DESCRIPTION

# MICROARRAY FOR DETECTION OF MUTATIONS IN β-GLOBIN GENES AND DETECTION METHOD THEREOF

## TECHNICAL FIELD

The present invention relates to a probe group for detecting mutations in β-globin gene, a microarray having the same, and method for detecting a mutation in β-globin gene using the same.

## BACKGROUND ART

The human genome is composed of approximately three billion genetic codes (bases), but it has been found that there exist many differences in the genetic codes (base sequences) between individuals. Currently, among these differences of base sequences, the concern over the studies of single nucleotide polymorphism (SNP) has risen.

Single nucleotide polymorphism (SNP) means that only a single base in the base sequence of DNA is different, and this corresponds to a minimum unit of human personality, such as whether one can hold one's drink or not, or whether one is sensitive to a medicine or not. It has been suggested that among the three billion base pairs of the human genome, there are about 3,000,000 (proportion corresponding to one out of 500 to 1000 base pairs) to 10,000,000 single nucleotide polymorphisms, and these cause blocking of the production of a particular protein or production of a protein that is different from that of other people, thereby bringing about differences such as differences among individuals (physical constitution) or differences among races. It is

believed that the study of individual differences in the genes of the human being enables order-made medicine, by which single nucleotide polymorphisms are analyzed to investigate the susceptibility to a disease or the responsiveness to a drug, and a drug that would cause less adverse side effects in an individual person is administered to the relevant person. Thus, research on the analysis of single nucleotide polymorphism (SNP) is underway.

The reason why single nucleotide polymorphism (SNP) is attracting attention is that an analysis of a large number of SNP's has been made possible by a progress in the nucleic acid analysis technologies, and the correlation between diseases and SNP has been revealed. The correlation with SNP is being disclosed in a wide variety of domains such as disease-related genes, analysis of individual differences in drug metabolism, and chronic diseases. It is also expected that the correlation of these factors with SNP will be further disclosed in the future.

The nucleic acid analysis technologies handle a very large number of samples, and include an enormous number of operations; however, the technology is complicated and time-consuming, and generally, high accuracy is required. Among the nucleic acid analysis technologies, it is known that a DNA chip for SNP detection (a DNA chip is also referred to as a DNA microarray; hereinafter, unless particularly stated otherwise, the terms will be considered to have the same meaning) is effective as a means for detecting a large number of genetic variations rapidly with high accuracy.

A DNA chip is a chip on which nucleic acid probes (probes) are respectively immobilized in defined compartments of a carrier, and usually, a single-stranded DNA or oligonucleotide molecule having a base sequence complementary to the nucleic acid fragment to be detected, is used as the nucleic acid probe.

In a DNA chip for SNP detection, complementary strands of the nucleic acid

fragments corresponding to the mutation sites of test nucleic acid are immobilized as the nucleic probes. The test object sites for mutation usually include one normal type and plural variants, and nucleic acid probes matching any of those are arranged within a plot. Regarding the sample to be tested, a specimen liquid in which only a nucleic acid fragment corresponding to a mutated test object site has been amplified by a nucleic acid amplification method represented by a PCR method, is used.

This specimen liquid is brought into contact with the surface of the DNA chip for SNP detection, on which the nucleic acid probes are immobilized, and the specimen nucleic acid fragment and the nucleic acid probe are hybridized. Then, the binding caused by this hybridization is detected as an optical or electrochemical signal, and thereby, the specimen nucleic acid fragments bound to the nucleic acid probes may be classified and quantitatively determined.

Here, when the combination of the nucleic acid probe and the specimen is a perfect match such as the combination of a wild type probe and a wild type specimen, or the combination of a variant probe and a variant specimen corresponding thereto, the hybridization forms a complete and strong bonding. On the other hand, when the combination of the nucleic acid probe and the specimen is a mismatch such as the combination of a wild type probe and a variant specimen, or the combination of a variant probe and a wild type specimen, since a site that is not capable of hydrogen bonding is inevitably accompanied, the hybridization is incomplete and forms a weak bonding.

Generally, hybridization is carried out under high-stringency conditions that are achieved by various combinations of temperature, a salt, a detergent, a solvent, a chaotropic agent, and a denaturant in order to maintain high specificity, and the difference in the signal intensity originating from the difference in the bonding force of

hybridization between a perfect match and a mismatch is detected. Thereby, the genotype in the specimen may be identified and determined.

Meanwhile, hemoglobin is an iron-containing complex allosteric protein that transports oxygen from the lungs to the cells, and carbon dioxide from the cells to the lungs. Hemoglobin A is a key mature hemoglobin protein, and includes four polypeptide chains (two α-globin chains and two β-globin chains).

Many human diseases are considered to be caused by genetic variations that affect one or more genes encoding the hemoglobin polypeptide chains. Such diseases include sickle cell anemia, and are caused by point mutations in the β-chain of hemoglobin. Furthermore, β-thalassemia symptoms relate to a blood-related disease caused by genetic variation that is significantly expressed in the phenotype by insufficient synthesis of one form of the globin chains, and cause excessive synthesis of the globin chains of the other form (see, for example, Non-Patent Document 1).

On the other hand, the recent development of the molecular biological techniques enables studies on the gene abnormalities causing or associated with the state and symptoms of particular human illness. The polymerase chain reaction (PCR) and many techniques modified therefrom serve as particularly useful tools for the studies on genetic abnormalities in the state and symptoms of illness (see, for example, Non-Patent Documents 2 and 3).

The use of the PCR method amplifies a particular target DNA or a portion of the DNA, and facilitates a new characterization of the amplified portion. Such a new characterization include gel electrophoresis for the determination of size, determination of the nucleotide sequence, studies on hybridization using particular probes, and the like (see, for example, Non-Patent Document 4).

In recent years, extensive studies have been carried out on the causal

relationship between the genotype (that is, genetic polymorphism) such as SNPs (single nucleotide polymorphisms) and diseases and the like, and thus, decisions have been made on whether or not genetic abnormalities exist in the genome of a particular individual.

Regarding a method for determining (detecting) single base mutations such as SNPs, or a genetic variation with a number of bases of 2 or higher, first, a PCR-SSP method is available. Since this technique involves synthesis of primers specific to the base sequence of a mutated gene to perform PCR, several ten primer pairs are needed in order to discriminate several tens of genetic variations.

This method is a convenient method with a short analysis time, but since scrupulous attention and knowledge, and time are required when such primers are designed , there are limitations in the analysis of a large amount of specimens. In addition, as there are more sites for which it is desired to detect mutation, the number of PCR-SSP also increases; therefore, there is a defect that it is difficult to handle a large number of specimens at a single time.

As a second method, a restriction enzyme fragment length polymorphism method (PCR-RFLP method) may be considered. Primers are designed in a consensus sequence site, polymorphisms are included within the PCR product. After the amplification, the amplification product is cut using various restriction enzymes, and mutations of the gene sequence are classified based on the size of the DNA fragments.

This method allows easy determination of the results, and the method is also simple; however, when the sites capable of recognizing the restriction enzyme are limited, discrimination is made difficult. Furthermore, since polyacrylamide gel should be used for the separation of the specimen, it is difficult to classify a large amount of specimen or several tens of mutations simultaneously. It has also been

reported that generally, when a whole blood specimen is directly subjected to PCR amplification and then fragmentized with restriction enzymes, whole blood-derived proteins remain in the amplified specimen, and cutting by restriction enzymes is achieved imperfectly. That is, since proteins bound to a DNA are not separated from the DNA, restriction enzymes cannot bind to the DNA, the cleavage reaction cannot proceed normally, and there is a need for a DNA extraction operation.

A PCR-SSCP method is available as a third method. This method is a method of modifying PCR products into single-stranded DNAs (ssDNAs) by adding a modifying agent such as formamide, and then performing electrophoresis using a non-modified polyacrylamide gel. In regard to the electrophoresis, the ssDNAs respectively assume characteristic structures based on their base sequences, and exhibit intrinsic migration velocities during the electrophoresis, thereby forming bands of respectively different types.

This method is a method of classifying mutations in a base sequence by utilizing the property that ssDNAs exhibit intrinsic migration velocities based on the base sequences; however, a complicated technology with a high degree of difficulty is required, and the analysis of the results also requires experience and knowledge.

As a fourth method, a PCR-SSO (sequence specific oligonucleotide) method may be considered. PCR-SSO is a method of hybridizing synthetic probes for a normal site and a mutation site with PCR products that have been dotted on a filter (a microplate may also be used), and thereby detecting the presence or absence of mutations. On the contrary, there is also a reverse dot blotting method of dotting probes, and hybridizing the PCR products. To compare this method with the antigen-antibody reaction, this is a method in which a DNA serves as an antigen, and an antibody specific to a mutated site and an antibody specific to a normal site are caused

to act as the antibodies has been bound to the DNA. Traditionally, radioactive isotopes have been used for the detection in this method, but due to a restriction on the facilities used and the like, detection is now achieved by chemiluminescence, color development method, or the like.

Although this method is simple, it is necessary to secure a significant amount of a sample, or else, it is necessary to adopt a technique for increasing the sensitivity (see, for example, Non-Patent Documents 5 and 6, and Patent Document 1).

As a fifth method, there is a direct base sequence determination method. The direct base sequence determination method is a method of directly determining a base sequence by using a PCR-amplified DNA strand as a template, without performing subcloning into a vector or the like.

This method performs secondary PCR, which is called asymmetric PCR, of a PCR-amplified DNA strand to amplify a single-stranded DNA, and thereby determines the base sequence generally using a dideoxy method. This secondary PCR performs PCR using one member of a primer pair in a limited amount (usually 1:10 to 1:100), and thereby a single-stranded DNA is amplified. Recently, a cycle sequencing method has been applied so that a sequencing reaction can now be carried out more simply. However, since the price of the kit is very expensive, highly expensive apparatuses are required, and the experimental procedure is also complicated, it is cost-consuming in order to analyze a large amount of a specimen.

CITATION LIST

PATENT DOCUMENT

Patent Document 1: JP 5-184398 A

NON-PATENT DOCUMENT

Non-Patent Document 1: Weatherall et al., The Thalassaemia Syndromes, 3rd Edition, Oxford Blackwell Scientific, 1981

Non-Patent Document 2: Erlich et al., Current Communications in Molecular Biology: Polymerase Chain Reaction, Cold Spring Harbor: Cold Spring Harbor Press (1989)

Non-Patent Document 3: Innis et al., PCR Protocols: A Guide to Methods and Applications. San Diego: Academic Press (1990)

Non-Patent Document 4: Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor: Cold Spring Harbor: Cold Spring Harbor Press (1989)

Non-Patent Document 5: Am. J. Hum. Genet., 43:095-100, 1988

Non-Patent Document 6: Blood, Vol. 81, No. 1 (January 1), 1993: pp. 239-242

## DISCLOSURE OF THE INVENTION

## PROBLEM TO BE SOLVED BY THE INVENTION

As described above, various detection methods are used in order to analyze and detect mutations of genes, but a common methods is that in order to simultaneously detect a large number of mutations, a very long time and enormous efforts are required, and it is even more difficult when it is intended to analyze a large amount of a specimen.

Therefore, a primary object of the present invention is to provide a microarray for detecting mutations in a β-globin gene, which can detect a large number of mutations (specimen) simply and conveniently in a short time.

In view of the problems of the related art, the inventors of the present invention conducted thorough investigations, and as a result, the inventors found that when plural kinds of probes having particular sequences are used, the object described above can be

achieved, thus completing the present invention.

That is, the present invention relates to a probe group for detecting mutations in a β-globin gene containing genes having the sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17, 18, and SEQ ID NOs:25 to 66, a microarray having the probe group immobilized thereon, and a method and a kit for detecting mutations using the microarray.

According to the present invention, since a hybridization solution can be mixed in so as to be brought directly into contact with and react with the microarray without purifying the PCR product, even in a case in which a large amount of a specimen is used, the treatment may be carried out in a short time. Furthermore, since 25 sites of mutation in a β-globin gene may be detected all at once, the present invention is excellent in practical usability and usefulness.

## BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 shows diagrams illustrating the correction method of the invention;

Fig. 2 shows diagrams illustrating the correction method of the invention;

Fig. 3 shows diagrams illustrating the correction method of the invention;

Fig. 4 shows diagrams illustrating the correction method of the invention;

Fig. 5 shows diagrams produced by plotting the results of the hybridization of a first control nucleic acid performed plural times, in a fluorescence coordinate system representing the signal intensities of the first and second probes for polymorphism detection, and presenting representative straight lines thereof;

Fig. 6 shows diagrams produced, in addition to Fig. 5, by plotting the results of the hybridization of a second control nucleic acid performed plural times, and presenting representative straight lines thereof;

Fig. 7 shows graphs for the correction values C and C2;

Fig. 8 shows the probe performance data obtained before and after making corrections using the correction values C and C2, and the angle of error;

Fig. 9 shows diagrams illustrating the data obtained before correction and after correction, from 25 kinds of plasmid-derived samples; and

Fig. 10 is a graph showing the results obtained by superimposing the data of Table 9 on the graph of Fig. 9 for the data obtained after correction.

## BEST MODE(S) FOR CARRYING OUT THE INVENTION

Hereinafter, the present invention will be described in detail. The following embodiment is an exemplary embodiment for explaining the present invention, and is not intended to limit the present invention to this embodiment. The present invention may be carried out in various forms as long as the gist is maintained.

Meanwhile, all publications, patent applications, and patent documents other than patent publications mentioned in this specification are incorporated herein by reference. Also, the present specification includes the subject matters described in the specification and the drawings of Japanese Patent Application (Japanese Patent Application No. 2012-077394) filed March 29, 2012, from which the present application claims priority.

Hereinafter, the present invention will be described in detail. The following embodiment is an exemplary embodiment for explaining the present invention, and is not intended to limit the present invention to this embodiment. The present invention may be carried out in various forms as long as the gist is maintained.

Furthermore, unless particularly stated otherwise, an amino acid sequence is defined to have the amino terminus at the left end and the carboxyl terminus at the right

end, and a base sequence is defined to have the 5'-terminus at the left end and the 3'-terminus at the right end.

1. Probe for polymorphism detection

A microarray is generally used for the detection of a polymorphism, but in order to perform detection with high sensitivity, there is a demand for a high performance probe which does not easily undergo non-specific hybridization. The performance of a probe is generally dependent on the Tm value of the probe (as the Tm value is higher, non-specific hybridization is likely to occur), the Tm value is determined by the sequence of the probe. Therefore, generally, the performance of the probe is constrained by the sequence of the peripheral region of the polymorphism to be detected.

However, the inventors of the present invention succeeded in enhancing the performance of a probe by regulating the Tm value of the probe by applying modification to the sequence of the probe to the extent that the intrinsic performance of the probe is not impaired.

Therefore, the present invention provides a probe as described below.

A probe for detecting a polynucleotide sequence having one or more polymorphisms,

characterized by being hybridized to the relevant sequence, and satisfying at least any one of the following requirements:

(1) the sequence contains one or more non-complementary bases at both ends or at any one end;

(2) the portion corresponding to the polymorphisms that are not targeted for detection, among the plural polymorphisms contained in the sequence, contains

universal bases; and

(3) the polymorphism that is targeted for detection is located at a position six or fewer bases away from any one terminus of the probe.

According to the present invention, the term "probe" means a compound capable of capturing a substance targeted for detection that is included in a specimen, and examples thereof include nucleic acids such as a deoxyribonucleic acid (DNA), a ribonucleic acid (RNA), and a peptide nucleic acid (PNA). These probes may be obtained from commercially available synthetic products such as a DNA synthesized in vitro by an enzyme or the like, and a chemically synthesized oligonucleotide, or from live cells. A DNA fragment that has been chemically modified or cleaved by a restriction enzyme may also be used.

