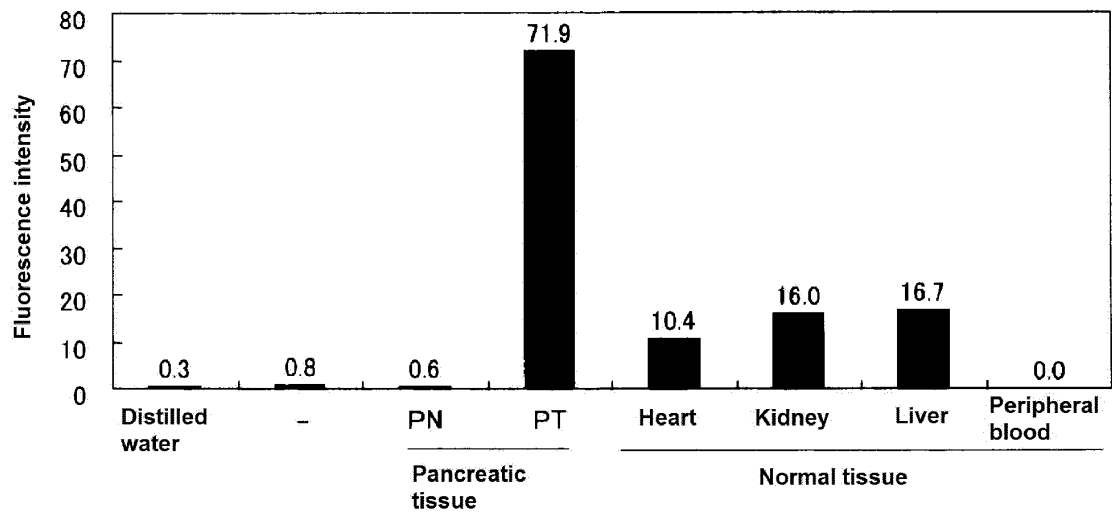
[Fig. 7]

SEQ ID NO: 17



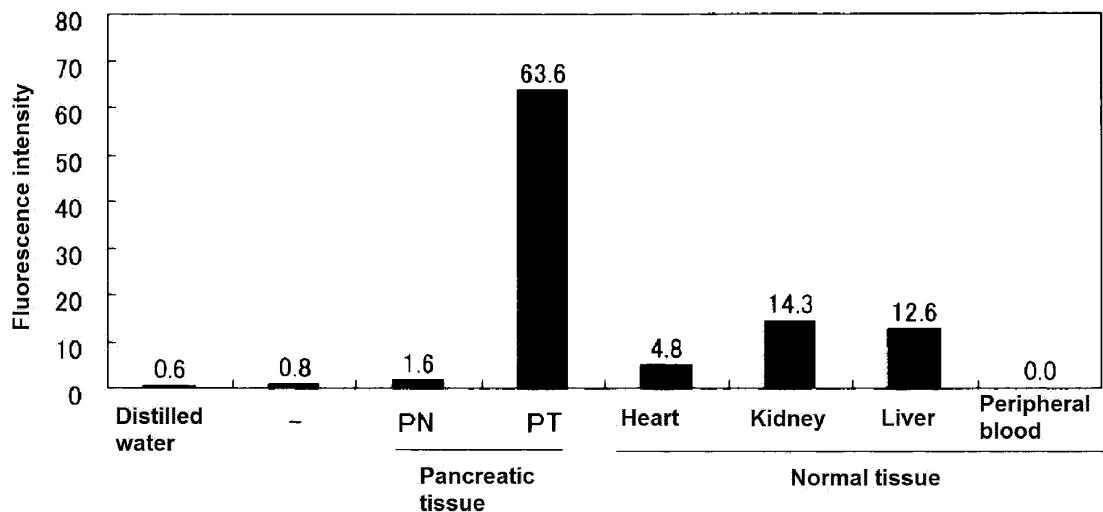
[Fig. 8]

SEQ ID NO: 18



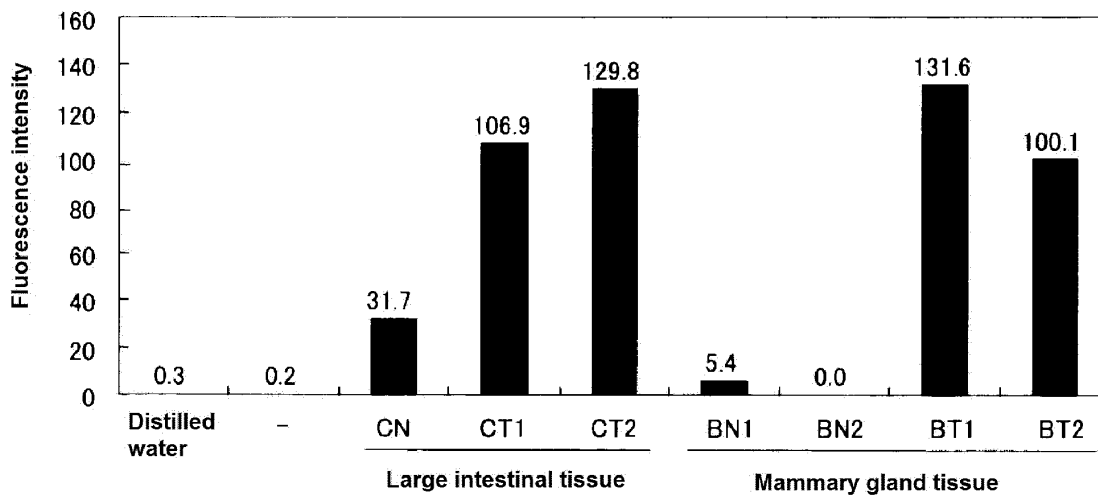
[Fig. 9]

SEQ ID NO: 19



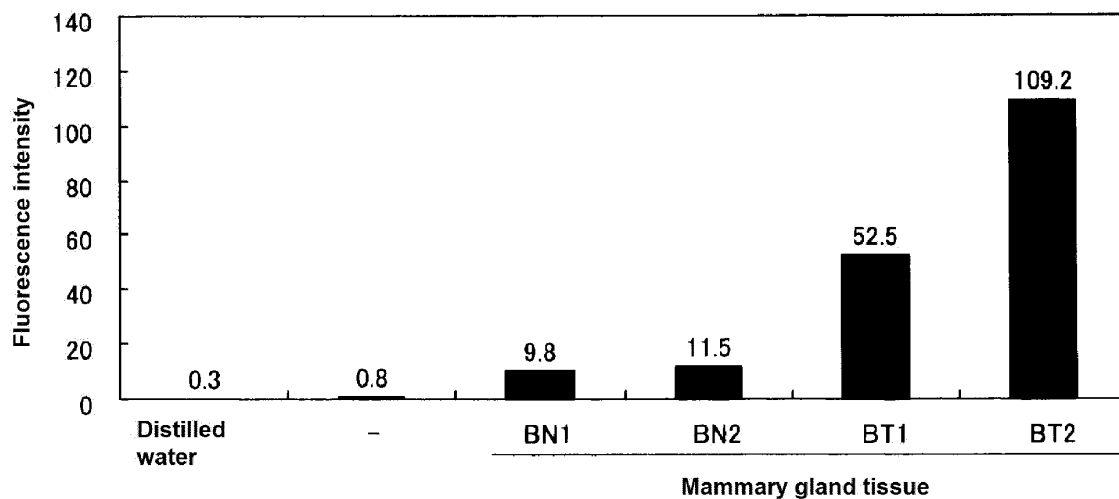
[Fig. 10]

SEQ ID NO: 15



[Fig. 11]

SEQ ID NO: 18



[Fig. 12-1]

SEQ ID NO: 7

CpG site		1*	2*	3**	4**	5*	6*	7*	8	9*	10*	11**	12**
MCF7	clone1	●	●	●	○	●	●	○	●	○	●	●	●
	clone2	○	●	○	●	○	●	○	○	○	○	●	●
	clone3	●	●	●	●	●	●	●	○	●	●	●	●
	clone4	●	●	●	●	○	●	○	●	●	●	●	●
	clone5	●	●	●	●	○	●	●	○	○	○	●	●
	clone6	●	●	●	●	●	●	●	○	●	○	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	●	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		13*	14*	15*	16*	17*	18*	19*	20*	21*	22*	23*	24*
MCF7	clone1	○	○	●	●	●	●	●	●	●	●	●	●
	clone2	○	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
	clone5	●	●	●	●	●	●	●	●	●	●	●	●
	clone6	●	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

[Fig. 12-2]

CpG site		25*	26*	27*	28*	29*	30*	31*	32*	33*	34*	35*	36*
MCF7	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
	clone5	●	●	●	●	●	●	●	●	●	●	●	●
	clone6	●	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	●	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		37*	38*	39*	40*	41*	42*	43*	44*	45*	46*	47*	48*
MCF7	clone1	●	●	●	●	○	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	●	●	●	○	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	○	●
	clone5	●	●	●	●	●	●	●	●	●	●	●	○
	clone6	●	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	●	○	○	○	○	○	○
	clone3	○	○	○	●	○	○	○	○	○	○	○	○
	clone4	○	○	●	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		49*	50*	51*	52*
MCF7	clone1	●	●	●	●
	clone2	●	●	●	●
	clone3	●	●	●	●
	clone4	●	●	●	○
	clone5	●	●	●	●
	clone6	●	●	●	●
HMEC	clone1	○	○	○	○
	clone2	○	○	○	○
	clone3	○	○	○	○
	clone4	○	○	○	○
	clone5	○	○	○	○

[Fig. 13]

SEQ ID NO: 8

CpG site		1*	2*	3*	4*	5*	6*	7*	8**	9**	10*	11*	12*
MCF7	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	●	●	○	●	○	○
	clone3	●	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		13*	14**	15**	16**	17	18*	19*	20*	21*	22*	23*	24*
MCF7	clone1	●	●	●	●	●	●	●	●	●	●	●	●
	clone2	○	●	●	●	●	●	●	●	○	●	●	●
	clone3	●	●	●	●	●	●	●	●	○	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	●	○	○	○	○	○	○	○
	clone2	○	○	○	○	●	○	○	○	○	○	○	○
	clone3	○	○	●	○	●	●	○	○	○	○	○	●
	clone4	○	○	○	○	●	○	○	○	○	○	○	○