The length of a probe is generally about 10 bp to 100 bp, but the length is preferably 10 bp to 80 bp, more preferably 10 bp to 50 bp, even more preferably 10 bp to 35 bp, and most preferably 12 bp to 28 bp.

Furthermore, the "polynucleotide sequence having one or more polymorphisms", which is the object of detection, means a polynucleotide included in a specimen, having one or more, two or more, or three or more polymorphisms in the base sequence.

The polynucleotide as an object of detection is primarily derived from a human specimen, but as long as the polynucleotide may be subjected to an amplification reaction, a polynucleotide of any biological species may be used.

The specimen may be any of cells, blood or body fluid derived from any tissue. Specific examples of the specimen include cells, blood and body fluid derived from various tissues such as brain, heart, lung, spleen, kidney, liver, pancreas, gall bladder,

esophagus, stomach, intestines, urinary bladder, and skeletal muscles of human being. More specifically, examples include blood, cerebrospinal fluid, urine, sputum, pleural fluid, ascitic fluid, gastric juice, and bullous fluid.

The polynucleotide as the object of detection may be purified before being supplied to the microarray. In regard to the method for purifying a polynucleotide, for example, various technologies according to the descriptions of Maniatis, et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, pp. 280, 281, 1982) may be employed.

Feature (1): one or more non-complementary bases are respectively contained at both ends or at any one end of the polynucleotide sequence containing polymorphisms, which is the object of detection

In general, a "GC-rich" polynucleotide having large contents of guanine (G) and cytosine (C) has a high Tm value and is likely to undergo non-specific hybridization.

Thus, the inventors of the present invention found that in a case in which the polynucleotide used as a probe contains a GC-rich region, a mismatch is caused between the probe and the polynucleotide as the object of detection by incorporating a GC-rich sequence and non-complementary bases at both ends of the same region, and thereby the Tm value of the probe may be decreased.

A "non-complementary base" means any base causing a mismatch with a corresponding base on the polynucleotide sequence as the object of detection. For example, when the corresponding base on the polynucleotide sequence as the object of detection is "C", the non-complementary base may be any of "A", "T" and "C". When the corresponding base on the polynucleotide sequence as the object of detection is "G",

the non-complementary base may be any of "A", "T" and "G". When the corresponding base on the polynucleotide sequence as the object of detection is "A", the non-complementary base may be any of "A", "C" and "G". Furthermore, when the corresponding base on the polynucleotide sequence as the object of detection is "T", the non-complementary base may be any of "T", "C" and "G".

The number of non-complementary bases contained at the two ends may be different between the 5'-terminus and the 3'-terminus.

The "non-complementary base" according to the present invention is incorporated into the two terminal sections of the polynucleotide sequence containing a polymorphism as the object of detection, or into the two terminal sections of the GC-rich sequence containing a polymorphism. For example, when the polynucleotide sequence containing a polymorphism as the object of detection is "GGCGCGGCGCGG" (the underlined part at the center is the polymorphism as the object of detection), the probe having the feature (1) may be "TGCGCGGCGCGA", may be "ATGCGCGGCGCGAA", or may be "ATGGCGCGGCGCGGAA".

Preferably, the "non-complementary base" includes one or more bases, two or more bases, or three or more bases, at either terminal section.

According to the present invention, the GC-rich region means a sequence region in which the content of G and C contained in the entire base sequence is 50% or more, 55% or more, 60% or more, 61% or more, 62% or more, 63% or more, 64% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, or 95% or more.

Feature (2): the portion corresponding to polymorphisms that are not intended

for detection, among the plural polymorphisms contained in the polynucleotide sequence containing polymorphisms as the objects of detection, contains universal bases.

A probe having this feature is useful when a second polymorphism which is a non-object of detection is contained, in addition to a first polymorphism which is the object of detection, in the polynucleotide sequence as the object of detection.

In this case, since the binding force between the probe and the polynucleotide as the object of detection is changed by the combination for the second polymorphism, in addition to the combination for the first polymorphism, the sensitivity of detection to the first polymorphism as the original object of detection is decreased.

When this polynucleotide sequence is used directly as the sequence of the probe, the polynucleotide as the object of detection in the first polymorphism portion matches the probe; however, there may occur an occasion in which the polynucleotide and the probe do not match (mismatching) in the second polymorphism portion, and an occasion in which the polynucleotide as the object of detection and the probe do not match (mismatching) in the first polymorphism portion, while the polynucleotide and the probe match in the second polymorphism portion. In the latter case, despite the first polymorphism does not match, since the probe and the polynucleotide as the object of detection are hybridized with a binding force at the same level as that of the former case, positive signals similar to those of the former case are emitted.

Therefore, in regard to the above-described cases, the former (true positive) and the latter (false positive) cannot be distinguished.

Thus, in order to nullify the influence of the second polymorphism, the probe of the present invention is characterized by containing universal bases in the portion corresponding to the second polymorphism.

According to the present invention, a universal base means a base which does not form a base pair with any of naturally occurring nucleic acid bases, namely, adenine, guanine, thymine, cytosine, and uracil.

Examples of such a universal base include, but are not limited to, 5-nitroindole, 3-nitropyrrole, 7-azaindole, 6-methyl-7-azaindole, pyrrolepyridine, imidazopyridine, isocarbostyryl, propynyl-7-azaindole, propynylisocarbostyryl, and allenyl-7-azaindole.

Other examples of the universal base include any one or more of the following compounds, including propynyl derivatives thereof:

8-aza-7-deaza-2'-deoxyguanosine, 8-aza-7-deaza-2'-dioxyadenosine, 2'-deoxycytidine, 2'-deoxyuridine, 2'-deoxyadenosine, 2'-deoxyguanosine, and pyrrolo[2,3-d]pyrimidine nucleotide.

Furthermore, the universal base may be formed from any of the following compounds, including derivatives thereof:

Deoxyinosine (for example, 2'-deoxyinosine), 7-deaza-2'-deoxyinosine, 2'-aza-2'-deoxyinosine, 3'-nitroazole, 4'-nitroindole, 5'-nitroindole, 6'-nitroindole, 4-nitrobenzimidazole, nitroindazole (for example, 5'-nitroindazole), 4-aminobenzimidazole, imidazo-4,5-dicarboxamide, 3'-nitroimidazole, imidazole-4-carboxamide, 3-(4-nitroazol-1-yl)-1,2-propanediol, and 8-aza-7-deazaadenine (pyrazolo[3,4-d]pyrimidine-4-amine).

According to another example, regarding the universal nucleic acid base, a universal nucleic acid base may be formed by combining a 3-methyl-7-propynylisocarbostyryl group, a 3-methylisocarbostyryl group, a 5-methylisocarbostyryl group, an isocarbostyryl group, a phenyl group or a pyrenyl group, with ribose or deoxyribose.

Feature (3): the polymorphism intended for detection is located at a position six or fewer bases away from any one terminus of the probe.

Furthermore, the probe of the present invention may also be designed such that the polymorphism intended for detection is located at a position six or fewer bases away from any one terminus (5'-terminus or 3'-terminus) of the probe.

A probe designed as such is useful in view of the following point.

Generally, when a probe for the detection of single nucleotide polymorphism is designed, the probe is designed so as to have a base length that is approximately equal on both the 5'-terminus side and the 3'-terminus side while centering the position of the polymorphism intended for detection. However, if the 5'-terminus side or the 3'-terminus side at the site of the polymorphism intended for detection is an extremely GC-rich region or an extremely AT-rich region, binding to the probe may occur, or may not occur, at the position of the polymorphism intended for detection, regardless of being a match or a mismatch.

Thus, the present invention is characterized by using a probe in which the position of the polymorphism intended for detection is set to a position six or fewer bases away from a terminus, and thus a GC-rich region or an AT-rich region is avoided. In a case in which the specificity of the probe cannot be enhanced by employing this feature only, the feature (1) may be employed in combination.

There are no particular limitations on the gene that serves as a basis of the polynucleotide of the present invention, and examples include G6PD gene, RAB27A gene, CHS1 gene, MTHFR gene, HMGCL gene, SLC2A1 gene, and H6PD gene.

In addition, in order to obtain information individually, access may be made to the OMIM Database (http://www.ncbi.nlm.nih.gov/omim), where the information on a disease, and causes thereof or genes that serve as risk factors may be obtained.

According to an embodiment, the polynucleotide of the present invention is prepared from the human β-globin gene.

According to the present invention, the polynucleotide sequence having the polymorphism intended for detection has a sum of the contents of guanine and cytosine of 63% or more (GC-rich), and has nucleotide sequences represented by from 99$^{th}$ to 117$^{th}$ nucleotides, from 127$^{th}$ to 142$^{nd}$ nucleotides, and from 1402$^{nd}$ to 1416$^{th}$ nucleotides of the human β-globin gene.

Alternatively, the polynucleotide sequence having the polymorphism intended for detection has a sum of the contents of guanine and cytosine of 45% or less (GC-poor), and has nucleotide sequences represented by from 1378$^{th}$ to 1399$^{th}$ nucleotides of the human β-globin gene.

The above-described region is a GC-rich region or an AT-rich region, and provides a probe capable of detecting a polymorphism in this site.

More specifically, the probe of the present invention is a probe specialized by a GC-rich region, and has a sequence set forth in SEQ ID NO:3, 4, 7, 8, 17 or 18.

On the other hand, the probe of the present invention is a probe specialized by a GC-poor region, and has a sequence set forth in SEQ ID NO:11 or 12.

Furthermore, the present invention provides a microarray having at least one of sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17 and 18.

2. Probe group

According to the present invention, sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17 and 18 and SEQ ID NOs:25 to 66 are used as probes (probe group of the present invention), and if necessary, genes other than the probe group of the present invention may also be used as probes.

3. Microarray

(1) Support

In order to actually put the above-described probe group to use, it is necessary to immobilize the probes to a support. There are no limitations on the kind of the support for immobilization, and any support that does not allow a probe to be eluted (released) into the reaction liquid at the time of a hybridization reaction, and enables characterization of which probe has reacted after the reaction, may be used.

Examples include a filter, beads, a gel, a chip, a slide glass, a multi-well plate, a membrane, and an optical fiber. More specifically, examples include a Western Blotting filter paper, a nylon membrane, a membrane made of polyvinylidene fluoride, a nitrocellulose membrane (Pierce Biotechnology, Inc.), affinity beads (Sumitomo Bakelite Co., Ltd.), MicroPlex (registered trademark) Microspheres, xMAP Multi Analyte LumAvidin Microspheres (Luminex Corp.), Dynabeads (Veritas Corp.), a 96-well plate kit for DNA immobilization (Funakoshi Co., Ltd.), a substrate for DNA immobilization (Sumitomo Bakelite Co., Ltd.), a coated slide glass for microarray (Matsunami Glass Industry, Ltd.), a hydrogel slide (PerkinElmer, Inc.), and Sentrix (registered trademark) Array Matrix (Illumina, Inc.).

(2) Immobilization

Immobilization of a probe may be carried out, in the case of using a filter, a membrane or the like, by directly spotting an unmodified probe, and irradiating the probe with a UV lamp or the like. Furthermore, in the case of using beads, a chip, a slide glass, a multi-well plate, a membrane, an optical fiber and the like, which have their surfaces chemically activated, it is preferable to use a probe having a terminus that may form chemically covalent bonding. More specifically, a probe having an amino group or the like introduced to the 5'-terminus or the 3'-terminus is used. Furthermore, in the case of immobilizing a probe onto a gel or the like, a probe having an unsaturated functional group that is capable of copolymerization reaction is used. When the probe has this introduced group, the probe is immobilized to the network structure of the gel by a copolymerization reaction with a substituted (meth)acrylamide derivative or an agarose derivative, and a crosslinking agent. In regard to the method of introducing an unsaturated functional group into the terminus of a nucleic acid strand, for example, the known method described in WO 02/062817 may be used.

In the present invention, it is preferable to immobilize the probe onto a gel or within a gel. It is because when the probe is immobilized within a gel, since the amount of the probe may be increased, the detection sensitivity of the microarray may be increased. Furthermore, in the present invention, it is preferable to maintain the gel inside through-holes, and to use a through-hole type microarray having a plural number of the relevant through-holes.

A through-hole type microarray may be obtained by forming through-holes on a foil plate, but a microarray obtainable by retaining gel carriers having probes immobilized therein, in the hollow sections of tubular bodies such as hollow fibers such

that different kinds of the gel carriers are retained in different tubular bodies, gathering and fixing all the tubular bodies such as hollow fibers, and then repeatedly cutting the tubular bodies along the longitudinal direction of the fibers, is preferred. It is because microarrays of stable quality may be produced in large quantities. In this manner, a microarray in which respective probes are immobilized within various through-holes in an independent manner (state in which a probe of one kind is immobilized within one through-hole), may be obtained.

Hereinafter, an embodiment of the method for producing a through-hole type microarray will be explained. The relevant microarray may be produced through the steps of (i) to (iv) described below.

Step (i): Step of arranging plural lines of hollow fibers three-dimensionally such that the fiber axes of the various hollow fibers will be in the same direction, fixing the arrangement with a resin, and thereby producing a hollow fiber bundle

The method for forming through-holes is not particularly limited, and for example, a method of producing an arranged body in which hollow fibers are arranged in the same axial direction, and then fastening the arranged body with a resin, as described in JP 2001-133453 A may be utilized. Regarding the hollow fibers, various materials may be used, but an organic material is preferred.

Examples of a hollow fiber formed from an organic material include polyamide-based hollow fibers of nylon 6, nylon 66, aromatic polyamide, and the like; polyester-based hollow fibers of polyethylene terephthalate, polybutylene terephthalate, polylactic acid, polyglycolic acid, polycarbonate, and the like; acrylic hollow fibers of

polyacrylonitrile, and the like; polyolefin-based hollow fibers of polyethylene, polypropylene, and the like; polymethacrylate-based hollow fibers of polymethyl methacrylate and the like; polyvinyl alcohol-based hollow fibers; polyvinylidene chloride-based hollow fibers; polyvinyl chloride-based hollow fibers; polyurethane-based hollow fibers; phenolic hollow fibers; fluorine-based hollow fibers formed from polyvinylidene fluoride, polytetrafluoroethylene, and the like; and polyalkylene para-oxybenzoate-based hollow fibers. The hollow fibers may be porous, and may be obtained by combining a melt spinning method or a solution spinning method with known porosification technologies such as a stretching method, a microphase separation method, and an extraction method. The porosity is not particularly limited, but from the viewpoint of increasing the density of the probes to be immobilized per unit length of the fiber material, a higher porosity is preferred as the specific surface area increases. The inner diameter of the hollow fiber may be arbitrarily set. The inner diameter may be adjusted preferably to 10 μm to 2000 μm, and more preferably 150 μm to 1000 μm.

The method for producing the relevant hollow fiber is not limited, and the hollow fiber may be produced by a known method such as described in JP 11-108928 A. For example, a melt spinning method is preferred, and regarding the nozzle, a horseshoe-shaped nozzle, a C-shaped nozzle, a double pipe nozzle, or the like may be used. According to the present invention, it is preferable to use a double pipe nozzle from the viewpoint that a continuous and uniform hollow section may be formed.

Furthermore, if necessary, a hollow fiber in which a black pigment such as carbon black has been incorporated in an appropriate amount, may also be used. When the hollow fiber contains a black pigment, optical noises originating from foreign

materials such as impurities may be reduced at the time of detection, or the strength of the resin may be increased. The content of the pigment is not limited, and the content may be appropriately selected according to the size of the hollow fiber, the purpose of use of the microarray, and the like. For example, the content may be adjusted to 0.1% to 10% by mass, preferably 0.5% to 5% by mass, and more preferably 1% to 3% by mass.

Production of a block body may be carried out using a method of fixing the block body with a resin such as an adhesive so that the arrangement of the arranged body would not be disrupted. For example, there may be mentioned a method of arranging plural lines of hollow fibers in parallel at a predetermined interval on a sheet-like object such as an adhesive sheet, fabricating the assembly into a sheet form, and then winding this sheet into a helical form (see JP 11-108928 A).

Another method may be a method of superimposing two sheets of porous plates each having plural holes provided at a predetermined interval, such that the respective hole areas of the plates would coincide, passing hollow fibers through those hole areas, opening a gap between the two sheets of porous plates, filling a curable resin raw material around the hollow fibers between the two sheets of porous plates, and curing the resin raw material (JP 2001-133453 A).

The curable resin raw material is preferably formed from an organic material such as a polyurethane resin or an epoxy resin. Specifically, the curable resin raw material is preferably formed from one or more kinds of materials consisting of organic polymers and the like. Examples of an organic polymer include rubber materials such

as polyurethane, a silicone resin, and an epoxy resin; polyamide-based resins such as nylon 6, nylon 66, and an aromatic polyamide; polyester-based resins such as polyethylene terephthalate, polybutylene terephthalate, polylactic acid, polyglycolic acid, and polycarbonate; acrylic resins such as polyacrylonitrile; polyolefin-based resins such as polyethylene and polypropylene; polymethacrylate-based resins such as polymethyl methacrylate; polyvinyl alcohol-based resins; polyvinylidene chloride-based resins; polyvinyl chloride-based resins; phenolic resins, fluorine-based resins such as polyvinylidene fluoride and polytetrafluoroethylene; and polyalkylene para-oxybenzoate-based resins. In the organic polymer, a black pigment such as carbon black may be incorporated in an appropriate amount. When a black pigment is added, optical noises originating from foreign materials such as impurities may be reduced at the time of detection, or the strength of the resin may be increased. The content of the pigment is not limited, and the content may be appropriately selected according to the size of the hollow fiber, the purpose of use of the microarray, and the like. For example, the content may be adjusted to 0.1% to 10% by mass, preferably 0.5% to 5% by mass, and more preferably 1% to 3% by mass.

The number of the hollow fibers that are arranged in the present invention, that is, the number of spots, is not limited and may be appropriately selected according to the intended experiment or the like. Therefore, the distance between the hollow fibers may also be appropriately selected according to the area of the microarray, the number of the hollow fibers to be arranged and the like.

Step (ii): Step of introducing a gel precursor solution containing a probe group into the hollow section of each hollow fiber of the hollow fiber bundle

The kind of the gel material that is filled in the hollow fibers is not particularly limited, and polysaccharides such as agarose and sodium alginate; and proteins such as gelatin and polylysine may be used as long as the gel material is a gel material obtainable from natural products. Regarding synthetic polymers, for example, a gel obtainable by allowing a polymer having a reactive functional group such as polyacryloylsuccinimide, to react with a crosslinking agent having reactivity with the polymer, may be utilized. In addition, preferred examples also include synthetic polymer gels obtainable by using polymerizable monomers such as acrylamide, N,N-dimethylacrylamide, N-isopropylacrylamide, N-acryloylaminoethoxyethanol, N-acryloylaminopropanol, N-methylolacrylamide, N-vinylpyrrolidone, hydroxyethyl methacrylate, (meth)acrylic acid and allyl dextrin as monomers, and copolymerizing the monomers with polyfunctional monomers, for example, methylenebis(meth)acrylamide and polyethylene glycol di(meth)acrylate.

The concentration of the gel used in the microarray of the present invention is not particularly limited, and the concentration may be appropriately selected according to the length or amount of the probe used. For example, the concentration n terms of the concentration of the monomer component, is preferably 2% to 10% by mass, more preferably 3% to 7% by mass, and even more preferably 3.5% to 5% by mass. The concentration is adjusted to 2% by mass or more because the probes may be securely immobilized so that detection of the target substance may be carried out with high efficiency. Furthermore, the concentration is adjusted to 10% by mass or less because even though the concentration is made higher than that, it may be difficult to obtain a dramatically improved effect.

In the case of retaining a synthetic polymer gel in the microarray of through-hole substrates described above, the synthetic polymer gel may be retained by filling a gel precursor solution of the synthetic polymer in the above-described block, and then gelating the gel precursor solution within the block. Regarding the method of filling a gel precursor solution inside the through-holes of the block, for example, the solution may be introduced by suctioning the solution into a syringe having a fine needle, and inserting the needle into the hollow section of each hollow fiber.

Furthermore, the hollow section of the fixed end of the hollow fiber bundle is sealed, and the hollow section of the other non-fixed end is left open. Next, a gel precursor solution containing a nucleic acid probe having a polymerization reaction point such as a methacryl group at a terminus is prepared, the gel precursor solution and the hollow fiber bundle are placed in a desiccator, subsequently the end of the hollow fiber bundle at which the hollow fibers are not fixed is immersed in this solution, the interior of the desiccators is brought to a state under reduced pressure, and then the pressure is returned to normal pressure. Thereby, this solution may be introduced into the hollow section of the hollow fibers through the ends of the hollow fibers immersed in the solution.

Step (iii): Step of causing the gel precursor solution that has been introduced into the hollow section of the hollow fiber bundle, to react, and thereby maintaining a gel-like object containing probes in the hollow section of the hollow fibers

By polymerizing the gel precursor solution that has been introduced into the hollow section of the hollow fibers, a gel-like object containing probes is retained in the hollow section of the hollow fibers. The conditions for polymerization are not particularly limited, and may be appropriately selected depending on the kind of the gel precursor used, or the like. For example, an acrylamide-based monomer may be polymerized using a radical initiator, and preferably, an acrylamide-based monomer may be polymerized by a thermal polymerization reaction using an azo-based initiator.

The kind and size of the probe are not limited, and may be appropriately selected according to the kind of the substance or compound that is the object of detection.

Step (iv): Step of cutting the hollow fiber bundle in a direction perpendicular to the longitudinal direction of the fibers, and thereby slicing the hollow fiber bundle

The method for cutting is not limited as long as slices may be obtained. For example, the cutting may be carried out using a microtome, a laser, or the like. The thickness of the slice thus obtainable is not limited, and may be appropriately selected according to the purpose of the experiment or the like. For example, the thickness may be adjusted to 5 mm or less, and preferably to 0.1 mm to 1 mm.

(3) Detection of mutation in β-globin gene

According to the present invention, detecting mutations in the β-globin gene means characterizing the base site having a mutated portion in the β-globin gene sequence, and the sequence, and it also means that it is determined which pair of alleles (diploid organism) has the mutation from among plural specified alleles.

(a) First, it is desirable to bring a specimen containing a human genome DNA into contact with a reaction solution containing a primer set for nucleic acid amplification, a nucleotide unit, and a DNA elongation enzyme.

<Specimen (nucleic acid serving as template of nucleic acid amplification)>

A specimen refers to a nucleic acid containing the gene sequence targeted for detection in the present invention, that is, the β-globin gene sequence. Any form of nucleic acid may be used as long as it contains a fragment of the β-globin gene sequence, and is capable of undergoing the amplification reaction described below.

The specimen is human-derived, and any material capable of an amplification reaction may be used. For the specimen, cells, blood or body fluid derived from any tissue may be used. Examples include cells, blood and body fluid derived from various tissues such as brain, heart, lung, spleen, kidney, liver, pancreas, gall bladder, esophagus, stomach, intestines, urinary bladder, and skeletal muscles. More specifically, examples include blood, cerebrospinal fluid, urine, sputum, pleural fluid, ascitic fluid, gastric juice, and bullous fluid.

Furthermore, it is preferable to prepare and purify the specimen as a DNA-containing sample that may be used for the nucleic acid amplification that will be described below, before the relevant amplification is carried out. This preparation and purification may be carried out according to a known nucleic acid extraction method, and for example, various technologies according to the descriptions of Maniatis, et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, pp. 280, 281, 1982) may be used.

<Primer set for nucleic acid amplification and probe position>

The present invention relates to a probe set for detecting mutations in the β-globin gene, and detects mutations in the β-globin gene sequence encoded on a complementary strand sequence of NCBI Reference Sequence: NC_000011.9 (sequence

length 135006516 bases).

Usually, it is preferable to have the site of mutation in the nucleic acid to be detected, from the viewpoint of detection. The primer pair may be designed to include any region, as long as an amplification product containing a site of mutation to be detected is produced. For example, when it is intended to detect plural sites of mutation that are contained in exons 1 and 2 all at once, nucleic acid amplification may be carried out using a forward primer on the upstream of exon 1 and a reverse primer on the downstream of exon 2, and the amplification products may be detected. Furthermore, when it is intended to detect plural regions at the same time, plural primer pairs may be used. In this case, it is preferable to check whether or not non-specific nucleic acid fragments have been produced in the stage of performing amplification, at the stage of setting the conditions. Once the conditions are determined, detection may be carried out with high reproducibility under those conditions.

For the base sequence of the oligonucleotide that serves as a probe, the probe sequence is determined so as to be included in the amplification product sequence (region sandwiched between the primers of the primer set). The length of the probe is usually set to be about 15 to 35 nucleotides, and for one mutated region to be detected, usually a pair (two kinds, namely, a probe for wild type detection and a probe for mutant detection) or more (one or more pairs) of probes are required.

For the purpose of facilitating the subsequent detection, the primer set to be used may have the termini labeled in advance with a fluorescent substance (Cy3, Cy5 or the like), biotin or the like. There are no particular limitations on the method of labeling,

and any method may be used as long as the method does not exhibit any phenomenon against the amplification reaction, such as marked inhibition of the reaction. After the reaction, color development may also be induced by further causing the reaction system to react with a complex with an enzyme, for example, streptavidin-alkaline phosphatase conjugate, and adding a substrate thereto.

<Amplification reaction>

The nucleic acid (specimen) that serves as the template for nucleic acid amplification is used as a template, and a nucleic acid fragment of the region to be detected from the site of mutation in the β-globin gene is amplified. Regarding the method for amplification of nucleic acid, various methods such as a PCR method, a LAMP method, an ICAN method, a TRC method, a NASBA method, and a PALSAR method may be used. Any method may be used for the nucleic acid amplification reaction as long as there is no particular problem in view of the detection of nucleic acid, but among these, a PCR method is preferred from the viewpoint of convenience.

For a temperature controlling apparatus that is used in the nucleic acid amplification reaction, a commercially available thermal cycler may be used. For example, GeneAmp 9600 or GeneAmp 9700 (Life Technologies Japan, Ltd.), or T Professional Series (Biometra GmbH), may be used, but an apparatus of any format may be used as long as there is no problem with thermal conductivity and the shape of the lid.

<Nucleotide unit used in amplification reaction>

An example of the nucleotide unit may be deoxyribonucleotide triphosphate or

the like, which is used in conventional amplification reactions. As for this, a derivative that facilitates detection later may be used as in the case of the primer set; however, it is preferable to use a nucleotide unit that does not inhibit the amplification reaction.

<DNA elongation enzyme used in amplification reaction, and others>

Regarding the DNA elongation enzyme, TaqDNA polymerase, TthDNA polymerase, PfuDNA polymerase and the like, which are DNA polymerases derived from heat resistant bacteria, may be used in the same manner as in the case of being used in a conventional PCR method.

Examples of an enzyme or a kit that may be used include Hot StarTaq DNA Polymerase (manufactured by Qiagen Corp.), PrimeStarMax DNA polymerase (Takara Bio, Inc.), SpeedSTAR HS DNA polymerase (Takara Bio, Inc.), KOD Plus Neo (Toyobo Co., Ltd.), KAPA2G FastHotStart PCR Kit (Nippon Genetics Co., Ltd.), and AmDirect kit (Shimadzu Corp.). In addition to these, any enzyme or kit capable of performing a nucleic acid amplification reaction directly from blood (body fluid or the like), is more preferred since the operation is made simple.

Regarding the method of mixing the constituent elements described above, any mixing method may be used as long as the enzyme is not deactivated, and mixing is achieved without causing the liquid to foam and leak from the tube. Usually, all of the constituent elements described above are dispensed in a tube for PCR use having a size of about 0.2 mL, the mixture is mixed using a vortex mixer, and then the mixture may be lightly centrifuged (spin-down) in order to cause the solution adhering to the lid to fall off. Furthermore, in the case of performing HotStart PCR, mixing is performed under the conditions in which the enzyme is not activated at the time of mixing.

Next, (b) the reaction liquid obtained in (a) may be subjected to a nucleic acid amplification reaction.

In regard to the nucleic acid amplification reaction, for example, in the case of performing amplification by a PCR reaction, an enzyme which dissociates the nucleic acid that serves as a template is activated at 90°C to 98°C for about 5 minutes, subsequently a cycle of 30 seconds at 94°C (dissociation of nucleic acid), 30 seconds at 60°C (annealing of primers), and 30 seconds at 72°C (elongation reaction from primers) is repeated 25 to 50 times, and thereby the nucleic acid may be amplified logarithmically. Furthermore, in the case of performing a nucleic acid amplification reaction under an isothermal condition instead of a PCR reaction, amplified nucleic acid may be obtained by incubating the reaction system at a constant temperature at about 40°C to 65°C.

(c) Subsequently, the nucleic acid fragment obtained in (b) is brought into contact with the microarray of the present invention, and thereby the targeted nucleic acid in the specimen may be detected.

In Step (c), the nucleic acid amplification reaction liquid may be brought into contact with the microarray by directly adding a hybridization solution, without purifying the nucleic acid amplification reaction liquid. A hybridization solution is a solution which enables the amplified nucleic acid to undergo a hybridization reaction with the probe immobilized in the microarray.

More specifically, the hybridization solution is a solution obtained by mixing a solution mixed with a salt, such as a NaCl solution or a $MgCl_2$ solution, a SSC solution, and a surfactant such as SDS or Tween 20, into a buffer solution such as Tris/HCl buffer.

In the case of performing a reaction with the probe used in the present invention, it is preferable to use TNT buffer (mixed solution of a Tris/HCl buffer solution, a NaCl solution, and a Tween solution), in which crystals of a surfactant, such as SDS, are not precipitated at the time of cooling.

The concentration of Tris/HCl or NaCl is preferably 0.06 M to 0.48 M, and more preferably 0.12 M to 0.24 M, as the final concentration of each. The final concentration of Tween may be adjusted to 0.01% to 0.2% by mass, preferably 0.02% to 0.15% by mass, and more preferably 0.03% to 0.12% by mass.

Furthermore, the temperature at the time of contact is preferably 45°C to 65°C, more preferably 50°C to 55°C, and is preferably 45°C to 70°C, and still more preferably 50°C to 65°C. The time for contacting is not limited as long as a hybridization reaction occurs, and mutation can be detected; however, a shorter time is preferred in order to suppress a non-specific reaction. For example, the contact time may be adjusted to 15 minutes to 4 hours, preferably 20 minutes to 3 hours, and more preferably 30 minutes to 2 hours.

<Detection of nucleic acid (amplification product)>

The nucleic acid captured by the probes in the microarray is detected by the above-described step (c).

The method for detection is not limited as long as the captured nucleic acid is detected, and any known method may be used. For example, a method of performing color development analysis or fluorescence intensity analysis using a fluorescent material or a luminescent material as a label substrate; or a method based on visual inspection may be used.

More specifically, determination of the presence or absence and quantitative determination of the captured nucleic acid may be carried out using a fluoroimaging analyzer, a CCD camera or the like. Quantitative determination of nucleic acid with higher reliability can be achieved by monitoring the amount of fluorescence over time using a quantitative real-time PCR analyzer that is being frequently used in recent years.