CpG site		25*	26	27*	28*	29*	30*	31*	32*	33*	34*	35*	36*
MCF7	clone1	●	●	●	●	●	●	●	●	●	●	○	●
	clone2	●	●	●	●	●	●	●	●	●	●	●	●
	clone3	●	●	○	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	○	●
HMEC	clone1	○	●	○	○	○	○	○	○	○	○	○	○
	clone2	○	●	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	●	○	●	○	○	○	○	○	-	●
	clone4	○	●	○	○	○	○	○	○	○	○	○	○

CpG site		37*	38*	39*
MCF7	clone1	●	●	●
	clone2	●	●	●
	clone3	●	●	○
	clone4	●	●	●
HMEC	clone1	○	○	○
	clone2	○	○	○
	clone3	○	○	○
	clone4	○	○	○

[Fig. 14]

SEQ ID NO: 9

CpG site		1*	2*	3	4*	5	6	7*	8*	9*	10*	11*	12
MCF7	clone1	●	●	●	○	●	○	●	●	●	●	●	●
	clone2	●	●	●	●	○	○	○	○	○	●	●	○
	clone3	●	●	●	●	●	○	●	●	●	●	●	●
	clone4	●	●	○	●	○	○	●	○	○	●	●	○
	clone5	●	●	○	●	○	●	●	●	●	●	●	○
	clone6	●	●	○	●	○	●	●	●	○	●	●	○
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	●	○	●	○	○	○	○	○	○	○	○	○
	clone3	●	○	●	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○	○
	clone6	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		13*	14*	15**	16**	17*	18*	19*	20*	21*	22**	23*
MCF7	clone1	●	●	●	●	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	○	●	○	○	●
	clone3	●	●	●	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	●	●	●	●	●	●	●
	clone5	●	●	●	●	●	○	●	●	●	●	●
	clone6	●	●	●	●	●	●	●	●	●	●	●
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	○
	clone6	○	○	○	○	○	○	○	○	○	○	○

[Fig. 15]

SEQ ID NO: 10

CpG site		1*	2**	3*	4*	5*	6**	7	8*	9	10*
MCF7	clone1	●	●	○	●	●	●	●	●	●	●
	clone2	●	●	●	●	●	●	●	●	○	○
	clone3	●	●	●	●	●	●	●	●	○	●
	clone4	●	●	●	●	●	●	●	●	●	●
	clone5	○	●	●	●	○	○	●	●	○	○
HMEC	clone1	●	○	○	○	○	○	●	○	○	○
	clone2	○	○	○	○	○	○	●	○	○	○
	clone3	○	○	○	○	○	○	●	○	○	○
	clone4	○	○	○	○	○	○	●	○	○	○
	clone5	○	○	○	○	○	○	●	○	○	○

[Fig. 16]

SEQ ID NO: 11

CpG site		1**	2**	3**	4	5*	6*	7*	8*	9*	10**	11*	12
MCF7	clone1	●	●	●	●	●	●	●	●	●	○	●	○
	clone2	●	●	●	●	●	●	●	●	●	●	●	○
	clone3	●	●	●	●	●	●	●	●	●	●	●	○
	clone4	○	○	○	●	●	●	○	●	●	●	●	●
HMEC	clone1	○	○	○	●	○	○	○	○	○	○	○	○
	clone2	○	○	○	●	○	○	-	-	○	○	○	○
	clone3	○	○	○	●	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	-	-	○	○	○	○

CpG site		13*	14	15*	16*	17*	18*	19*	20	21*	22*	23*
MCF7	clone1	●	○	●	●	●	●	●	○	●	●	●
	clone2	●	●	●	●	●	●	●	○	○	○	●
	clone3	●	○	●	●	●	●	●	○	●	●	●
	clone4	○	○	●	●	●	●	●	●	○	○	○
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○

[Fig. 17]

SEQ ID NO: 12

CpG site		1*	2*	3*	4*	5*	6*	7	8*	9	10**	11**	12*
MCF7	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	●	●	●	●	●	●	○	●	●	●	●	●
	clone3	○	●	●	●	●	●	○	●	●	●	●	●
	clone4	●	●	●	●	●	●	○	●	●	●	●	●
	clone5	○	○	○	○	○	○	○	○	○	○	○	○
	clone6	○	○	○	○	○	○	○	○	○	○	○	○
HMEC	clone1	-	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	●	○	○	○
	clone4	○	○	○	○	●	●	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	●	○	○	○
	clone6	○	○	○	○	○	○	○	○	○	○	○	○

CpG site		13*	14*	15*	16**	17**	18**	19**	20*	21*	22*	23	24
MCF7	clone1	○	○	●	●	○	○	○	-	○	○	-	○
	clone2	●	●	●	●	●	●	●	●	●	●	●	○
	clone3	●	●	●	●	●	●	●	●	●	●	●	○
	clone4	●	●	●	●	●	●	●	●	●	●	●	○
	clone5	○	○	●	○	●	○	○	-	○	○	-	●
	clone6	○	○	●	○	○	○	○	-	○	○	-	○
HMEC	clone1	○	○	○	○	-	○	○	○	○	○	-	○
	clone2	○	○	○	○	○	○	○	○	○	○	-	○
	clone3	○	○	○	○	○	○	○	○	○	○	-	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○	○	○	-	○
	clone6	○	○	○	○	-	○	○	○	○	○	-	○

CpG site		25*	26*	27*	28*	29	30	31	32*
MCF7	clone1	○	○	○	○	●	●	○	●
	clone2	●	●	●	●	○	○	○	-
	clone3	●	●	●	●	●	●	●	●
	clone4	●	●	●	●	○	○	○	○
	clone5	●	●	○	○	○	○	○	●
	clone6	●	○	●	○	○	○	○	-
HMEC	clone1	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○
	clone5	○	○	○	○	○	○	○	○
	clone6	○	○	○	○	○	○	○	○

[Fig. 18]

SEQ ID NO: 13

CpG site		1*	2*	3*	4	5	6	7**	8**	9*	10*	11*	12*
MCF7	clone1	●	●	○	○	○	○	○	●	●	●	●	●
	clone2	●	●	●	●	○	○	●	●	●	●	●	●
	clone3	●	●	●	●	○	○	●	●	●	●	●	●
	clone4	○	○	○	○	○	●	●	●	●	●	●	●
	clone5	●	●	●	○	○	○	●	●	○	●	●	●
HMEC	clone1	●	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	●	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	●	●	○	○	○	○	○	○	○	○	●

CpG site		13*	14	15**	16	17	18	19*
MCF7	clone1	●	○	●	●	○	●	●
	clone2	●	○	●	○	○	○	●
	clone3	●	●	●	●	●	●	●
	clone4	●	○	●	○	○	○	●
	clone5	●	○	●	○	○	○	●
HMEC	clone1	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○

[Fig. 19]

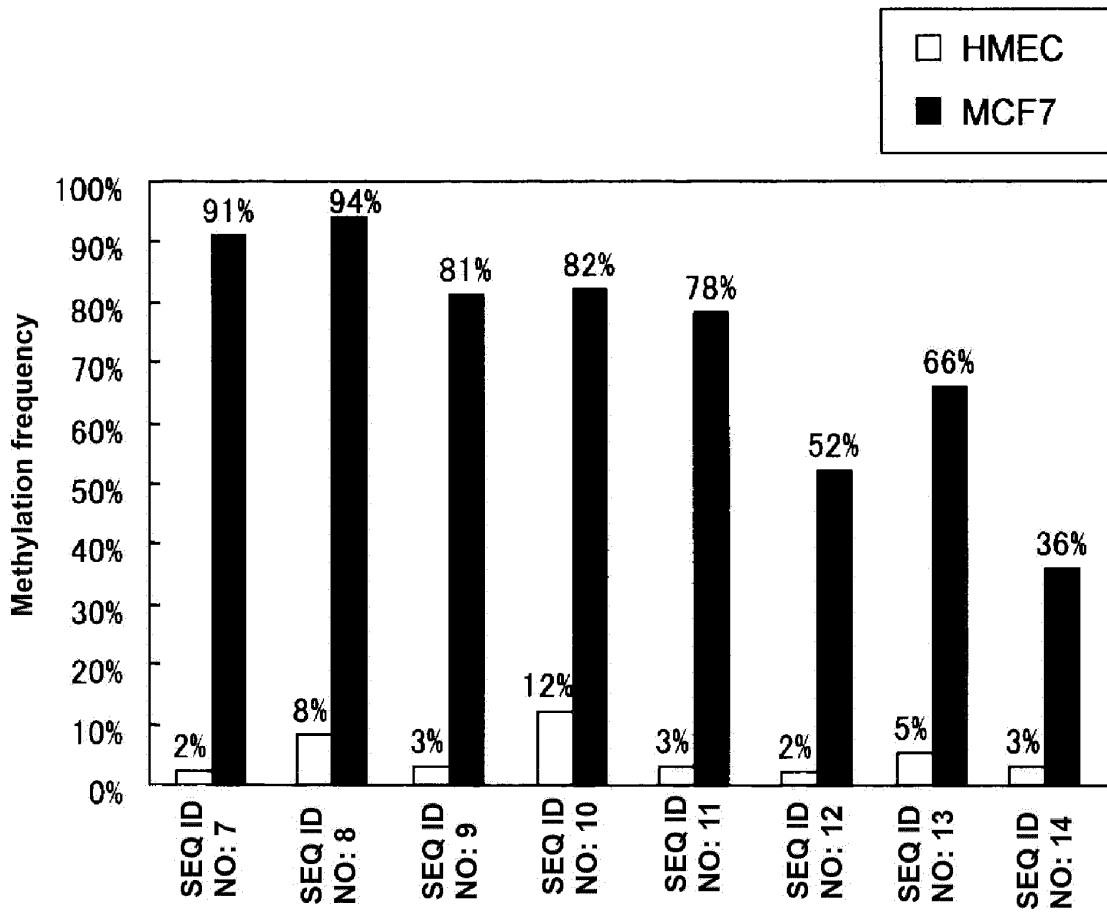
SEQ ID NO: 14

CpG site		1**	2**	3*	4**	5	6	7	8	9	10	11	12
MCF7	clone1	●	●	●	●	○	○	○	○	○	○	○	○
	clone2	●	○	●	●	○	○	○	○	○	○	○	○
	clone3	●	○	○	●	○	○	○	○	○	○	○	○
	clone4	●	●	●	●	○	○	○	○	○	○	○	○
HMEC	clone1	○	○	○	○	●	○	○	○	○	○	○	○
	clone2	○	○	○	○	●	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	●	○	○	○	○	○	○	○