Furthermore, color development method may also be carried out using a color developing reagent that does or does not utilize an enzymatic reaction, or the like. Such a method may involve direct observation by visual inspection, or scanning with an optical scanner.

The method for detecting a nucleic acid of the present invention may be applied to an analysis of 30 sites of mutation in the β-globin gene as disclosed in the Sequence Listing, but a detection kit for sites other than the sites of mutation disclosed in the present invention can also be produced by designing and using appropriate probes.

(4) Kit

According to the present invention, a microarray having a primer set and the

probe group of the present invention may also be used as a kit for detecting mutations in the β-globin gene. Regarding the primer set, a set of an oligonucleotide primer having the sequence set forth in SEQ ID NO:21 and an oligonucleotide primer having the sequence set forth in SEQ ID NO:22; or a set of an oligonucleotide primer having the sequence set forth in SEQ ID NO:23 and an oligonucleotide primer having the sequence set forth in SEQ ID NO:24 may be more suitably used.

4. Method for evaluating microarray probe

As discussed previously, when detection of polymorphism is carried out using a microarray, it is preferable that the probe used does not cause non-specific hybridization. That is, a probe that does not cause non-specific hybridization is evaluated to have high performance.

Thus, the present invention provides the following method as a method for quantitatively evaluating the performance of a probe.

A method for evaluating a microarray probe, the method including the following steps:

(1) a step of plotting the fluorescence coordinates obtained by hybridizing a control nucleic acid for first polymorphism with a probe pair for polymorphism detection consisting of a probe for first polymorphism detection and a probe for second polymorphism detection, in a fluorescence coordinate system which includes a Y-axis representing the signal intensity obtainable when the probe for first polymorphism detection is hybridized, and an X-axis representing the signal intensity obtainable when the probe for second polymorphism detection is hybridized;

(2) a step of defining a value which is inversely proportional to the gradient of a straight line that passes through the intersection O between the Y-axis and the X-axis

and the fluorescence coordinates plotted in the step (1), as a correction value C; and

(3) a step of carrying out steps (1) and (2) on plural probe pairs for polymorphism detection, comparing the correction values C between the various probes, and determining a probe pair having the minimum correction value C as probes appropriate for first polymorphism detection.

According to the present invention, the first polymorphism and the second polymorphism are different alleles for a same polymorphism. That is, the first polymorphism is a first allele, and the second polymorphism is a second allele corresponding to the first allele.

Hereinafter, a summary of the various steps will be described.

Step (1): Plotting step

First, in step (1), the signal intensities obtainable when a control nucleic acid for first polymorphism is hybridized to a probe pair for polymorphism detection consisting of a probe for first polymorphism detection and a probe for second polymorphism detection, are plotted in a fluorescence coordinate system. In regard to the fluorescence coordinate system of the present invention, the Y-axis represents the signal intensity obtainable when the probe for first polymorphism detection is hybridized, and the X-axis represents the signal intensity obtainable when the probe for second polymorphism detection is hybridized. Here, the intersection between the Y-axis and the X-axis is designated as O. Furthermore, the Y-axis and the X-axis may perpendicularly intersecting each other, or may not perpendicularly intersecting each other.

Through the plotting process described above, fluorescence coordinates $P(x_1,y_1)$

representing the fluorescence characteristics of the probe for first polymorphism detection are obtained.

The fluorescence coordinates P of an ideal probe that does not cause non-specific hybridization, are such that $x_1 = 0$ and $y_1 > 0$ (Fig. 1: Panel A).

However, in reality, many probes cause non-specific hybridization to a certain extent. Therefore, the fluorescence coordinates P of many probes are such that $x_1 > 0$ and $y_1 > 0$ (Fig. 1: Panel B).

Hybridization is achieved in a hybridization solution. A hybridization solution is a solution which enables a hybridization reaction between a control nucleic acid and a probe, but more specifically, a hybridization solution is a solution obtained by mixing a solution mixed with a salt, such as a NaCl solution or a $MgCl_2$ solution, or a SSC solution, and a surfactant such as SDS or Tween 20, with a buffer solution such as Tris/HCl buffer. Generally, when a reaction is carried out, it is preferable to use TNT buffer (mixed solution of a Tris/HCl buffer solution, a NaCl solution, and a Tween solution), in which crystals of a surfactant such as SDS are not precipitated at the time of cooling. The final concentration of the hybridization solution is preferably 0.06 M to 0.48 M, and more preferably, the final concentration is 0.12 M to 0.24 M. Furthermore, the temperature at the time of contact is preferably 45°C to 70°C, and more preferably 50°C to 65°C. Regarding the contact time, a shorter contact time is more preferred, as long as a hybridization reaction occurs, and detection can be made. The contact time is usually 15 minutes to 4 hours, preferably 20 minutes to 3 hours, and more preferably 30 minutes to 2 hours.

A signal is a value obtained by digitizing the amount of control nucleic acids captured by the probes as a result of the hybridization described above. In general, the signal may be obtained by causing a fluorescent substance or a luminescent substance to bind to the nucleic acid that is hybridized to a probe, and measuring the intensity of the fluorescence or developed color emitted from the probe region. Specifically, the signal may be obtained using a fluoroimaging analyzer, a CCD camera, or the like.

Signals "corresponding to" a probe may include signals originating from the background in the signals, but signals "originating from" a probe mean signals originating from the intrinsic specificity of the probe.

Step (2): Determination of correction value

Next, a straight line L that passes through the fluorescence coordinates P thus plotted and the intersection O is determined, and a value that is inversely proportional to the gradient of this straight line L may be designated as a correction value C ($C > 0$).

As a specific example, the correction value C may also be a value inversely proportional to the radian angle $\alpha$ ($0 \leq \alpha \leq \pi/2$) formed by the straight line L and the X-axis (Fig. 1: Panel C). That is, when the fluorescence coordinates P are corrected to exist on the Y-axis, it is necessary to amplify the angle $\alpha$ to $[(\pi/2) \div \alpha]$ times (Fig. 1: Panel D), but the correction value C may also be defined as Correction value $C = (\pi/2) \div \alpha$ ($C \geq 1$), based on this degree of amplification.

Step (3): Comparison of probe pairs

Usually, in order to detect a single polymorphism, plural candidate probes are prepared. Therefore, it is necessary to carry out the above-described steps (1) and (2) on plural candidate probes, and to thereby determine the correction value C of each

probe.

When a comparison is made between the correction values C obtained in this manner, the performance of the candidate probes may be compared and evaluated (Fig. 1: Panel E).

As discussed above, in an ideal probe, since the fluorescence coordinates P exist on the Y-axis, the relationship $\alpha = \pi/2$ is established. Therefore, in an ideal probe, the correction value C is as follows: Correction value $C = (\pi/2) \div (\pi/2) = 1$.

On the other hand, in the case of a probe causing non-specific hybridization to a certain extent, since $\alpha < \pi/2$, the correction value C is larger than 1. For example, in the case of $\alpha = \pi/4$, the correction value C is equal to 2, and in the case of $\alpha = \pi/6$, the correction value C is equal to 3.

Therefore, according to the method of the present invention, it is considered that as the value of the correction value C is smaller, the probe has superior performance. That is, a probe having the minimum value of the correction value C (that is, the angle $\alpha$ is the maximum) is determined as a probe appropriate for the first polymorphism detection. For example, in the example of Fig. 1 Panel E, the probe pair No. 3 is determined as a probe appropriate for the first polymorphism detection.

The processes described above may also be subjected to various modifications.

For example, in regard to step (1), two or more points of fluorescence coordinates may be obtained by repeating hybridization between a control nucleic acid and a probe two or more times (Fig. 2: Panel A: in this case, three points of fluorescence coordinates). In this case, a representative value M of the two or more points of the fluorescence coordinates thus obtained is determined, and the straight line L according to step (2) may be a median straight line that passes through the intersection O and the

representative value M (Fig. 2: Panel B). Here, the representative value is a value representing plural values. Examples include an average value, a median value, and a weighted average value, but from the viewpoint of robustness against outliers, a median value is preferred.

Furthermore, in regard to step (1), among the various straight lines (since there are two or more points of fluorescence coordinates, there are also two or more straight lines) that each pass through the intersection O and the fluorescence coordinates, a straight line having a difference in the gradient with the median straight line is selected, and this may be designated as an error straight line (when a straight line having the maximum difference in the gradient is selected, this is designated as a first error straight line) (Fig. 2: Panel B).

When a first error straight line is determined, step (2) includes:

(a) a process of determining the angle $\alpha$ (radian) between the median straight line and the X-axis, and

determining the correction value $C = \pi/2 \div \alpha$ (Fig. 2: Panel C); and

(b) a process of designating the angle formed by the median straight line and the error straight line (when a straight line having the maximum difference in the gradient is selected, this is done with the first error straight line) as an error angle $\theta$ (radian), and

defining that correction error angle $\theta'$ (radian) = $\theta$ (radian) $\times$ correction value C.

The error angle may be subjected to constant multiplication described above as necessary. Also, a straight line having the largest difference as the straight line having a difference may be designated as the first error straight line, and a range of the error angle added with a confidence interval may be determined from the angle differences

with plural straight lines having a difference.

On the other hand, steps (4) to (6) corresponding to the steps (1) to (3) may also be carried out for a second probe for polymorphism detection, using a control nucleic acid for second polymorphism.

Specifically, steps (4) to (6) are as follows.

(4) a step of plotting fluorescence coordinates obtained by hybridizing a control nucleic acid for second polymorphism with a probe pair for polymorphism detection consisting of a probe for first polymorphism detection and a probe for second polymorphism detection;

(5) a step of designating a value which is proportional to the gradient of the straight line that passes through the intersection O and the fluorescence coordinates plotted in step (4), as a correction value $C_2$; and

(6) a step of carrying out steps (4) and (5) on plural probe pairs for polymorphism detection, comparing the correction values $C_2$ between various probes, and determining a probe pair having the minimum correction value $C_2$ as a probe appropriate for second polymorphism detection.

Step (4): Plotting step

In step (4), the signal intensities obtainable when a control nucleic acid for second polymorphism is hybridized to a probe pair for polymorphism detection consisting of a probe for first polymorphism detection and a probe for second polymorphism detection, are plotted in a fluorescence coordinate system.

Through the plotting process described above, fluorescence coordinates $P_2(x_2,y_2)$ representing the fluorescence characteristics of the second probe for polymorphism

detection are obtained (Fig. 3: Panel A).

Step (5): Determination of correction value

Next, a straight line $L_2$ that passes through the fluorescence coordinates $P_2$ thus plotted and the intersection O is determined, and a value that is proportional to the gradient of this straight line $L_2$ may be designated as a correction value $C_2$ ($C_2 > 0$).

As a specific example, the correction value $C_2$ may be a value inversely proportional to the radian angle $\beta$ ($0 \leq \beta \leq \pi/2$) formed by the straight line $L_2$ and the Y-axis (that is, proportional to the gradient of $L_2$ ($\pi/2\beta$)) (Fig. 3: Panel B). When the fluorescence coordinates $P_2$ are corrected to exist on the X-axis, it is necessary to amplify the angle $\beta$ to $[(\pi/2) \div \beta]$ times (Fig. 3: Panel C), but the correction value $C_2$ may also be defined as Correction value $C_2 = (\pi/2) \div \beta$ ($C_2 \geq 1$), based on this degree of amplification.

Step (6): Comparison of probe pairs

Similarly to the step (3) described above, it is necessary to determine the correction values $C_2$ of various probes by carrying out the above-described steps (4) and (5) on plural candidate probes.

When a comparison is made between the correction values $C_2$ obtained in this manner, the performance of the candidate probes may be compared and evaluated (Fig. 3: Panel D).

As discussed above, in an ideal probe pair, since the fluorescence coordinates $P_2$ exist on the X-axis, the relationship $\beta = \pi/2$ is established. Therefore, in an ideal probe, the correction value $C_2$ is as follows: Correction value $C_2 = (\pi/2) \div (\pi/2) = 1$.

On the other hand, in the case of a probe which causes non-specific

hybridization to a certain extent, since $\beta < \pi/2$, the correction value $C_2$ is larger than 1. For example, in the case of $\beta = \pi/4$, the correction value $C_2$ is equal to 2, and in the case of $\beta = \pi/6$, the correction value $C_2$ is equal to 3.

Therefore, it is considered that as the value of the correction value $C_2$ is smaller, the probe has superior performance. That is, a probe having the minimum value of the correction value $C_2$ is determined as a probe appropriate for the second polymorphism detection. For example, in the example of Fig. 3 Panel D, the probe pair No. 4' is determined as a probe appropriate for the second polymorphism detection.

The processes described above may also be subjected to various modifications.

For example, in regard to step (4), two or more points of fluorescence coordinates may be obtained by repeating hybridization between a control nucleic acid and a probe two or more times (Fig. 4: Panel A: in this case, three points of fluorescence coordinates). In this case, a representative value $M_2$ of the two or more points of the fluorescence coordinates thus obtained is determined, and the straight line $L_2$ according to step (2) may be a second median straight line that passes through the intersection O and the representative value $M_2$ (Fig. 4: Panel B).

Here, the representative value is a value representing plural values. Examples include an average value, a median value, and a weighted average value, but from the viewpoint of robustness against outliers, a median value is preferred.

Furthermore, in regard to step (4), among the various straight lines (since there are two or more points of fluorescence coordinates, there are also two or more straight lines) that each pass through the intersection O and the fluorescence coordinates, a straight line having a difference in the gradient with the second median straight line is selected, and this may be designated as an error straight line (when a straight line having

the maximum difference in the gradient is selected, this is designated as a second error straight line) (Fig. 4: Panel B).

When a second error straight line is determined, step (2) includes:

(a) a process of determining the angle β (radian) between the second median straight line and the Y-axis, and

determining the correction value $C_2 = \pi/2 \div \beta$; and

(b) a process of designating the angle formed by the second median straight line and the error straight line (when a straight line having the maximum difference in the gradient is selected, this is done with the second error straight line) as an error angle $\theta_2$ (radian), and

defining that correction error angle $\theta_2'$ (radian) = $\theta_2$ (radian) × correction value $C_2$.

The error angle may be subjected to constant multiplication described above as necessary. Also, a straight line having the largest difference as the straight line having a difference may be designated as the second error straight line, and a range of the error angle added with a confidence interval may be determined from the angle differences with plural straight lines having a difference.

Furthermore, the present invention provides a method of displaying the correction value C (or $C_2$) of the probe evaluated by the method described above, and the performance of the probe is evaluated. The present invention also provides a method of displaying corrected coordinates and a corrected error range that have been corrected using the correction value C (or $C_2$).

In the evaluation method of the present invention, the performance between

various probes can be easily compared, and it is also possible to determine the genotype by considering the error range.

Hereinafter, the present invention will be described more specifically by way of Examples, but these Examples are only for illustrative purposes and are not intended to limit the present invention.

## EXAMPLES

[Example 1]

An investigation was conducted on detecting the mutation at 25 sites in the β-globin gene all at once using a DNA microarray. The sites of mutation to be detected are presented in the following Table 1.

Table 1 Sites of mutation in β-globin

| Mutation Site | HGVS nomenclature |
|---|---|
| 1 | c-137C>A |
| 2 | c-81A>G |
| 3 | c-80T>C |
| 4 | c-78A>G |
| 5 | c 2T>G |
| 6 | c 5T>C |
| 7 | c 19G>A |
| 8 | c 27_28insG |
| 9 | c 46delT |
| 10 | c 52A>T |
| 11 | c 59A>G |
| 12 | c 79G>A |
| 13 | c 84_85insC |

| Mutation Site | HGVS nomenclature |
|---|---|
| 14 | c.92+1G>T |
| 15 | c.92+5G>C |
| 16 | c.108C>A |
| 17 | c.170G>A |
| 18 | c.216_217insA |
| 19 | c.251G>A |
| 20 | c.316-197C>T |
| 21 | c.364G>C |
| 22 | c.370_3777delACCCCACC |
| 23 | c.380T>G |
| 24 | c.410G>A |
| 25 | c.441_442insAC |

Among these, for probes that detect mutation c.52A>T, c.84_85insC, c.364G>C, and c.380T>G, since the difficulty in the probe design is high due to the characteristics

of vicinal base sequences, the investigation was conducted first.