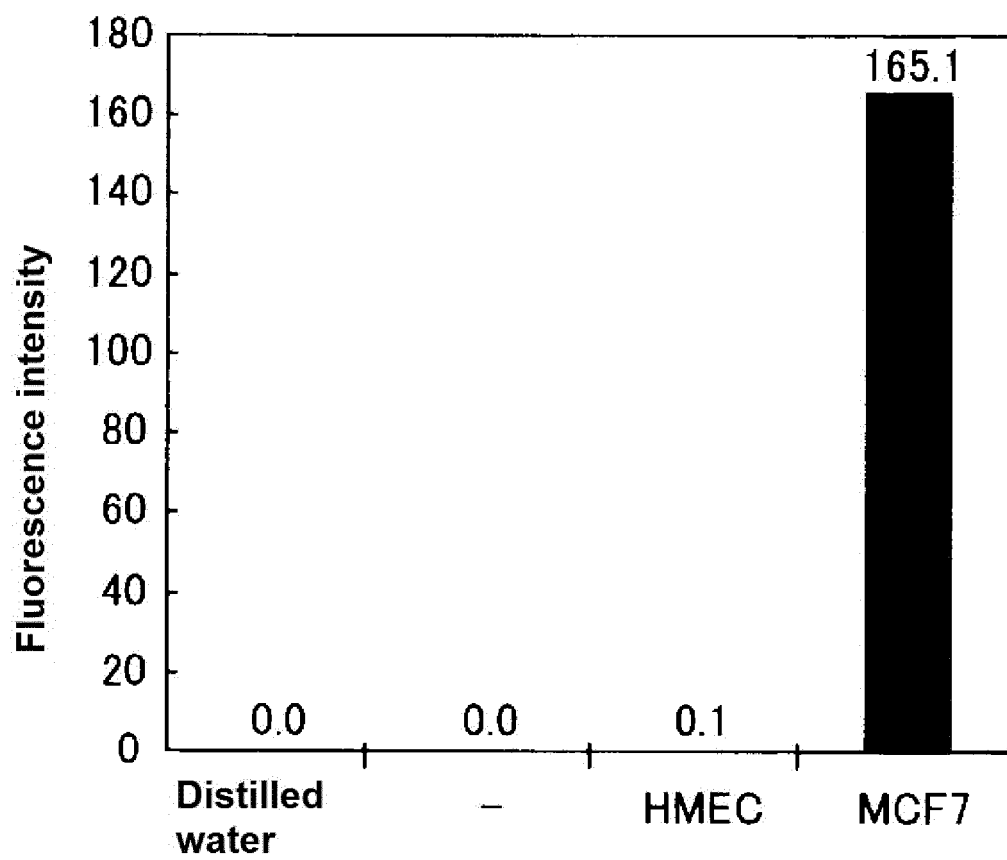
CpG site		13	14*	15*	16*	17	18*	19*	20	21*	22*	23	24
MCF7	clone1	○	○	●	●	○	●	●	●	●	●	●	●
	clone2	○	●	●	●	○	●	○	○	○	○	○	○
	clone3	○	●	●	○	○	●	○	○	○	○	○	○
	clone4	○	○	●	●	○	●	●	○	●	●	○	○
HMEC	clone1	○	○	○	○	○	○	○	○	○	○	○	○
	clone2	○	○	○	○	○	○	○	○	○	○	○	○
	clone3	○	○	○	○	○	○	○	○	○	○	○	○
	clone4	○	○	○	○	○	○	○	○	○	○	○	○

[Fig. 20]



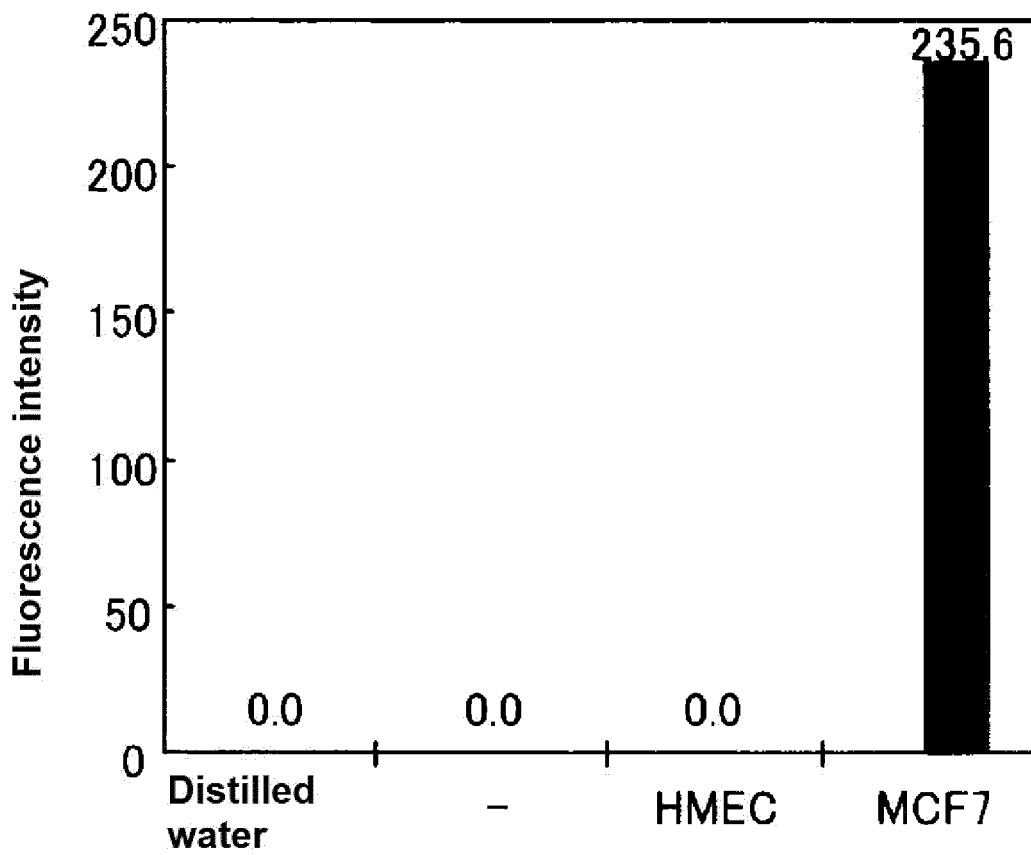
[Fig. 21]

SEQ ID NO: 20



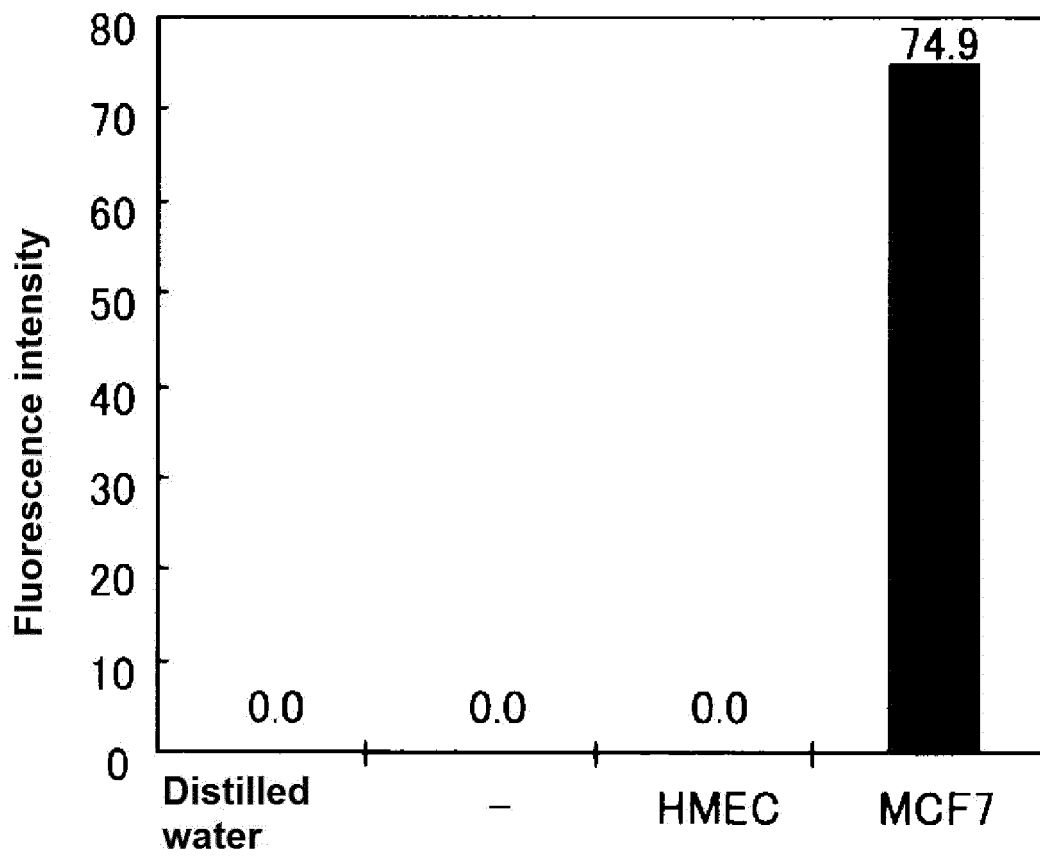
[Fig. 22]