1. Production of through-hole type DNA microarray

A DNA microarray was produced as follows.

<1-1. Preparation of probe>

Oligonucleotides having the sequences set forth in SEQ ID NO:1 to 18 that served as probes were synthesized.

These were synthesized as oligonucleotides each having an aminohexyl group introduced at the 5'-terminus of the oligonucleotide. After the synthesis, the oligonucleotide was caused to react with methacrylic anhydride, and the product was further purified and fractionated by HPLC. Thus, 5'-terminal vinylated oligonucleotides having the base sequences set forth in SEQ ID NOs:1 to 18 of Table 2 were obtained. Regarding the features of the sequences, SEQ ID NOs:1 and 2 are probes that are affected by the mutation of mutation c.59A>T adjacent to mutation c.52A>T, while SEQ ID NOs:3 and 4 have inosine introduced therein as a universal base that is hybridized to the mutation of mutation c.59A>T adjacent to mutation c.52A>T.

Similarly, SEQ ID NOs:5 and 6 are probes that are affected by the mutation of mutation c.79G>A adjacent to mutation c.84_85insC, while SEQ ID NOs:7 and 8 have inosine introduced therein as a universal base that is hybridized to the mutation of mutation c.79G>A adjacent to mutation c.84_85insC.

Furthermore, in SEQ ID NOs:11 and 12, a site of a different base for detecting mutation is located at the position six bases away from the 3-terminus of the probe, and SEQ ID NOs:17 and 18 have "AA" introduced at both termini.

Table 2 Candidate probe sequences

| | | |
|---|---|---|
| Probe pair candidate 1 for mutation c.52A>T detection | | |
| 12_1_c.52A>T | CTGTGGGGCAAGGTGAACG | SEQ ID NO:1 |
| 12_2_c.52A>T | CTGTGGGGCTAGGTGAACG | SEQ ID NO:2 |
| Probe pair candidate 2 for mutation c.52A>T detection | | |
| 12_c.52A>T①kail | GGCAAGGTGAICGTGGATG | SEQ ID NO:3 |
| 12_c.52A>T②kail | GGCTAGGTGAICGTGGATG | SEQ ID NO:4 |
| Probe pair candidate 1 for mutation c.84_85insC detection | | |
| 15_1_c.84_85insC | TGGTGAGGCCCTGGGCAGG | SEQ ID NO:5 |
| 15_2_c.84_85insC | TGGTGAGGCCCCTGGGCAG | SEQ ID NO:6 |

| | | |
|---|---|---|
| Probe pair candidate 2 for mutation c.84_85insC detection | | |
| 15_c.84_85insC①kail | GTIAGGCCCTGGGCAG | SEQ ID NO:7 |
| 15_c.84_85insC②kail | TIAGGCCCCTGGGCAG | SEQ ID NO:8 |
| | | |
| Probe pair candidate 1 for mutation c.364G>C detection | | |
| 26_1_c.364G>C | TTTGGCAAAGAATTCACCC | SEQ ID NO:9 |
| 26_2_c.364G>C | TTTGGCAAACAATTCACCC | SEQ ID NO:10 |
| Probe pair candidate 2 for mutation c.364G>C detection | | |
| 26_c.364G>C①kail | CCATCACTTTGGCAAAGAATTC | SEQ ID NO:11 |
| 26_c.364G>C②kail | CCATCACTTTGGCAAACAATTC | SEQ ID NO:12 |
| | | |
| Probe pair candidate 1 for mutation c.380T>G detection | | |
| 28_1_c.380T>G | ACCCCACCAGTGCAGGCTG | SEQ ID NO:13 |
| 28_2_c.380T>G | ACCCCACCAGGGCAGGCTG | SEQ ID NO:14 |
| Probe pair candidate 2 for mutation c.380T>G detection | | |
| 28_1_c.380T>G_20111104 | CAGTGCAGGCTGCCTATCAGA | SEQ ID NO:15 |
| 28_2_c.380T>G_20111104 | CAGGGCAGGCTGCCTATCAGA | SEQ ID NO:16 |
| Probe pair candidate 3 for mutation c.380T>G detection | | |
| 27_28 probe ① (Wt-T) | AACCCACCAGTGCAGGCAA | SEQ ID NO:17 |
| 27_28 probe ② (Wt-G) | AACCCACCAGGGCAGGCAA | SEQ ID NO:18 |

The oligonucleotides having the sequences set forth in SEQ ID NOs:1 to 18 may be hybridized to portions of the human β-globin gene sequences.

<1-2. DNA microarray>

In the present Example, nucleic acid microarrays ((GENOPAL: registered trademark), Mitsubishi Rayon Co., Ltd.) which used the probes described in Table 1 (SEQ ID NOs:1 to 18), and used water instead of the nucleic acid probes for those sites that were not mounted with probes, were used.

2. Evaluation of probes for mutation detection in β-globin gene

<2-1. Production of plasmid template DNA>

The β-globin gene sequence is encoded on the complementary strand sequence from the 5246730$^{th}$ base to the 5248465$^{th}$ base of NCBI Reference Sequence: NC_000011.9 (sequence length: 135006516 bases). The sequence is set forth in SEQ ID NO:19. Furthermore, the exon region in the genomic DNA sequence was characterized by comparing with the sequence of NM_000518.4|Homo sapiens hemoglobin, beta (HBB), mRNA set forth in SEQ ID NO:20 (protein coding region 51$^{st}$ base to 494$^{th}$ base, exon 1: 1$^{st}$ base to 142$^{nd}$ base, exon 2: 143$^{rd}$ base to 365$^{th}$ base, exon 3: 366$^{th}$ base to 626$^{th}$ base)

In SEQ ID NO:19, the sequence sites described in italicized characters represents the positions of the primer sequences of SEQ ID NOs:21 to 24, the underlined sequences represent exon regions, and the regions surrounded by rectangles represent UTR regions.

A wild type template for the β-globin gene was prepared by synthesizing a plasmid containing the sequence set forth in SEQ ID NO:19 (inserted into a pUC57 vector using the artificial gene synthesis service provided by BEX Co., Ltd.), and the

template was prepared into a solution having a concentration of 10 ng/l.

Furthermore, similarly to this, individual plasmid DNAs (25 kinds) having mutations introduced at the positions of the notation of mutation according to the HGVS nomenclature as shown in Table 1, were produced, and those were also prepared into solutions having a concentration of 10 ng/l.

Primer pair <SEQ ID NOs:21, 22, 23 and 24>

| | | |
|---|---|---|
| SEQ ID NO:21 | Amplicon1F | ACTCCTAAGCCAGTGCCAGA |
| SEQ ID NO:22 | Amplicon1R | cy5-CACTCAGTGTGGCAAAGGTG |
| SEQ ID NO:23 | MRC-Amplicon2F | GTATCATGCCTCTTTGCACCATTC |
| SEQ ID NO:24 | MRC-Amplicon2R | cy5-CAGATGCTCAAGGCCCTTCATA |

<SEQ ID NO:19>

>gi|224589802:c5248465-5246730 homo sapiens chromosome 11 GRCh37.5 Primary Assembly

A*ACTCCTAAGCCAGTGCCAGA*AGAGCCAAGGACAGGTACGGCTGTCATCACTTAGACCTCACC
CTGTGG
AGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGCTG
GGCATAAA
AGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACC
TCAAA
CAGACACCATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAG
GTGAAC
CTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAAGGTTACAAGACAGGTTTAA
GGAGACC
AATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTC
TCTGCCT
ATTGGTCTATTTTCCCACCCTTAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAG
TCCTT
TGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAA
AGTGCT
CGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGG*CACCTTTGCCACACTGAG*
*TG*AGCT
GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGCTTGAT
GTTTTCT
TTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGGAAGGGGATAAGTAACAGGGTACAGTTT
AGAAT
GGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTTATTTG
CTGTT

CATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCTTTTTTTTTCTTCTCCGCAATTTTTACTAT
TA
TACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAGTAACTTA
AAAAA
AAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAATATATGTGTGCTTATTTGCATATTCAT
AAT
CTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAATCATTATACATATTTATGGGTTAAAGT
GT
AATGTTTTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTTAAAAA
ATGC
TTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTTCAG
GG
CAATAATGATACAAT*GTATCATGCCTCTTTGCACCATTC*TAAAGAATAACAGTGATAATTTCTGGG
TTA
AGGCAATAGCAATATCTCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAGGTT
TCAT
ATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAGGCTGG
ATTA
TTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCTTCCTCCCACAGCTC
CTGG
GCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCACCAGTGCAG
GCTGCCT
ATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCACTAAGCTCGCTTTC
TTGCT
GTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATAT*TATGA*
*AGG*
*GCCTTGAGCATCGG*

<SEQ ID NO:20>

ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGA
CTCCTG
AGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGA
GGCCCTGG
GCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCA
CTCCTG
ATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTG
ATGGCC
TGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAG
CTGCACG
TGGATCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTG
GCAAAG
AATTCACCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGG
CCCACA
AGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCC
AACT
ACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTA
TTTTC
ATTGC

<PCR reaction>

PCR reactions were carried out using five kinds of plasmid DNAs in total, namely, the wild type plasmid DNA, and the four kinds of mutant plasmid DNAs of Nos. 10, 13, 21 and 23 (including mutations c.52A>T, c.84_85insC, c.364G>C, c.380T>G) described in Table 1 as templates, and using two pairs of primers having the sequences of SEQ ID NOs:21 to 24. For the PCR reaction, a KOD FX Neo kit (Toyobo Co., Ltd.) was used.

<PCR reaction liquid composition>

| | |
|---|---|
| Plasmid DNA solution (10 ng/μL) | 1 μL (wild type or mutant) |
| Amplicon1F primer (20 μM) | 1 μL |
| Amplicon1R primer (20 μM) | 1 μL |
| MRC-Amplicon2F primer (20 μM) | 0.5 μL |
| MRC-Amplicon2R primer (20 μM) | 0.5 μL |
| 2×buffer | 50 μl |
| 2 mM dNTPs | 20 μl |
| MILLI Q water | 24 μl |
| KOD FX Neo | 2 μl |
| Reaction volume | 100 μl |

For the PCR reaction, a GeneAmp9700 thermal cycler was used, and the reaction was carried out in the Max mode. The temperature conditions are shown below.

<PCR reaction temperature conditions>

95°C for 10 minutes

(94°C for 30 seconds, 68°C for 30 seconds, and 72°C for 30 seconds) × 35 cycles

4°C end of reaction

The following buffer solution was added to 100 μl of the reaction liquid obtained after the reaction to obtain a final volume of 200 μl.

| | | |
|---|---|---|
| Reaction liquid | 100 μl | |
| 1M Tris/HCl (pH 7.5) buffer | 48 μl | |
| 5M NaCl solution | | 9.6 μl |
| 0.5% aqueous solution of Tween20 | 20 μl | |
| MILLI Q water | | 22.4 μl |
| Total | | 200 μl |

Thereafter, 200 μl of this solution was introduced into a chamber for exclusive use (described in http://www.mrc.co.jp/genome/about/usage.html), subsequently the DNA microarrays were introduced therein, the chamber was covered with a lid, and the mixture was incubated at 55°C for 2 hours.

After the incubation, each of the chips was immersed in 10 ml of 0.24 M TNT buffer at 55°C for 20 minutes. Thereafter, subsequently, each of the chips was immersed in 10 ml of 0.24 M TN buffer at 55° for 10 minutes to perform washing. After the washing, detection was performed.

The detection was carried out using an automated DNA microarray detection apparatus of a cooled CCD camera system. The DNA microarrays were subjected to image-capturing from the top of the wells for an exposure time of 4 seconds, and the fluorescent signals of Cy5 at various spots were detected. A spot where the probes on the microarrays were not mounted was designated as a blank spot, and the median value of the fluorescence intensity thereof was designated as the background value. Values obtained by subtracting the background value from the fluorescence intensities at all of the spots were designated as the signals of the various probes.

The results obtained by performing the experiment several times by employing the sequence of the wild type as a first polymorphism, the sequence of a mutant as a second polymorphism, and the control nucleic acid for first polymorphism as a wild type plasmid, are presented in Table 3. Furthermore, Fig. 5 shows the results of plotting the results of Table 3 in a fluorescence coordinate system which included a Y-axis representing the signal intensity obtainable when the probe for first polymorphism detection was hybridized, and an X-axis representing the signal intensity obtainable when the probe for second polymorphism detection was hybridized, with the X-axis and the Y-axis perpendicularly intersecting each other (Fig. 5: a diagram obtained by plotting the results of performing hybridization of the first control nucleic acid several times in a fluorescence coordinate system representing the signal intensities of the probes for first and second polymorphism detection, and showing representative straight lines thereof).

A dotted line in Fig. 5 is a straight line (considered as a representative straight line) that links between the average signal intensity of the results of plural experiments (2 times or 3 times) for each candidate probe pair, and the zero point. A mathematical formula in the graph represents the formula for such straight lines.

Regarding the selection of the probe, a value that is inversely proportional to the gradient of the representative straight line is designated as a correction value C, this is carried out for plural probe pairs for polymorphism detection to compare the correction

values, and a probe pair having the minimum correction value C is selected. In the present investigation, the correction value C was calculated by the formula: $\pi/2 \div$ (angle (radian) formed by the representative straight line and the X-axis). Among the various probe pairs, the following probe pairs could be selected among the probe candidates as probe pairs having favorable performance:

as a probe for detecting mutation of c.52A>T, the pair of SEQ ID NOs:3 and 4 was selected between the pair of SEQ ID NOs:1 and 2 and the pair of SEQ ID NOs:3 and 4;

as a probe for detecting mutation of c.84_85insC, the pair of SEQ ID NOs:7 and 8 was selected between the pair of SEQ ID NOs:5 and 6 and the pair of SEQ ID NOs:7 and 8;

as a probe for detecting mutation of c.364G>C, the pair of SEQ ID NOs:11 and 12 was selected between the pair of SEQ ID NOs:9 and 10 and the pair of SEQ ID NOs:11 and 12; and

as a probe for detecting mutation of c.380T>G, the pair of SEQ ID NOs:17 and 18 was selected among the pair of SEQ ID NOs:13 and 14, the pair of SEQ ID NOs:15 and 16, and the pair of SEQ ID NOs:17 and 18.