SEQ ID NO: 21



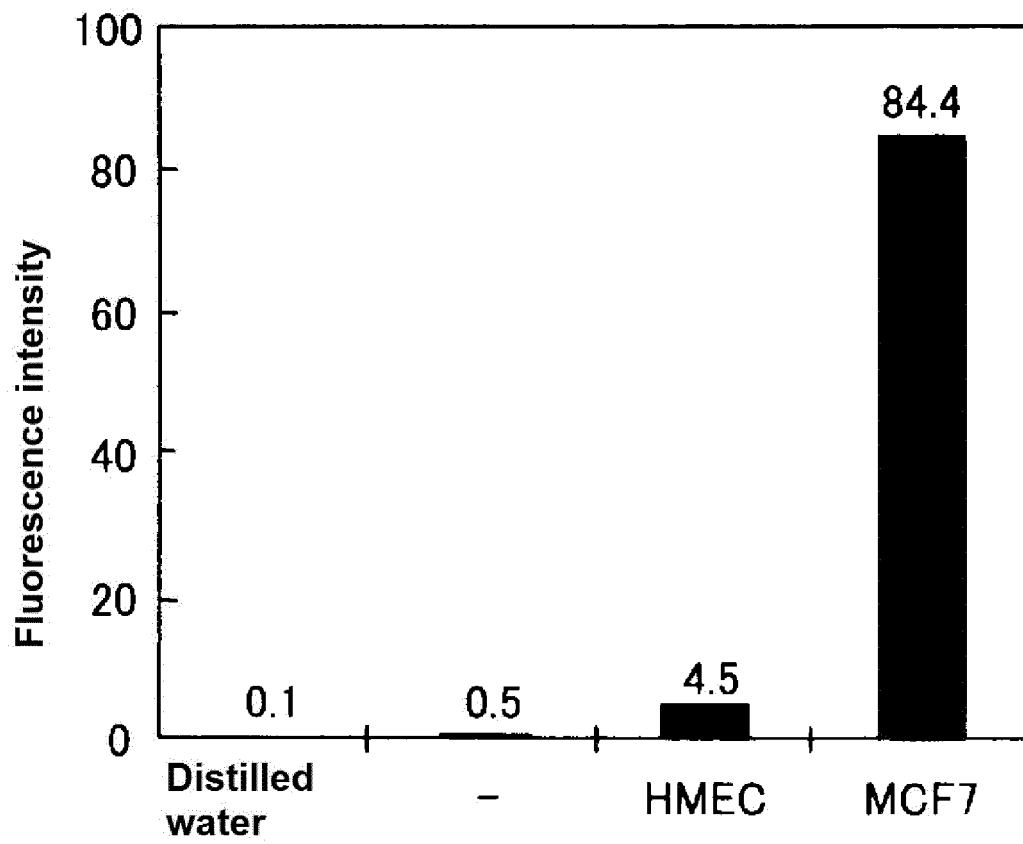
[Fig. 23]

SEQ ID NO: 22



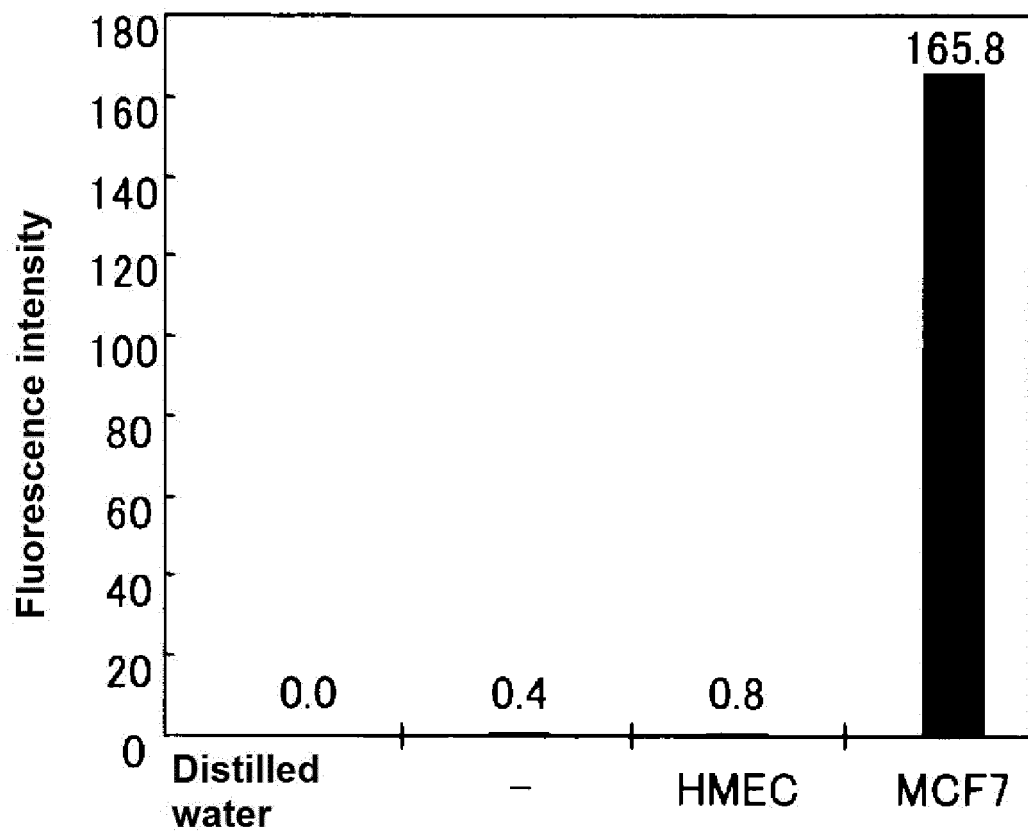
[Fig. 24]

SEQ ID NO: 23



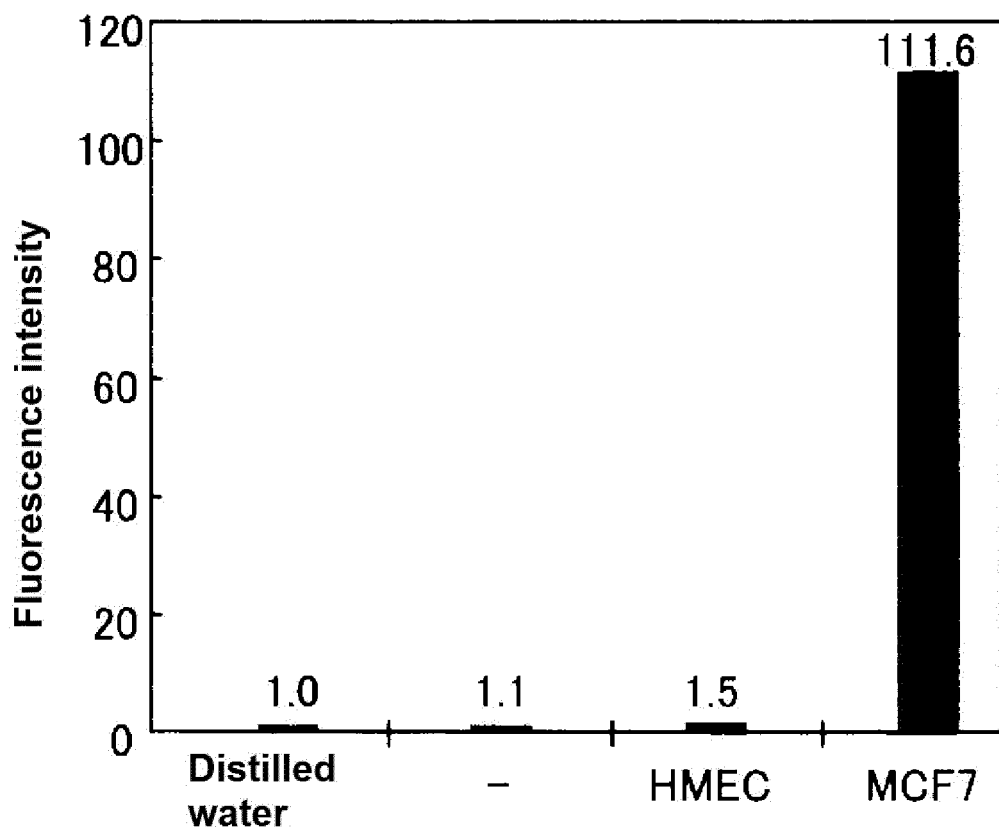
[Fig. 25]

SEQ ID NO: 24



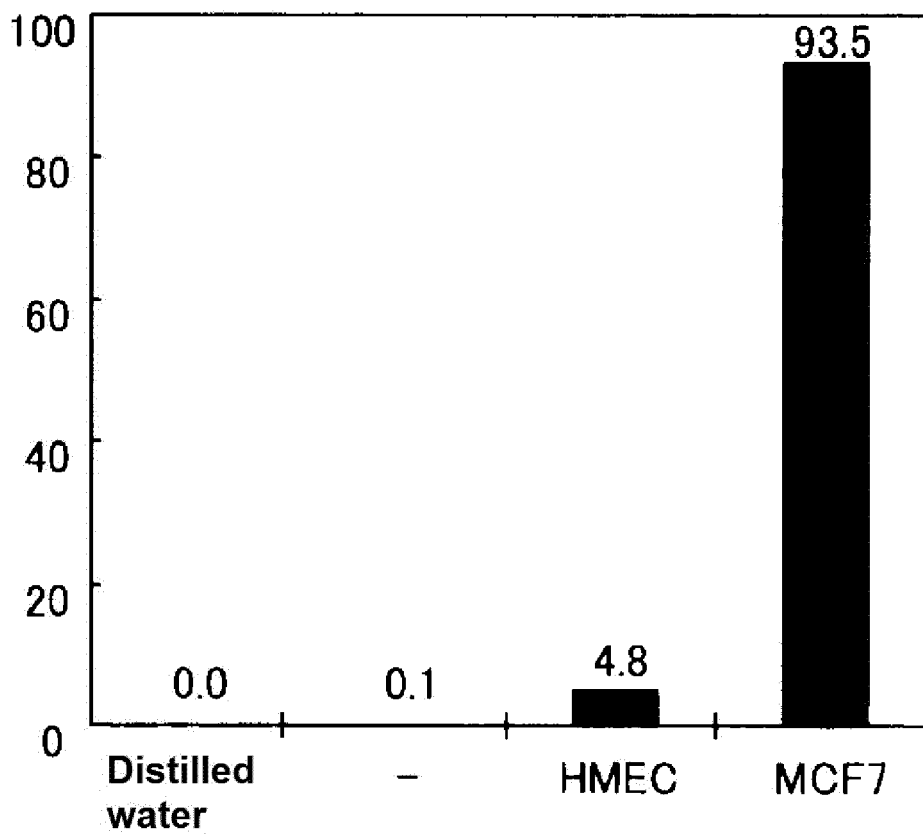
[Fig. 26]

SEQ ID NO: 25



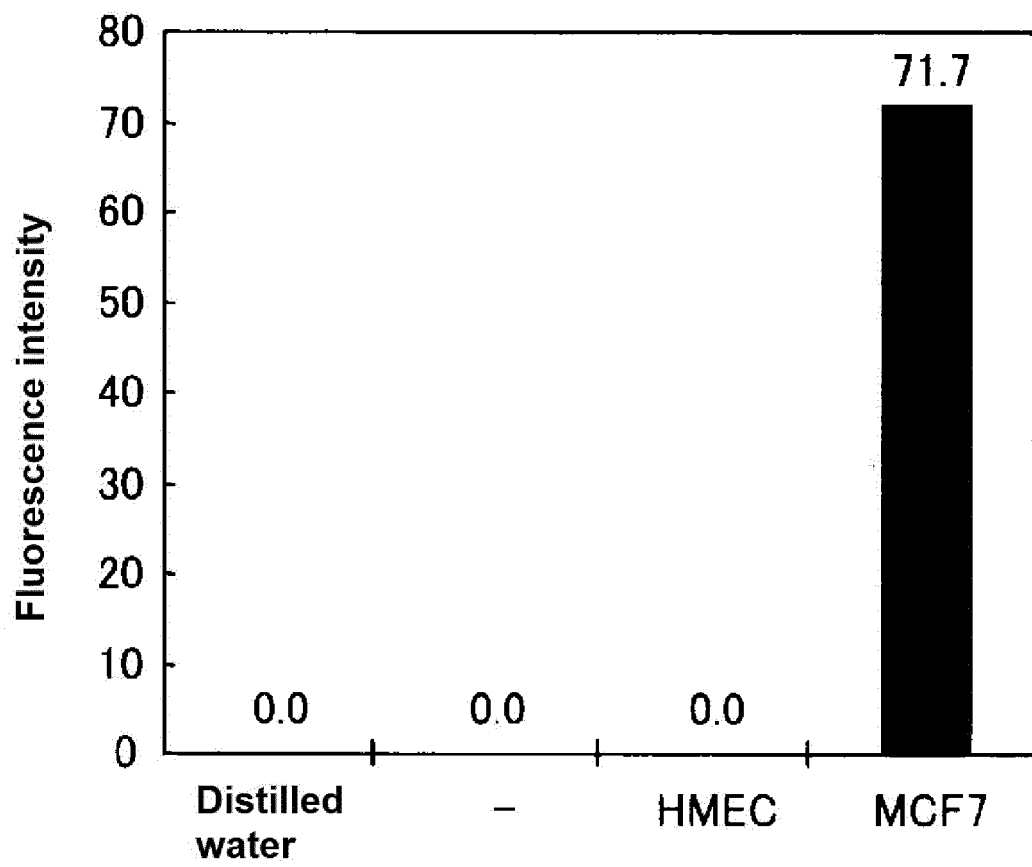
[Fig. 27]

SEQ ID NO: 26



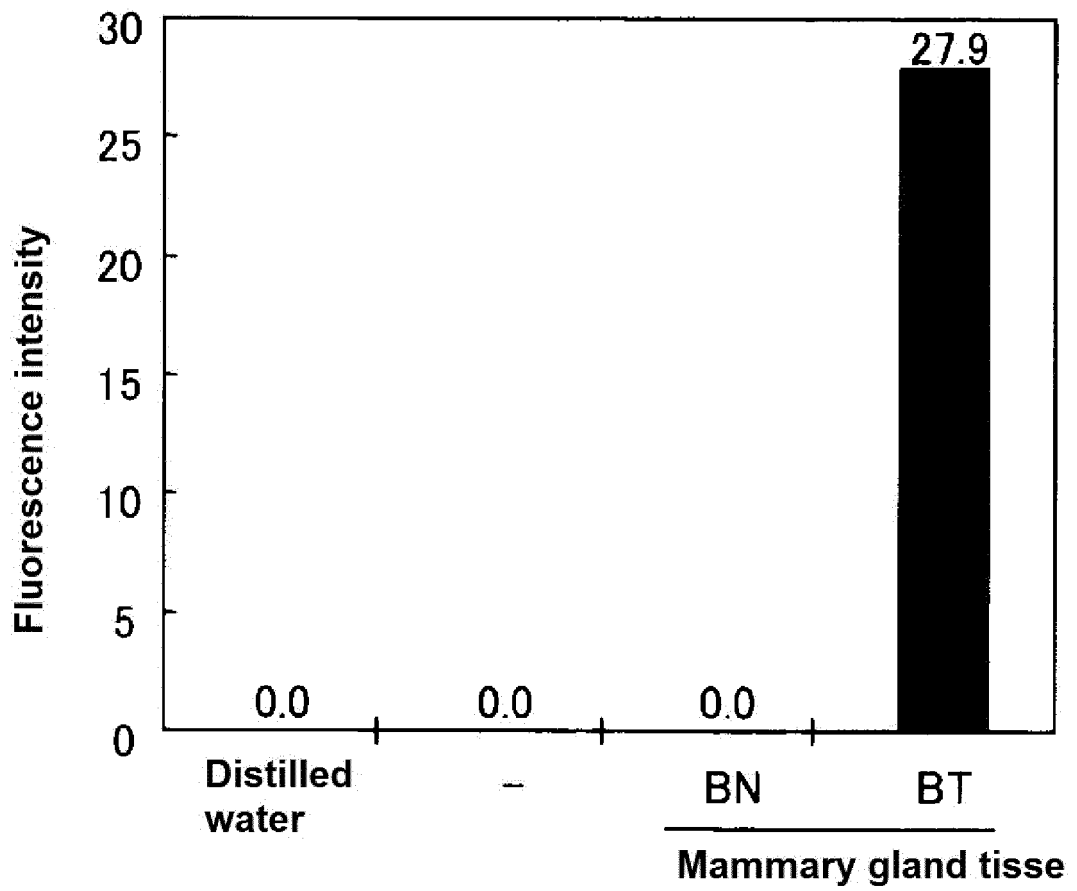
[Fig. 28]

SEQ ID NO: 27



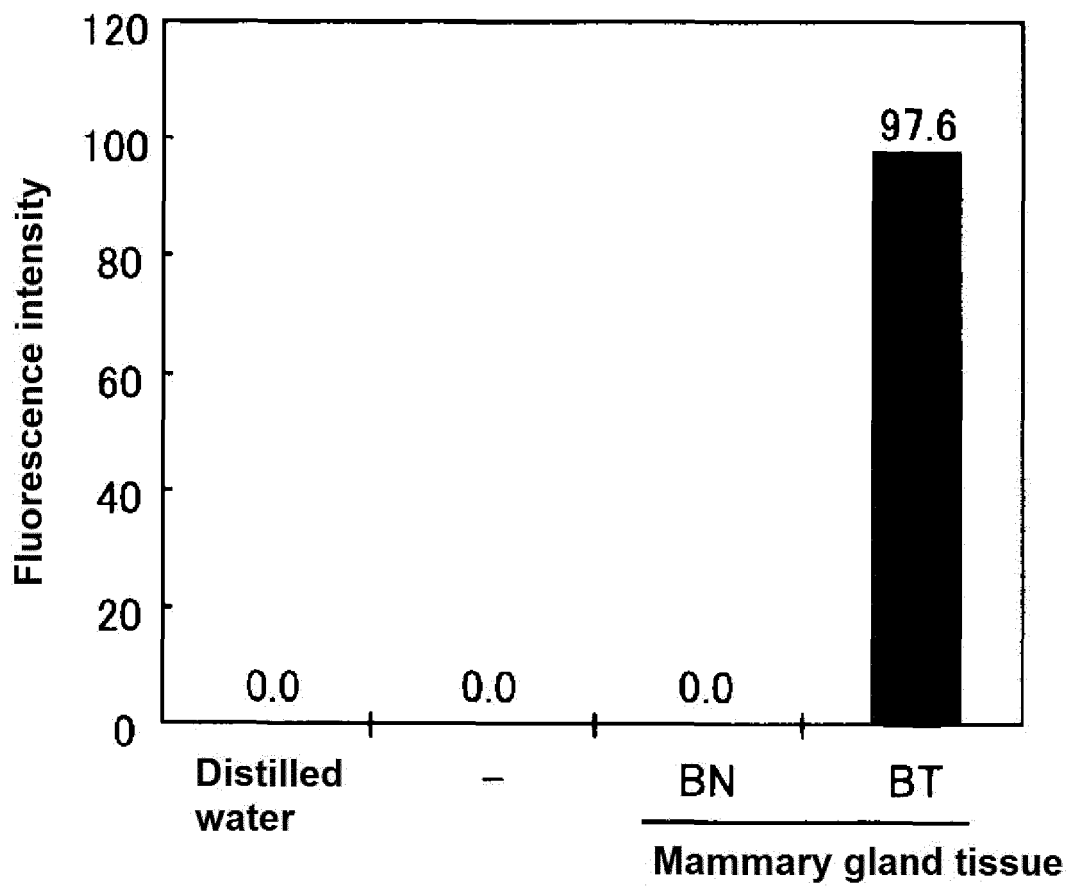
[Fig. 29]