Table 3 Results obtained by performing the experiment plural times using a wild type plasmid (control nucleic acid for first polymorphism)

| Site to be detected | | | Probe name | Probe sequence | Fluorescence obtained by hybridizing control nucleic acid for first polymorphism (wild type) with probe pair for polymorphism detection: Signal intensity of 1st test | Signal intensity of 2nd test | Signal intensity of 3rd test |
|---|---|---|---|---|---|---|---|
| c.52A>T | Pair of candidate 1 | Probe for first polymorphism detection | 12_1_c.52A>T | CTGTGGGGC**A**AGGTGAACG | 9034 | 7672 | |
| | | Probe for second polymorphism detection | 12_2_c.52A>T | CTGTGGGGC**T**AGGTGAACG | 6783 | 5918 | |
| | Pair of candidate 2 | Probe for first polymorphism detection | 12_c.52A>T①kail | GGC**A**AGGTGA**I**CGTGGATG | 3106 | 2391 | |
| | | Probe for second polymorphism detection | 12_c.52A>T②kail | GGC**T**AGGTGA**I**CGTGGATG | 289 | 242 | |
| Site to be detected | | | | | | | |
| c.84_85insC | Pair of candidate 1 | Probe for first polymorphism detection | 15_1_c.84_85insC | TGGTGAGGCCCTGGGCAGG | 14940 | 12913 | 11755 |
| | | Probe for second polymorphism detection | 15_2_c.84_85insC | TGGTGAGG**C**CCCTGGGCAG | 11353 | 10148 | 11091 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 15_c.84_85insC①kail | GT**I**AGGCCCTGGGCAG | 5021 | 4024 | 4228 |
| | | Probe for second polymorphism detection | 15_c.84_85insC②kail | T**I**AGG**C**CCCTGGGCAG | 102 | 83 | 85 |
| Site to be detected | | | | | | | |
| c.364G>C | Pair of candidate 1 | Probe for first polymorphism detection | 26_1_c.364G>C | TTTGGCAAA**G**AATTCACCC | 9912 | 9880 | 8661 |
| | | Probe for second polymorphism detection | 26_2_c.364G>C | TTTGGCAAA**C**AATTCACCC | 214 | 183 | 196 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 26_c.364G>C①kail | CCATCACTTTGGCAAA**G**AATTC | 17501 | 17171 | 17563 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Probe for second polymorphism detection | 26_c.364G>C②kail | CCATCACTTTGGCAAA**C**AATTC | 285 | 286 | 278 |

| Site to be detected | | | | | | | |
|---|---|---|---|---|---|---|---|
| c.380T>G | Pair of candidate 1 | Probe for first polymorphism detection | 28_1_c.380T>G | ACCCCACCAG**T**GCAGGCTG | 35457 | 35876 | 26406 |
| | | Probe for second polymorphism detection | 28_2_c.380T>G | ACCCCACCAG**G**GCAGGCTG | 22184 | 20900 | 17312 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 28_1_c.380T>G_20111104 | CAG**T**GCAGGCTGCCTATCAGA | 29308 | 24727 | 23549 |
| | | Probe for second polymorphism detection | 28_2_c.380T>G_20111104 | CAG**G**GCAGGCTGCCTATCAGA | 20927 | 18841 | 17785 |
| | Pair of candidate 3 | Probe for first polymorphism detection | 27_28 probe ① (Wt-T) | A**A**CCCACCAG**T**GCAGGC**AA** | 16676 | 16009 | 15860 |
| | | Probe for second polymorphism detection | 27_28 probe ② (Wt-T) | A**A**CCCACCAG**G**GCAGGC**AA** | 5830 | 5776 | 5505 |

In regard to the probe sequences presented in Table 3, a single-underlined base is the polymorphism to be detected, and a double-underlined base is a base that has been subjected to the modification of the present invention (inosine substitution or adenine insertion).

Similarly, the results obtained by performing the experiment several times by employing the sequence of the wild type as a first polymorphism, the sequence of a mutant as a second polymorphism, and the control nucleic acid for second polymorphism as a mutant plasmid, are presented in Table 4. Furthermore, Fig. 6 shows the results obtained by plotting the results of Table 3 and Table 4 in a fluorescence coordinate system which included a Y-axis representing the signal intensity obtainable when the probe for first polymorphism detection was hybridized, and an X-axis representing the signal intensity obtainable when the probe for second polymorphism detection was hybridized, with the X-axis and the Y-axis perpendicularly intersecting each other (in addition to Fig. 5, Fig. 6 is also a diagram obtained by plotting the results of performing hybridization of the second control nucleic acid several times, and showing representative straight lines thereof).

Table 4 Results obtained by performing the experiment plural times using a mutant plasmid (control nucleic acid for second polymorphism)

| Site to be detected | | | Probe name | Probe sequence | Fluorescence obtained by hybridizing control nucleic acid for first polymorphism (wild type) with probe pair for polymorphism detection: Signal intensity of 1st test | Signal intensity of 2nd test | Signal intensity of 3rd test |
|---|---|---|---|---|---|---|---|
| c.52A>T | Pair of candidate 1 | Probe for first polymorphism detection | 12_1_c.52A>T | CTGTGGGGC**A**AGGTGAACG | 9034 | 7672 | |
| | | Probe for second polymorphism detection | 12_2_c.52A>T | CTGTGGGGC**T**AGGTGAACG | 6783 | 5918 | |
| | Pair of candidate 2 | Probe for first polymorphism detection | 12_c.52A>T①kail | GGC**A**AGGTGA**I**CGTGGATG | 3106 | 2391 | |
| | | Probe for second polymorphism detection | 12_c.52A>T②kail | GGC**T**AGGTGA**I**CGTGGATG | 289 | 242 | |
| Site to be detected | | | | | | | |
| c.84_85insC | Pair of candidate 1 | Probe for first polymorphism detection | 15_1_c.84_85insC | TGGTGAGGCCCTGGGCAGG | 14940 | 12913 | 11755 |
| | | Probe for second polymorphism detection | 15_2_c.84_85insC | TGGTGAGG**C**CCCTGGGCAG | 11353 | 10148 | 11091 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 15_c.84_85insC①kail | GT**I**AGGCCCTGGGCAG | 5021 | 4024 | 4228 |
| | | Probe for second polymorphism detection | 15_c.84_85insC②kail | T**I**AGG**C**CCCTGGGCAG | 102 | 83 | 85 |
| Site to be detected | | | | | | | |
| c.364G>C | Pair of candidate 1 | Probe for first polymorphism detection | 26_1_c.364G>C | TTTGGCAAA**G**AATTCACCC | 9912 | 9880 | 8661 |
| | | Probe for second polymorphism detection | 26_2_c.364G>C | TTTGGCAAA**C**AATTCACCC | 214 | 183 | 196 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 26_c.364G>C①kail | CCATCACTTTGGCAAA**G**AATTC | 17501 | 17171 | 17563 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Probe for second polymorphism detection | 26_c.364G>C②kail | CCATCACTTTGGCAAA**C**AATTC | 285 | 286 | 278 |

| Site to be detected | | | | | | | |
|---|---|---|---|---|---|---|---|
| c.380T>G | Pair of candidate 1 | Probe for first polymorphism detection | 28_1_c.380T>G | ACCCCACCAG**T**GCAGGCTG | 35457 | 35876 | 26406 |
| | | Probe for second polymorphism detection | 28_2_c.380T>G | ACCCCACCAG**G**GCAGGCTG | 22184 | 20900 | 17312 |
| | Pair of candidate 2 | Probe for first polymorphism detection | 28_1_c.380T>G_20111104 | CAG**T**GCAGGCTGCCTATCAGA | 29308 | 24727 | 23549 |
| | | Probe for second polymorphism detection | 28_2_c.380T>G_20111104 | CAG**G**GCAGGCTGCCTATCAGA | 20927 | 18841 | 17785 |
| | Pair of candidate 3 | Probe for first polymorphism detection | 27_28 probe ① (Wt-T) | A**A**CCCACCAG**T**GCAGGC**AA** | 16676 | 16009 | 15860 |
| | | Probe for second polymorphism detection | 27_28 probe ② (Wt-G) | A**A**CCCACCAG**G**GCAGGC**AA** | 5830 | 5776 | 5505 |

In regard to the probe sequences indicated in Table 4, a single-underlined base is the polymorphism to be detected, and a double-underlined base is a base that has been subjected to the modification of the present invention (inosine substitution or adenine insertion).

The series including the "hybridized to control nucleic acid for second polymorphism" in the graph of Fig. 6 are the results obtained by hybridizing the mutant plasmid. A dotted line or a solid line is a representative straight line that links between the average signal intensity of the results of plural experiments (2 times or 3 times) for each candidate probe pair, and the zero point.

Regarding the selection of these probes, a value that is proportional to the gradient of the representative straight line is designated as a correction value $C_2$, this is carried out for plural probe pairs for polymorphism detection to compare the correction values, and a probe pair having the minimum correction value $C_2$, which is appropriate for the detection of second polymorphism (mutant), is selected.

In the present investigation, the correction value $C_2$ was calculated by the formula: $\pi/2 \div (\pi/2$ - angle (radian) formed by the representative straight line and the X-axis). Among the various probe pairs, the following probe pairs could be selected among the probe candidates as probe pairs having favorable performance:

as a probe for detecting mutation of c.52A>T, the pair of SEQ ID NOs:3 and 4 was selected between the pair of SEQ ID NOs:1 and 2 and the pair of SEQ ID NOs:3 and 4;

as a probe for detecting mutation of c.84_85insC, the pair of SEQ ID NOs:7 and 8 was selected between the pair of SEQ ID NOs:5 and 6 and the pair of SEQ ID NOs:7 and 8;

as a probe for detecting mutation of c.364G>C, the pair of SEQ ID NOs:11 and

12 was selected between the pair of SEQ ID NOs:9 and 10 and the pair of SEQ ID NOs:11 and 12; and

as a probe for detecting mutation of c.380T>G, the pair of SEQ ID NOs:17 and 18 was selected among the pair of SEQ ID NOs:13 and 14, the pair of SEQ ID NOs:15 and 16, and the pair of SEQ ID NOs:17 and 18.

Graphs of the correction values C and $C_2$ described so far are presented in Fig. 7.

Subsequently to the evaluation of the probes, the error range that would be useful at the time of determining the genotype was set as shown in the following Table 5. The average value was calculated from the signal intensities obtained by repeating the procedure two or more times using the first control nucleic acid or the second control nucleic acid, and the average value was designated as the representative coordinates given by the probe pair. Furthermore, the straight line passing through the representative coordinates and the zero point was designated as a representative straight line, the angle between the X-axis and the representative straight line was designated as a representative coordinate angle, and the angle (radian unit) between a straight line that linked the individual data and the zero point, and the representative straight line was calculated. The maximum angle was designated as an error angle. Fig. 8 shows the probe performance data obtained before and after the correction made using the correction values C and $C_2$, and the error angle.

Table 5 Specific examples of correction method of present invention

Fluorescence obtained by hybridizing control nucleic acid for first polymorphism with probe pair for polymorphism detection

| Site to be detected | Probe name | Probe name | Signal intensity of 1st test | Signal intensity of 2nd test | Signal intensity of 3rd test | Test 1 angle | Test 2 angle | Test 3 angle | Representative coordinates | Representative coordinates angle | Difference in angle with representative straight line | | | Angle of maximum difference | Correction value C | Error angle after correction |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c.52A>T | Pair of candidate 1 | 12_1_c.52A>T | 9034 | 7672 | | 0.9268 | 0.9138 | | 8353 | 0.921 | 0.006 | 0.007 | | 0.007 | 1.706 | 0.0119 |
| | | 12_2_c.52A>T | 6783 | 5918 | | | | | 6351 | | | | | | | |
| | Pair of candidate 2 | 12_c.52A>Tjkail | 3106 | 2391 | | 1.4781 | 1.4701 | | 2748 | 1.475 | 0.0035 | 0.045 | | 0.0045 | 1.0652 | 0.0048 |
| | | 12_c.52A>Tkkail | 289 | 242 | | | | | 265 | | | | | | | |
| c.84_85insC | Pair of candidate 1 | 15_1_c.84_85insC | 14940 | 12913 | 11755 | 0.921 | 0.9047 | 0.8145 | 13203 | 0.882 | 0.0387 | 0.0225 | 0.0678 | 0.0678 | 1.7804 | 0.1207 |
| | | 15_2_c.84_85insC | 11353 | 10148 | 11091 | | | | 10864 | | | | | | | |
| | Pair of candidate 2 | 15_c.84_85insCjkail | 5021 | 4024 | 4228 | 1.5505 | 1.5503 | 1.5506 | 4425 | 1.55 | 0.0001 | 0.0002 | 0.0001 | 0.0002 | 1.0131 | 0.0002 |
| | | 15_c.84_85insCkkail | 102 | 83 | 85 | | | | 90 | | | | | | | |
| c.364G>C | Pair of candidate 1 | 26_1_c.364G>C | 9912 | 9880 | 8661 | 1.5492 | 1.5523 | 1.5482 | 9484 | 1.55 | 0.0007 | 0.0023 | 0.0018 | 0.0023 | 1.0134 | 0.0023 |
| | | 26_2_c.364G>C | 214 | 183 | 196 | | | | 198 | | | | | | | |
| | Pair of candidate 2 | 26_c.364G>Cjkail | 17501 | 17171 | 17563 | 1.5545 | 1.5542 | 1.555 | 17412 | 1.555 | 0 | 0.0004 | 0.0004 | 0.0004 | 1.0105 | 0.0004 |
| | | 26_c.364G>Ckkail | 285 | 286 | 278 | | | | 283 | | | | | | | |
| c.380T>G | Pair of candidate 1 | 28_1_c.380T>G | 35457 | 35876 | 26406 | 1.0117 | 1.0433 | 0.9905 | 32580 | 1.017 | 0.0056 | 0.026 | 0.0268 | 0.0268 | 1.5441 | 0.0414 |
| | | 28_2_c.380T>G | 22184 | 20900 | 17312 | | | | 20132 | | | | | | | |
| | Pair of candidate 2 | 28_1_c.380T>G_20111104 | 29308 | 24727 | 23549 | 0.9507 | 0.9197 | 0.924 | 25861 | 0.933 | 0.0182 | 0.0129 | 0.0086 | 0.0182 | 1.6844 | 0.0306 |
| | | 28_2_c.380T>G_20111104 | 20927 | 18841 | 17785 | | | | 19184 | | | | | | | |
| | Pair of candidate 3 | 27_28 probe j (Wt-T) | 16676 | 16009 | 15860 | 1.2345 | 1.2245 | 1.2367 | 16182 | 1.232 | 0.0026 | 0.0074 | 0.0048 | 0.0074 | 1.2751 | 0.0094 |
| | | 27_28probe k (Wt-G) | 5830 | 5776 | 5505 | | | | 5704 | | | | | | | |

Fluorescence obtained by hybridizing control nucleic acid for second polymorphism with probe pair for polymorphism detection

| Site to be detected | | Probe name | Signal intensity of 1st test | Signal intensity of 2nd test | Signal intensity of 3rd test | Test 1 angle | Test 2 angle | Test 3 angle | Representative coordinates | Representative coordinates angle | Difference in angle with representative straight line | | | Angle of maximum difference | Correction value C2 | Error angle after correction |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c.52>T | Pair of candidate 1 | 12_1_c.52A>T | 6051 | 6400 | | 0.6718 | 0.6492 | | 6226 | 0.66 | 0.0118 | 0.0108 | | 0.0118 | 1.7247 | 0.0204 |
| | | 12_2_c.52A>T | 7609 | 8433 | | | | | 8021 | | | | | | | |
| | Pair of candidate 2 | 12_c.52A>Tjkai1 | 220 | 194 | | 0.0732 | 0.0924 | | 207 | 0.081 | 0.0079 | 0.0113 | | 0.0113 | 1.0544 | 0.012 |
| | | 12_c.52A>Tkkai1 | 2997 | 2089 | | | | | 2543 | | | | | | | |
| c.84_85insC | Pair of candidate 1 | 15_1_c.84_85insC | 13649 | 14420 | 14327 | 0.7043 | 0.7163 | 0.6842 | 14132 | 0.701 | 0.0029 | 0.0149 | 0.0172 | 0.0172 | 1.8067 | 0.0311 |
| | | 15_2_c.84_85insC | 16065 | 16565 | 17567 | | | | 16732 | | | | | | | |
| | Pair of candidate 2 | 15_c.84_85insCjkai1 | 561 | 509 | | 0.241 | 2642 | | 535 | 0.251 | 0.0105 | 0.0127 | | 0.0127 | 1.1906 | 0.0151 |
| | | 15_c.84_85insCkkai1 | 2283 | 1882 | | | | | 2083 | | | | | | | |
| c.364G>C | Pair of candidate 1 | 26_1_c.364G>C | 2544 | 2281 | 2406 | 0.2197 | 0.2202 | 0.1964 | 2410 | 0.212 | 0.0082 | 0.0086 | 0.0151 | 0.0151 | 1.1556 | 0.0174 |
| | | 26_2_c.364G>C | 11394 | 10193 | 12090 | | | | 11226 | | | | | | | |
| | Pair of candidate 2 | 26_c.364G>Cjkai1 | 2970 | 2784 | | 0.1587 | 0.181 | | 2877 | 0.169 | 0.0101 | 0.0122 | | 0.0122 | 1.1203 | 0.0137 |
| | | 26_c.364G>Ckkai1 | 18563 | 15216 | | | | | 16889 | | | | | | | |
| c.380T>G | Pair of candidate 1 | 28_1_c.380T>G | 14565 | 13692 | 14086 | 0.4628 | 0.4693 | 0.4649 | 14114 | 0.466 | 0.0027 | 0.0037 | 0.0007 | 0.0037 | 1.4212 | 0.0052 |
| | | 28_2_c.380T>G | 29190 | 27004 | 28087 | | | | 28094 | | | | | | | |
| | Pair of candidate 2 | 28_1_c.380T>G_20111104 | 11250 | 10456 | | 0.6087 | 0.5863 | | 10853 | 0.598 | 0.0109 | 0.0114 | | 0.0114 | 1.6143 | 0.0184 |
| | | 28_2_c.380T>G_20111104 | 16143 | 15741 | | | | | 15942 | | | | | | | |
| | Pair of candidate 3 | 27_28probe j (Wt-T) | 188 | 137 | | 0.0137 | 0.01 | | 162 | 0.012 | 0.0019 | 0.0019 | | 0.0019 | 1.0076 | 0.0019 |
| | | 27_28probe k (Wt-G) | 13677 | 13626 | | | | | 13651 | | | | | | | |

[Example 2]

In order to detect all at once the mutations at 25 sites in the β-globin gene using a DNA microarray, an array mounted with probes having the sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17 and 18, and SEQ ID NOs:25 to 66 was produced.