SEQ ID NO: 20



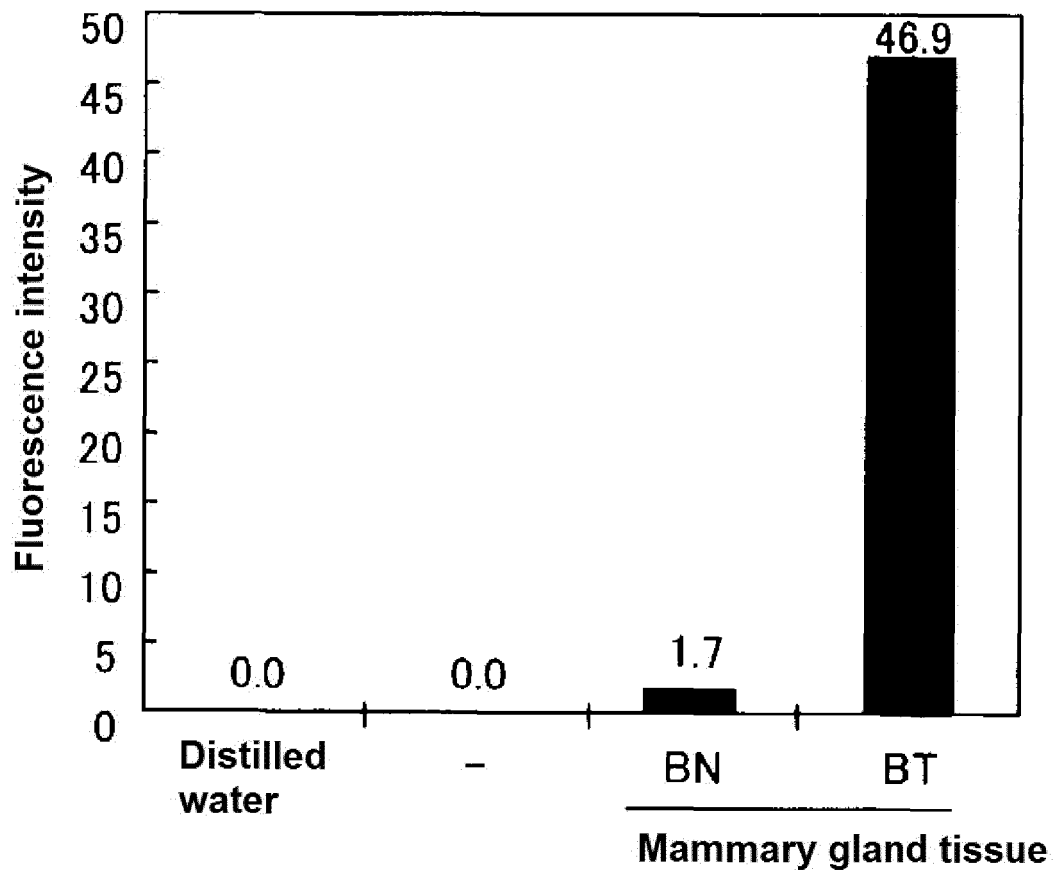
[Fig. 30]

SEQ ID NO: 21



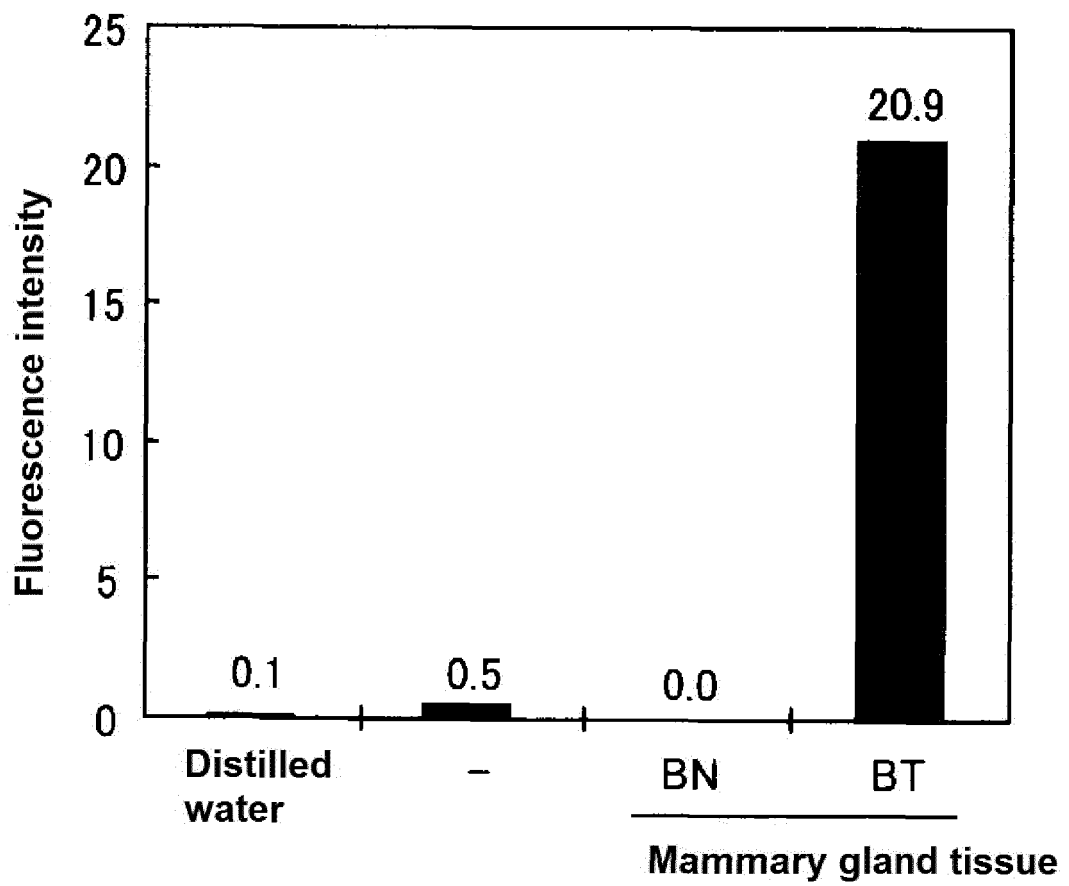
[Fig. 31]

SEQ ID NO: 22



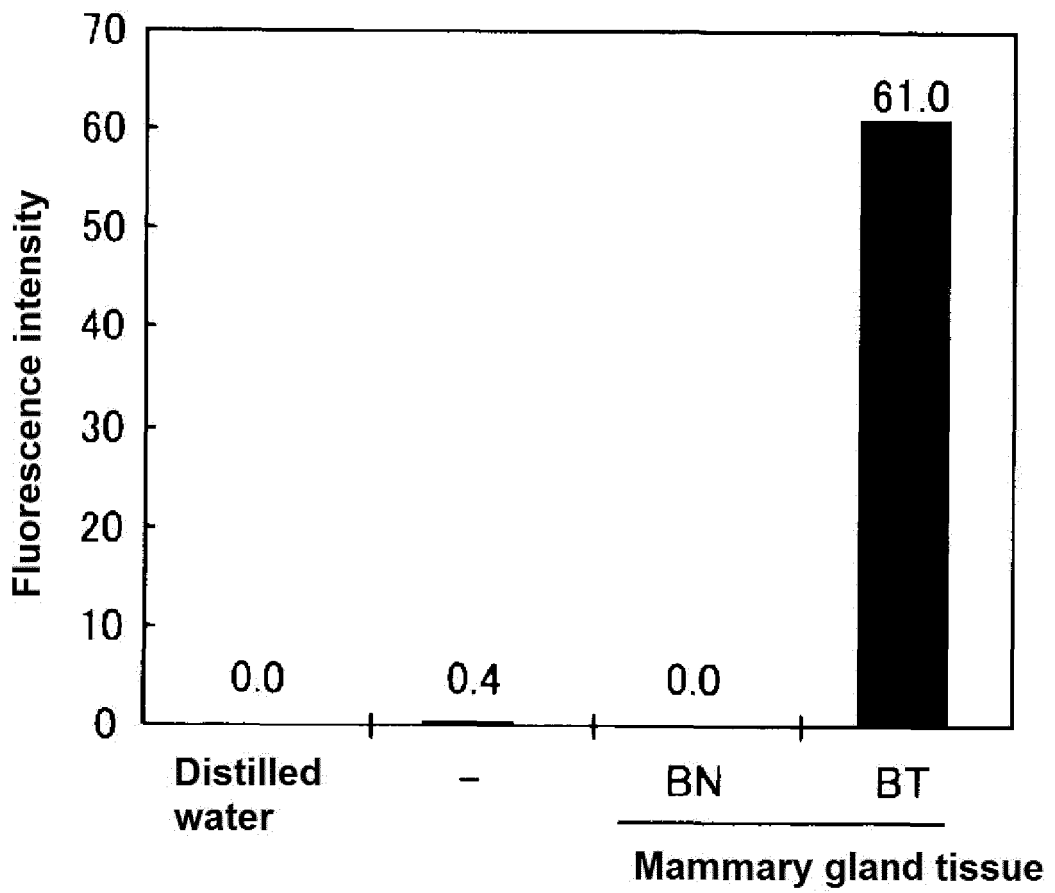
[Fig. 32]

SEQ ID NO: 23



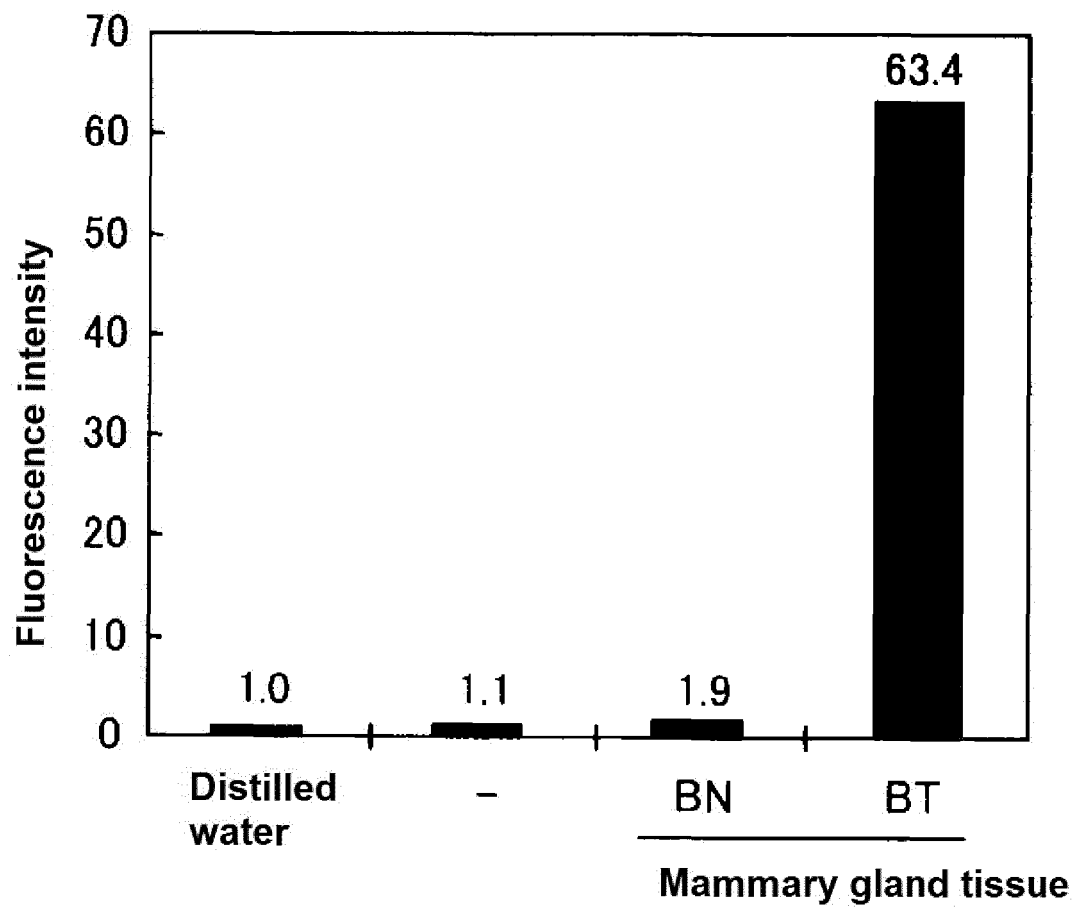
[Fig. 33]

SEQ ID NO: 24



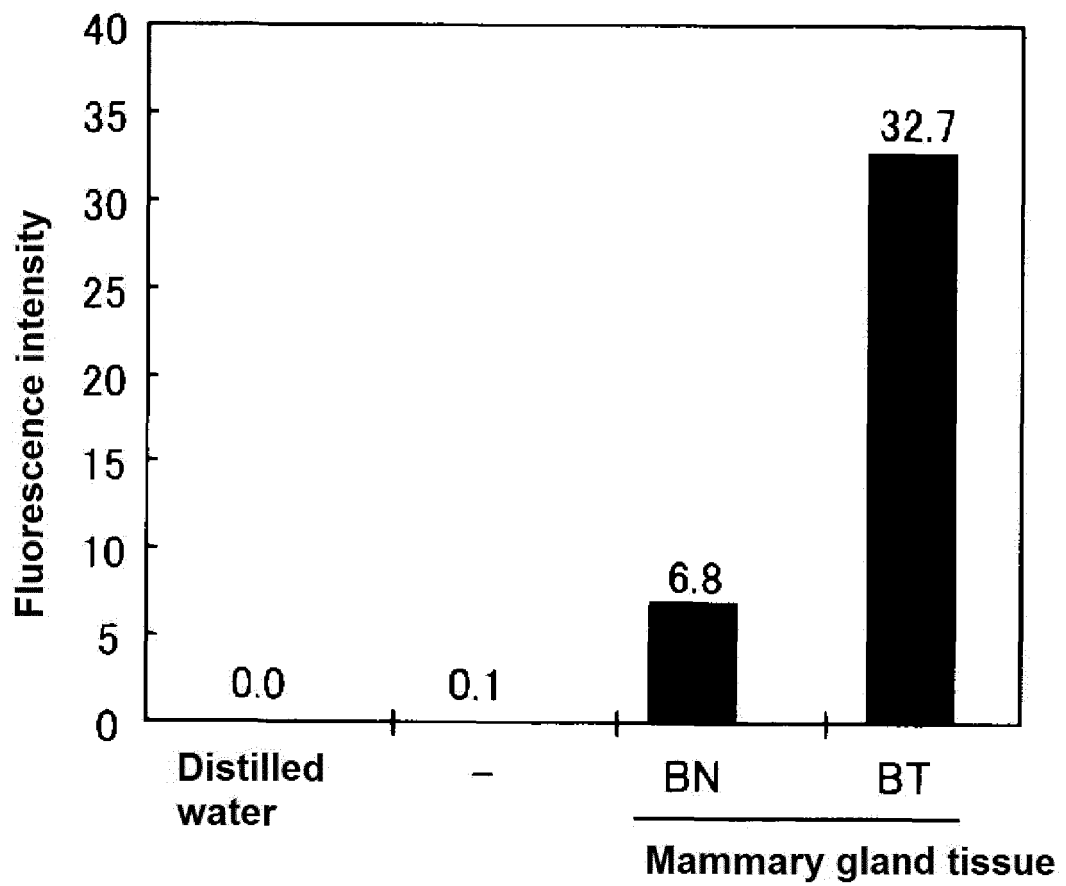
[Fig. 34]

SEQ ID NO: 25



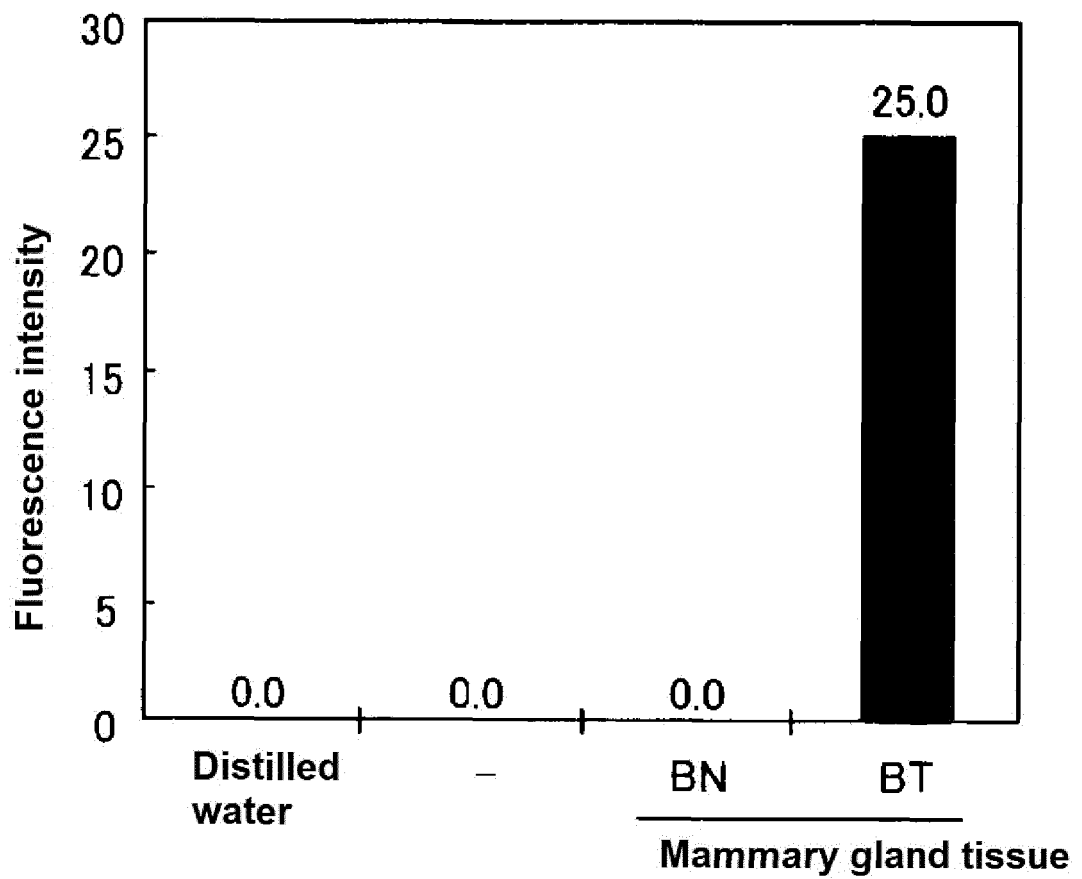
[Fig. 35]

SEQ ID NO: 26

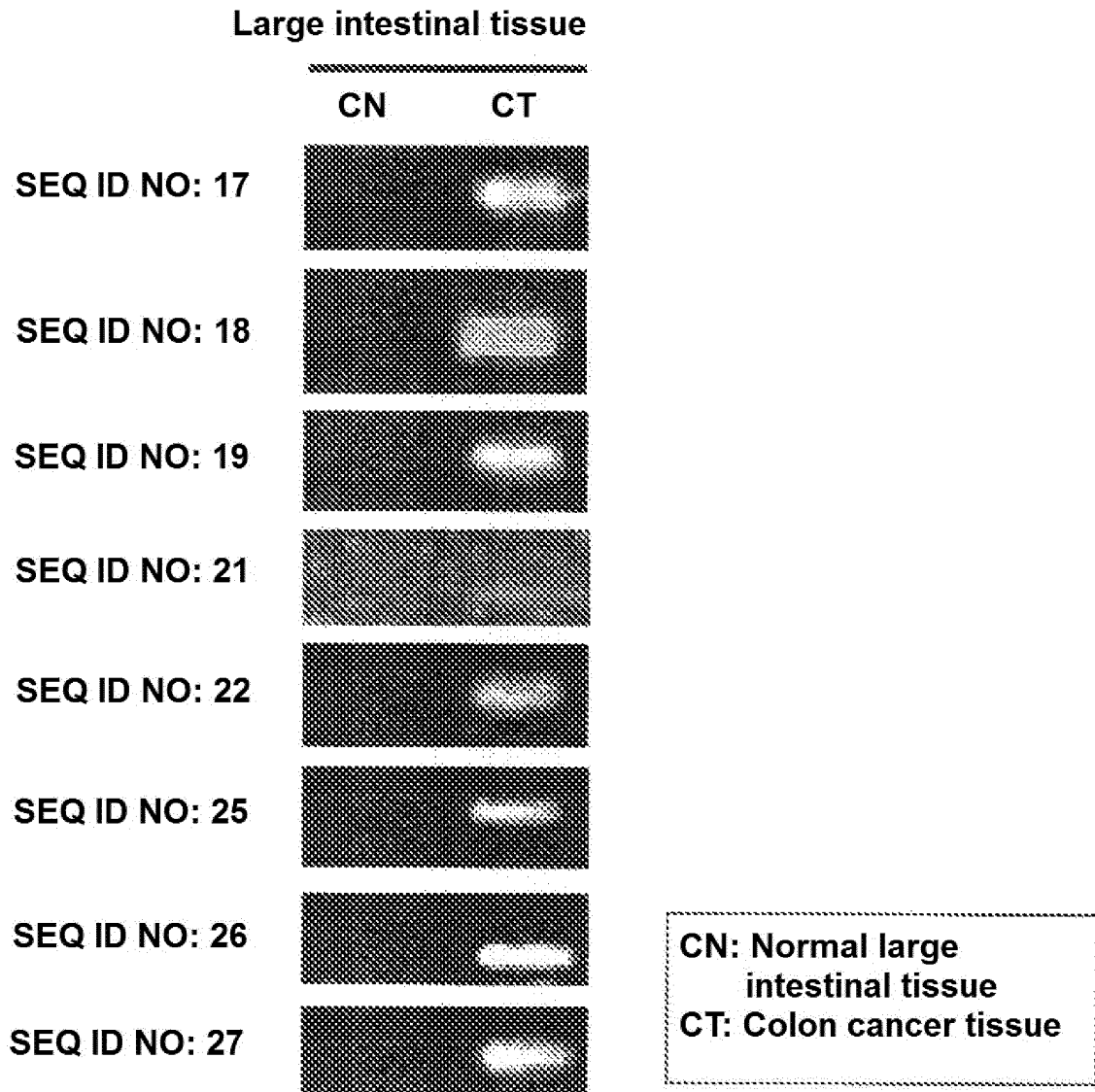


[Fig. 36]