The sites of mutation to be detected were the same as shown in Table 1 of Example 1, and the DNA microarray was also produced in the same manner as in Example 1.

<PCR reaction>

PCR reactions were carried out using the mutant plasmid DNAs of Nos. 1 to 25 described in Table 1 as templates, and using two pairs of primers having the sequences of SEQ ID NOs:21 to 24. For the PCR reactions, an Ampdirect Plus kit (Shimadzu Corp.) was used.

<PCR reaction liquid composition>

| | |
|---|---|
| Plasmid DNA solution (10 ng/μL) | 1 μL (wild type or mutant) |
| Amplicon1F primer (20 μM) | 1 μL |
| Amplicon1R primer (20 μM) | 1 μL |
| MRC-Amplicon2F primer (20 μM) | 0.5 μL |
| MRC-Amplicon2R primer (20 μM) | 0.5 μL |
| 2×Ampdirect buffer | 50 μl |
| BioTaq (accompanying Ampdirect Plus kit) | 1 μl |
| MILLI Q water | 45 μl |
| Total | 100 μl |

For the PCR reaction, a GeneAmp9700 thermal cycler was used, and the reaction was carried out in the Max mode. The temperature conditions are shown below.

<PCR reaction temperature conditions>

95°C for 10 minutes

(94°C for 30 seconds, 68°C for 30 seconds, and 72°C for 30 seconds) × 35 cycles

4°C end of reaction

The following buffer solution was added to 100 μl of the reaction liquid obtained after the reaction to obtain a final volume of 200 μl.

| | |
|---|---|
| Reaction liquid | 100 μl |

| | |
|---|---|
| 1M Tris/HCl (pH 7.5) buffer | 48 μl |
| 5M NaCl solution | 9.6 μl |
| 0.5% aqueous solution of Tween20 | 20 μl |
| MILLI Q water | 22.4 μl |
| Total | 200 μl |

Thereafter, 200 μl of this solution was introduced into a chamber for exclusive use (described in http://www.mrc.co.jp/genome/about/usage.html), subsequently the DNA microarrays were introduced therein, the chamber was covered with a lid, and the mixture was incubated at 55°C for 2 hours.

After the incubation, each of the chips was immersed in 10 ml of 0.24 M TNT buffer at 55°C for 20 minutes. Thereafter, subsequently, each of the chips was immersed in 10 ml of 0.24 M TN buffer at 55° for 10 minutes to perform washing. After the washing, detection was performed.

The detection was carried out using an automated DNA microarray detection apparatus of a cooled CCD camera system. The DNA microarrays were subjected to image-capturing from the top of the wells for an exposure time of 4 seconds, and the fluorescent signals of Cy5 at various spots were detected. A spot where the probes on the microarrays were not mounted was designated as a blank spot, and the median value of the fluorescence intensity thereof was designated as the background value. Values obtained by subtracting the background value from the fluorescence intensities at all of the spots were designated as the signals of the various probes.

The results are summarized in Table 6.

Table 6 Detection results obtained using 25 kinds of mutant plasmids: signal values

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Site of mutation 1 reference nucleic acid sample | Site of mutation 2 reference nucleic acid sample | Site of mutation 3 reference nucleic acid sample | Site of mutation 4 reference nucleic acid sample | Site of mutation 5 reference nucleic acid sample | Site of mutation 6 reference nucleic acid sample | Site of mutation 7 reference nucleic acid sample | Site of mutation 8 reference nucleic acid sample | Site of mutation 9 reference nucleic acid sample | Site of mutation 10 reference nucleic acid sample | Site of mutation 11 reference nucleic acid sample | Site of mutation 12 reference nucleic acid sample | Site of mutation 13 reference nucleic acid sample | Site of mutation 14 reference nucleic acid sample | Site of mutation 15 reference nucleic acid sample | Site of mutation 16 reference nucleic acid sample | Site of mutation 17 reference nucleic acid sample | Site of mutation 18 reference nucleic acid sample | Site of mutation 19 reference nucleic acid sample | Site of mutation 20 reference nucleic acid sample | Site of mutation 21 reference nucleic acid sample | Site of mutation 22 reference nucleic acid sample | Site of mutation 23 reference nucleic acid sample | Site of mutation 24 reference nucleic acid sample | Site of mutation 25 reference nucleic acid sample |
| Probe name | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal | Signal |
| 1_1_c.-137C>A | 367 | 3517 | 3026 | 2114 | 2724 | 2722 | 2201 | 2610 | 2447 | 2783 | 2803 | 3046 | 2268 | 2853 | 2620 | 2058 | 2455 | 1973 | 2597 | 2615 | 2597 | 2632 | 2396 | 2909 | 319 |
| 1_2_c.-137C>A | 5647 | 1353 | 1205 | 1091 | 1217 | 1056 | 1353 | 1161 | 1061 | 1207 | 1032 | 1313 | 1060 | 1216 | 1252 | 1211 | 1425 | 861 | 1062 | 1335 | 1022 | 1212 | 1013 | 1410 | 57 |
| 2_c.-81A>Gj | 7023 | 797 | 2733 | 382 | 5829 | 5056 | 7095 | 7045 | 6300 | 7353 | 5837 | 6874 | 8281 | 6205 | 6513 | 5893 | 6482 | 5885 | 6052 | 7428 | 6188 | 5244 | 6071 | 1899 | 402 |
| 2_c.-81A>Gk | 3466 | 9753 | 612 | 40 | 2942 | 2888 | 3395 | 3235 | 2313 | 3252 | 2459 | 3408 | 3195 | 3139 | 2817 | 2474 | 2845 | 2549 | 2392 | 2517 | 2805 | 2553 | 2672 | 2172 | 31 |
| 3_1_c.-80T>C | 6521 | 944 | 3129 | 528 | 6117 | 5509 | 8983 | 6744 | 6755 | 5785 | 5832 | 7087 | 6299 | 7734 | 8598 | 6009 | 7544 | 8444 | 6006 | 6735 | 6735 | 5458 | 6900 | 6043 | 5430 |
| 3_2_c.-80T>C | 1009 | 288 | 8303 | 23 | 883 | 868 | 1057 | 958 | 662 | 944 | 942 | 1001 | 917 | 922 | 983 | 1221 | 945 | 781 | 767 | 744 | 873 | 777 | 846 | 760 | 873 |
| 4_1_c.-78A>G | 4544 | 275 | 1170 | 93 | 4030 | 3258 | 4239 | 4590 | 3612 | 4492 | 4022 | 4360 | 4018 | 3304 | 3616 | 4136 | 3843 | 3607 | 3797 | 3757 | 4396 | 3331 | 3228 | 3604 | 2450 |
| 4_2_c.-78A>G | 681 | 28 | 70 | 5508 | 591 | 587 | 695 | 627 | 572 | 605 | 598 | 648 | 605 | 594 | 599 | 627 | 607 | 546 | 511 | 524 | 594 | 467 | 560 | 526 | 294 |
| 5_1_c.2T>G | 9074 | 10161 | 9770 | 8194 | 499 | 4607 | 8834 | 9947 | 6398 | 9504 | 6097 | 9129 | 5304 | 8788 | 6836 | 9612 | 8476 | 7805 | 8211 | 7642 | 8628 | 7626 | 5084 | 7889 | 6308 |
| 5_2_c.2T>G | 4342 | 4817 | 2895 | 3141 | 12083 | 1389 | 4775 | 4396 | 3136 | 3700 | 3599 | 3870 | 3092 | 3989 | 4036 | 4206 | 3890 | 3087 | 3164 | 3488 | 3309 | 3207 | 3209 | 3614 | 2339 |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6_1_c.5T>C | 9800 | 10890 | 9549 | 7388 | 729 | 4192 | 10194 | 10103 | 8559 | 7183 | 5247 | 9240 | 6017 | 7997 | 10777 | 8681 | 10972 | 7638 | 7660 | 8002 | 9069 | 7759 | 8982 | 7216 | 6706 |
| 6_2_c.5T>C | 2198 | 2604 | 2318 | 1821 | 190 | 9473 | 2471 | 2482 | 1588 | 2246 | 1858 | 2332 | 1950 | 1933 | 1998 | 2327 | 1985 | 1748 | 1947 | 1667 | 1893 | 1739 | 1700 | 1892 | 1579 |
| 7_1_c.19G>A | 4827 | 4987 | 4427 | 4171 | 4006 | 3227 | 273 | 2962 | 4250 | 4388 | 4259 | 4595 | 4035 | 4612 | 4153 | 5224 | 4501 | 4610 | 3808 | 3670 | 4738 | 3710 | 4359 | 3867 | 4005 |
| 7_2_c.19G>A | 33 | 34 | 30 | 28 | 25 | 30 | 2283 | 20 | 24 | 30 | 27 | 30 | 25 | 27 | 29 | 32 | 23 | 26 | 25 | 23 | 28 | 22 | 28 | 23 | 30 |
| 10_c.27_28insGjkail | 9499 | 10850 | 9449 | 8289 | 7707 | 7402 | 10597 | 2247 | 8354 | 10383 | 8306 | 9689 | 8693 | 9502 | 10121 | 11058 | 10459 | 9239 | 9847 | 9157 | 9511 | 8424 | 7913 | 7732 | 8355 |
| 10_c.27_28insGkkail | 470 | 475 | 418 | 452 | 376 | 365 | 477 | 8134 | 472 | 483 | 454 | 431 | 580 | 411 | 450 | 480 | 359 | 422 | 339 | 337 | 447 | 317 | 432 | 350 | 514 |
| 11_c.46delTjkail | 3804 | 4095 | 3647 | 3815 | 3191 | 3214 | 4109 | 3847 | 182 | 4949 | 3633 | 3626 | 3076 | 3496 | 3506 | 4255 | 3938 | 3473 | 3178 | 3299 | 3750 | 3258 | 3544 | 3183 | 3125 |
| 11_c.46delTkkail | 24 | 27 | 21 | 22 | 19 | 18 | 23 | 23 | 1341 | 51 | 25 | 21 | 15 | 22 | 21 | 25 | 20 | 24 | 19 | 18 | 24 | 17 | 20 | 19 | 18 |
| 12_c.52A>Tjkail | 3475 | 3851 | 3220 | 3048 | 2644 | 3045 | 3481 | 3326 | 3088 | 183 | 10079 | 2967 | 2613 | 3173 | 3107 | 3477 | 3141 | 2901 | 2394 | 2669 | 3137 | 2617 | 2685 | 3009 | 2903 |
| 12_c.52A>Tkkail | 268 | 271 | 230 | 232 | 209 | 225 | 244 | 257 | 232 | 2960 | 2172 | 224 | 178 | 228 | 256 | 276 | 212 | 213 | 179 | 185 | 233 | 180 | 224 | 200 | 280 |
| 13_1_c.59A>G | 8725 | 9773 | 8451 | 7991 | 7151 | 7026 | 9160 | 8616 | 7715 | 2555 | 715 | 8069 | 5754 | 6842 | 6846 | 9032 | 7656 | 6163 | 8874 | 7058 | 8412 | 8520 | 7413 | 7120 | 225 |
| 13_2_c.59A>G | 2228 | 2330 | 1877 | 2149 | 1888 | 1781 | 2293 | 2194 | 2034 | 242 | 12924 | 1945 | 1095 | 2059 | 2027 | 1774 | 1863 | 1850 | 1660 | 1604 | 2168 | 1604 | 1970 | 1770 | 23 |
| 14_c.79G>Ajkail | 22596 | 21159 | 20425 | 16856 | 17072 | 16052 | 21019 | 21241 | 18038 | 16623 | 16795 | 10811 | 17108 | 18279 | 18962 | 18616 | 16476 | 20190 | 16715 | 15989 | 22382 | 16638 | 19455 | 16170 | 18624 |
| 14_c.79G>Ajkail | 9364 | 9832 | 8309 | 8031 | 7350 | 6411 | 9806 | 9265 | 7448 | 8294 | 7467 | 17721 | 5833 | 8241 | 8991 | 9018 | 7887 | 7118 | 7437 | 6883 | 7821 | 6811 | 7152 | 7118 | 1086 |
| 15_c.84.85insCjkail | 5878 | 6152 | 5253 | 4964 | 4795 | 4582 | 5927 | 5583 | 4181 | 5710 | 4410 | 1988 | 525 | 5388 | 4969 | 6009 | 5061 | 4653 | 4069 | 4261 | 4887 | 4311 | 4312 | 4468 | 4620 |
| 15_c.84.85insCjkail | 54 | 59 | 46 | 44 | 51 | 46 | 55 | 50 | 38 | 44 | 45 | 33 | 2248 | 67 | 50 | 60 | 39 | 34 | 39 | 37 | 46 | 37 | 39 | 42 | 57 |
| 16_1_c.92+1G>T | 20573 | 20321 | 18349 | 12658 | 16266 | 16982 | 21497 | 18925 | 16927 | 18510 | 8620 | 17466 | 15085 | 5443 | 7235 | 16971 | 17706 | 18311 | 16971 | 16295 | 20013 | 10314 | 17728 | 15377 | 15324 |
| 16_2_c.92+1G>T | 1519 | 1543 | 1311 | 1443 | 1197 | 1198 | 1528 | 1433 | 1195 | 1295 | 1190 | 1327 | 660 | 14798 | 58 | 1548 | 1231 | 1285 | 993 | 1013 | 1329 | 981 | 1227 | 1071 | 1467 |
| 17_1_c.92+5G>C | 18143 | 18262 | 15981 | 12881 | 14656 | 14905 | 18069 | 18485 | 15656 | 15413 | 11141 | 15190 | 13642 | 5154 | 7048 | 17050 | 17157 | 12096 | 14669 | 14649 | 15406 | 13222 | 14927 | 15338 | 14633 |