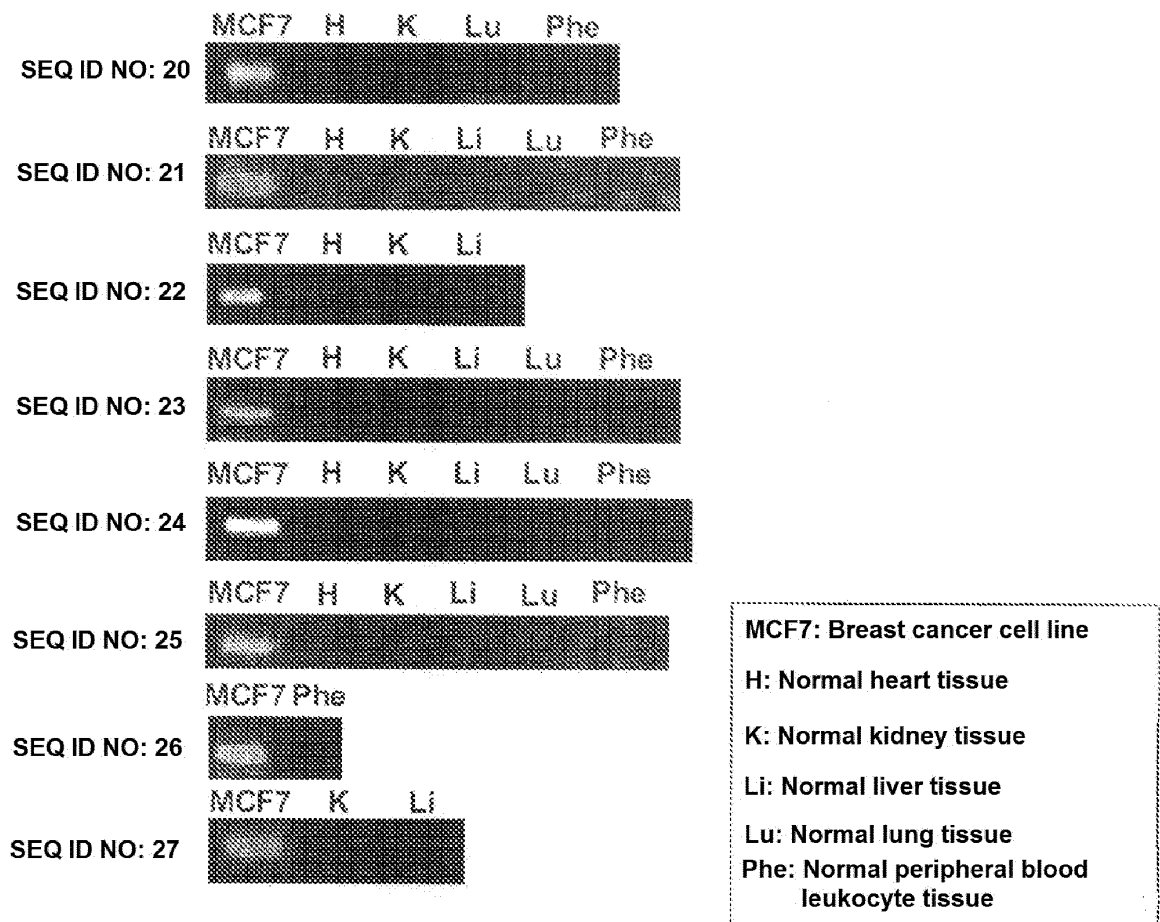
SEQ ID NO: 27



[Fig. 37]



[Fig. 38]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/054363

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> <i>C12Q1/68(2006.01) i, C12N15/09(2006.01) i, G01N33/50(2006.01) i, G01N33/53(2006.01) i</i>  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) <i>C12Q1/68, C12N15/09, G01N33/50, G01N33/53</i>  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) <i>CA/BIOSIS/MEDLINE/WPIDS (STN), JSTPlus (JDreamII), GenBank/EMBL/DDBJ/GeneSeq</i>		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2003/074736 A1 (Institute of Gene and Brain Science), 12 September 2003 (12.09.2003), abstract; claims & EP 1489189 A1 & US 2006/0105330 A1	1-13
X	WU Q., et al., DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. <i>Molecular Cancer</i> , 2007, Vol.6, No.45, pp.1-10	1-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* 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
Date of the actual completion of the international search 25 May, 2011 (25.05.11)		Date of mailing of the international search report 07 June, 2011 (07.06.11)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

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

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/054363

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAUCH T., et al., Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. Proceedings of the National Academy of Sciences of USA, 2007, Vol.104, No.13, pp.5527-5532	1-13
A	RUIKE Y., et al., Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. BMC Genomics, 25 February 2010, Vol.11, No.137, pp.1-11	1-13
A	JACINTO F.V., et al., Methyl-DNA immunoprecipitation (MeDIP): Hunting down the DNA methylome. BioTechniques, 2008, Vol.44, No.1, pp.35, 37, 39, 41, 43	1-13
A	Yoshihiko TAJIMA et al., "Shinki Kaiseki Shuho o Mochiita MeDIP on chip assay no Yuyosei Hyoka to, Nyugan ni Okeru Shinki Methyl-ka Idenshi Marker Kensaku", Dai 68 Kai Proceedings of the Japanese Cancer Association, 2009, vol.68, page 179, #P-0383	1-13

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

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/054363

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

The inventions in claims 1-8 relate to a method for determining the presence or absence of a cancer cell by analyzing the methylated state at the CpG site in SEQ ID NOS:1-14. These inventions have a common technical matter that is "determining the presence or absence of a cancer cell by analyzing the methylated state at the CpG site". However, this technical feature does not make a contribution over prior art in the light of WO2004/08339, WO2006/007830 and WO2003/074736 (document 1). Thus, the aforesaid technical feature cannot be considered as a special technical feature. Among these inventions, there is no other the same or corresponding special technical feature. (Continued to extra sheet.)

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
The parts relating to SEQ ID NO:1 in claims 1-13.

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/054363

Continuation of Box No.III of continuation of first sheet (2)

The same applies to the inventions in claims 9 and 10 relating to a marker for determining the presence or absence of a cancer cell and the inventions in claims 11-13 relating to a kit for determining the presence or absence of a cancer cell.

Thus, claims 1-13 have 13 invention groups as will be shown hereinbelow. Invention group 1: inventions in claims 1-13 relating to SEQ ID NO:1 (HOXA9). Invention group 2: inventions in claims 1-13 relating to SEQ ID NO:2 (KCNQ1DN). Invention group 3: inventions in claims 1-13 relating to SEQ ID NO:3 (RAPSN). Invention group 4: inventions in claims 1-13 relating to SEQ ID NOS:4 and 5 (FZD9). Invention group 5: inventions in claims 1-13 relating to SEQ ID NO:6 (CDH22). Invention group 6: inventions in claims 1-13 relating to SEQ ID NO:7 (LBX2). Invention group 7: inventions in claims 1-13 relating to SEQ ID NO:8 (PAX9). Invention group 8: inventions in claims 1-13 relating to SEQ ID NO:9 (ADD3). Invention group 9: inventions in claims 1-13 relating to SEQ ID NO:10 (CCDC61). Invention group 10: inventions in claims 1-13 relating to SEQ ID NO:11 (ZIC4). Invention group 11: inventions in claims 1-13 relating to SEQ ID NO:12 (CGB7). Invention group 12: inventions in claims 1-13 relating to SEQ ID NO:13 (TOX). Invention group 13: inventions in claims 1-13 relating to SEQ ID NO:14 (TNNI3).

**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 2007524369 W [0006]
- WO 2004083399 A [0006]
- JP 2008506407 W [0006]
- WO 2006008128 A [0006]

专利名称(译)	用于确定生物样品中癌细胞的存在或不存在的方法，以及用于测定的分子标记和试剂盒		
公开(公告)号	<a href="#">EP2540839A4</a>	公开(公告)日	2013-07-31
申请号	EP2011747526	申请日	2011-02-25
[标]申请(专利权)人(译)	希森美康株式会社		
申请(专利权)人(译)	希森美康公司		
当前申请(专利权)人(译)	希森美康公司		
[标]发明人	SAKAI AYAKO TASHIMA YOSHIHIKO KAJITA MASAHIRO		
发明人	SAKAI, AYAKO TASHIMA, YOSHIHIKO KAJITA, MASAHIRO		
IPC分类号	C12Q1/68 C12N15/09 G01N33/50 G01N33/53		
CPC分类号	C12Q1/6886 C12Q2600/154 C12Q2600/156		
优先权	2010042814 2010-02-26 JP		
其他公开文献	EP2540839A1 EP2540839B1		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明涉及一种基于通过分析从生物样品中提取的DNA的甲基化状态获得的分析结果来确定生物样品中癌细胞的存在的方法，所述分析结果允许确定是否存在癌细胞。