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 17_2_c.92+5G>C | 361 | 364 | 317 | 283 | 292 | 285 | 370 | 341 | 241 | 315 | 287 | 321 | 168 | 27 | 18102 | 391 | 197 | 289 | 218 | 205 | 283 | 209 | 265 | 223 | 455 |
| 18_1_c.108C>A | 28815 | 26751 | 23997 | 25117 | 21054 | 16219 | 26933 | 25770 | 23959 | 24981 | 22745 | 22450 | 21808 | 23859 | 26370 | 3364 | 26131 | 23658 | 22180 | 21951 | 25795 | 22257 | 25747 | 24017 | 3430 |
| 18_2_c.108C>A | 9530 | 8861 | 7089 | 8229 | 6965 | 6857 | 8749 | 8332 | 7174 | 7707 | 6871 | 7468 | 6859 | 7544 | 8894 | 27039 | 8433 | 7662 | 6674 | 6624 | 8141 | 6859 | 7712 | 7616 | 40 |
| 22_1_c.170G>A | 44489 | 40949 | 32685 | 31346 | 30074 | 29301 | 36853 | 33581 | 29159 | 34889 | 31320 | 31218 | 29885 | 32018 | 36160 | 38387 | 3190 | 28560 | 27122 | 29880 | 31137 | 29637 | 30360 | 34614 | 31336 |
| 22_2_c.170G>A | 648 | 696 | 557 | 465 | 522 | 471 | 634 | 566 | 415 | 591 | 548 | 595 | 499 | 639 | 664 | 696 | 18645 | 413 | 389 | 384 | 466 | 393 | 447 | 420 | 885 |
| 23_1_c.216_217insA | 51526 | 48050 | 39041 | 41876 | 36450 | 32941 | 44280 | 42107 | 37509 | 40506 | 38641 | 36690 | 35353 | 38357 | 43289 | 51424 | 41621 | 9216 | 36476 | 35057 | 42262 | 38085 | 39458 | 41594 | 41954 |
| 23_2_c.216_217insA | 693 | 665 | 594 | 582 | 586 | 584 | 630 | 611 | 556 | 607 | 595 | 577 | 559 | 636 | 664 | 800 | 545 | 30894 | 736 | 478 | 609 | 526 | 591 | 578 | 806 |
| 24_1_c.251G>A | 73021 | 89958 | 62328 | 64337 | 58878 | 59951 | 67878 | 67139 | 59925 | 66339 | 58110 | 61318 | 59560 | 62759 | 62683 | 72883 | 61225 | 53731 | 32131 | 58739 | 75954 | 59929 | 62499 | 65083 | 32075 |
| 24_2_c.251G>A | 10187 | 10541 | 9702 | 8731 | 8611 | 9808 | 10000 | 9707 | 8607 | 9409 | 9196 | 9078 | 8896 | 9136 | 9866 | 10832 | 8802 | 8999 | 63983 | 7687 | 8849 | 8084 | 8513 | 8053 | 135 |
| 25_c.316-197C>Tjkail | 7220 | 12183 | 12580 | 5420 | 8189 | 6183 | 10752 | 10287 | 6432 | 8960 | 4710 | 7345 | 6987 | 9094 | 6904 | 4483 | 11438 | 4995 | 11353 | 1666 | 6636 | 8614 | 5535 | 5764 | 3236 |
| 25_c.316-197C>Tkkail | 4602 | 7823 | 6109 | 4000 | 4708 | 4111 | 7126 | 6018 | 4929 | 3351 | 3083 | 4353 | 3860 | 6222 | 4441 | 1820 | 7223 | 3276 | 6240 | 6098 | 4570 | 4948 | 3815 | 3425 | 2070 |
| 26_c.364G>Cjkail | 19594 | 26496 | 25820 | 17444 | 20801 | 19213 | 24646 | 22617 | 17515 | 23936 | 20040 | 21268 | 19157 | 24150 | 21584 | 17328 | 24748 | 11984 | 25580 | 22833 | 2906 | 21480 | 14338 | 17083 | 12602 |
| 26_c.364G>Ckkail | 283 | 337 | 319 | 228 | 283 | 296 | 318 | 304 | 230 | 294 | 292 | 317 | 289 | 325 | 304 | 278 | 249 | 222 | 268 | 233 | 18499 | 283 | 241 | 250 | 280 |
| 27_1_c.370_377delACCCCACC | 25089 | 31786 | 31546 | 16299 | 26634 | 27060 | 30531 | 28110 | 21154 | 28988 | 20088 | 25692 | 27409 | 20297 | 25859 | 21402 | 30440 | 19789 | 29638 | 25344 | 22189 | 18 | 902 | 20944 | 15373 |
| 27_2_c.370_377delACCCCACC | 148 | 164 | 158 | 114 | 145 | 150 | 157 | 145 | 119 | 150 | 152 | 151 | 140 | 157 | 156 | 142 | 120 | 108 | 129 | 115 | 121 | 20925 | 8 | 129 | 159 |
| 27_28probej(Wt-T) | 18905 | 23977 | 23327 | 16619 | 19777 | 19494 | 22704 | 22050 | 15961 | 19591 | 18366 | 19773 | 20116 | 21041 | 20124 | 15775 | 23247 | 13352 | 24350 | 20229 | 16921 | 31 | 137 | 16370 | 21515 |
| 27_28probek(Wt-T) | 8344 | 8669 | 8023 | 5773 | 6936 | 7119 | 8157 | 7834 | 5728 | 7541 | 6627 | 7208 | 7496 | 7661 | 6838 | 5604 | 7829 | 5201 | 8180 | 6767 | 5524 | 23 | 13626 | 6293 | 15880 |
| 29_1_c.410G>A | 35122 | 31793 | 42622 | 26902 | 34117 | 29675 | 36053 | 38323 | 28175 | 38669 | 31192 | 33985 | 34129 | 38668 | 35325 | 22187 | 41454 | 22218 | 41717 | 38885 | 20211 | 30957 | 28507 | 10596 | 21385 |
| 29_2_c.410G>A | 5059 | 8107 | 6027 | 4666 | 5445 | 5235 | 5957 | 5514 | 4425 | 5668 | 5260 | 5426 | 5510 | 5475 | 5489 | 4627 | 5406 | 4343 | 5394 | 4904 | 4403 | 6099 | 4426 | 26876 | 1954 |

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 30_1_c.441_442 insAC | 338 91 | 436 19 | 409 06 | 281 78 | 353 69 | 365 78 | 351 53 | 366 98 | 278 09 | 371 54 | 317 16 | 353 98 | 363 26 | 397 50 | 332 89 | 286 14 | 388 12 | 247 03 | 426 16 | 340 08 | 284 92 | 263 85 | 280 55 | 265 46 | 467 0 |
| 30_2_c.441_442 insAC | 35 | 42 | 44 | 26 | 31 | 44 | 41 | 33 | 24 | 41 | 31 | 38 | 32 | 35 | 31 | 37 | 33 | 25 | 40 | 30 | 24 | 29 | 25 | 27 | 184 03 |

Subsequently, the ratio of signal originating from the wild type probe/signal originating from the mutant probe was calculated for each probe pair, and then the ratio was converted to radian unit (left side of Table 7). Thereafter, among the data obtained by performing hybridization 25 times, the median value (radian) and the standard error (radian) of the Wild Type 24 data were calculated. The correction value C was computed by calculating the value of ($\pi/2 \div$ median value of wild type), and the correction value $C_2$ was computed by calculating the value of ($\pi/2 \div (\pi/2$ - mutant)) (right side of Table 7).

Thereafter, regarding the error range, the standard error of the Wild Type 24 data was multiplied by the correction value C or the correction value $C_2$, and thereby the error range after correction was computed (right end of Table 7).

The data obtained before and after the correction using the upper limit or lower limit of the error range as the range of determination for the genotype, are presented in Fig. 9 (Fig. 9: before correction and after correction of the data obtained from 25 kinds of plasmid-derived samples).

Table 7Data, correction values and standard errors obtained from 25 kinds of plasmid-derived samples

| Site of mutation | Reference nucleic acid 1 | Reference nucleic acid 2 | Reference nucleic acid 3 | Reference nucleic acid 4 | Reference nucleic acid 5 | Reference nucleic acid 6 | Reference nucleic acid 7 | Reference nucleic acid 8 | Reference nucleic acid 9 | Reference nucleic acid 10 | Reference nucleic acid 11 | Reference nucleic acid 12 | Reference nucleic acid 13 | Reference nucleic acid 14 | Reference nucleic acid 15 | Reference nucleic acid 16 | Reference nucleic acid 17 | Reference nucleic acid 18 | Reference nucleic acid 19 | Reference nucleic acid 20 | Reference nucleic acid 21 | Reference nucleic acid 22 | Reference nucleic acid 23 | Reference nucleic acid 24 | Reference nucleic acid 25 | | Wild type data median value | Mutant data | Coefficient of correction C | Coefficient of correction C2 | Standard error from 24 data | Standard error 1 | Error range after correction (wild type) | Error range after correction (mutant) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Site of mutation detection 1 | 0.06 | 1.20 | 1.19 | 1.09 | 1.15 | 1.20 | 1.02 | 1.15 | 1.16 | 1.16 | 1.22 | 1.16 | 1.13 | 1.17 | 1.13 | 1.04 | 1.04 | 1.16 | 1.18 | 1.10 | 1.20 | 1.14 | 1.17 | 1.12 | 1.39 | Site of mutation detection 1 | 1.16 | 0.06 | 1.35 | 1.04 | 0.07 | 0.07 | 0.10 | 0.08 |
| Site of mutation detectio | 1.11 | 0.08 | 1.35 | 1.47 | 1.10 | 1.05 | 1.12 | 1.14 | 1.22 | 1.15 | 1.17 | 1.11 | 1.10 | 1.10 | 1.16 | 1.17 | 1.16 | 1.16 | 1.19 | 1.24 | 1.15 | 1.12 | 1.16 | 1.15 | 1.49 | Site of mutation detectio | 1.16 | 0.08 | 1.36 | 1.05 | 0.11 | 0.11 | 0.15 | 0.11 |

| | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Site of mutation detection 6 | 1.35 | 1.34 | 1.33 | 1.33 | 1.32 | 0.42 | 1.33 | 1.33 | 1.39 | 1.27 | 1.23 | 1.32 | 1.26 | 1.33 | 1.39 | 1.31 | 1.39 | 1.35 | 1.32 | 1.37 | 1.37 | 1.35 | 0.38 | 1.34 | 1.34 | Site of mutation detection 6 | 1.33 | 0.42 | 1.18 | 1.36 | 0.04 | 0.04 | 0.05 | 0.05 |
| Site of mutation detection 7 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 0.12 | 1.56 | 1.57 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.57 | 1.57 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | 1.56 | Site of mutation detection 7 | 1.56 | 0.12 | 1.00 | 1.08 | 0.00 | 0.00 | 0.00 | 0.00 |
| Site of mutation detection 8 | 1.52 | 1.53 | 1.53 | 1.52 | 1.52 | 1.52 | 1.53 | 0.27 | 1.51 | 1.52 | 1.52 | 1.53 | 1.51 | 1.53 | 1.53 | 1.53 | 1.54 | 1.53 | 1.54 | 1.53 | 1.52 | 1.53 | 1.52 | 1.53 | 1.51 | Site of mutation detection 8 | 1.53 | 0.27 | 1.03 | 1.21 | 0.01 | 0.01 | 0.01 | 0.01 |
| Site of | 1.56 | 1.56 | 1.57 | 1.56 | 1.56 | 1.57 | 1.57 | 1.56 | 0.13 | 1.56 | 1.56 | 1.56 | 1.57 | 1.56 | 1.56 | 1.56 | 1.57 | 1.56 | 1.56 | 1.57 | 1.56 | 1.57 | 1.57 | 1.56 | 1.57 | Site of | 1.56 | 0.13 | 1.00 | 1.09 | 0.00 | 0.00 | 0.00 | 0.00 |

| | | |

| | | |

| | | |

| | | |

| | | |

on
de
tec
tio
n
25
on
de
tec
tio
n
25

Apart from the present investigation, ECACC Ethnic Diversity DNA Panels (EDP-1) (Sigma Catalogue No: 07020701) were purchased, and one sample was subjected to an analysis of the base sequence using a sequencer. It was found that the sample was a sample having a heterogeneous mutation at Mutation Site 12. Thus, subsequently, the data of the DNA chips were obtained by the same method as the method described in Example 2, using this sample, and the signal intensities shown in Table 8 were obtained.

Subsequently, the signal intensities of Table 8 were used to calculate the ratio of (signal originating from wild type probe)/(signal originating from mutant probe) for each probe pair, and the resultant value was converted to radian unit, and then further multiplied by the correction value C in Table 7. The results are presented in Table 9. Furthermore, the data of Table 9 were superimposed on the graph for the data after correction of Fig. 9 (Fig. 10: results obtained by superimposing the data of Table 9 on the graph for the data after correction of Fig. 9), and only for Mutation Site No. 12, the data were plotted in the space indicated between the error bar of the wild type and the error bar of the mutant error bar. Thus, it could be determined that the mutation was heterozygous.

Table 8DNA chip signal values originating from sample in which site of mutation 12 is heterozygous

| Probe | Signal intensity |
|---|---|
| 1_1_c.-137C>A | 1909 |
| 1_2_c.-137C>A | 807 |
| 2_c.-81A>G① | 8805 |
| 2_c.-81A>G② | 3997 |
| 3_1_c.-80T>C | 7676 |
| 3_2_c.-80T>C | 1228 |
| 4_1_c.-78A>G | 5293 |

| | |
|---|---|
| 4_2_c.-78A>G | 799 |
| 5_1_c.2T>G | 9591 |
| 5_2_c.2T>G | 5016 |
| 6_1_c.5T>C | 14262 |
| 6_2_c.5T>C | 2827 |
| 7_1_71c.19G>A | 6619 |
| 7_2_c.19G>A | 38 |
| 10_c.27_28insG①kail | 13476 |
| 10_c.27_28insG②kail | 715 |
| 11_c.46delT①kail | 5258 |
| 11_c.46delT②kail | 28 |
| 12_c.52A>T①kail | 4699 |
| 12_c.52A>T②kail | 347 |
| 13_1_c.59A>G | 12209 |
| 13_2_c.59A>G | 2773 |
| 14_c.79G>A①kail | 22830 |
| 14_c.79G>A②kail | 14559 |
| 15_c.84_85insC①kail | 5048 |
| 15_c.84_85insC②kail | 51 |
| 16_1_c.92+1G>T | 27551 |
| 16_2_c.92+1G>T | 2101 |
| 17_1_c.92+5G>C | 23286 |
| 17_2_c.92+5G>C | 457 |
| 18_1_c.108C>A | 35463 |
| 18_2_c.108C>A | 11120 |
| 22_1_c.170G>A | 48518 |
| 22_2_c.170G>A | 703 |
| 23_1_c.216_217insA | 57316 |
| 23_2_c.216_217insA | 504 |
| 24_1_c.251G>A | 82857 |
| 24_2_c.251G>A | 12711 |
| 25_c.316-197C>T①kail | 4843 |

| | |
|---|---|
| 25_c.316-197C>T②kail | 3308 |
| 26_c.364G>C①kail | 22720 |
| 26_c.364G>C②kail | 304 |
| 27_1_c.370_377delACCCCACC | 30203 |
| 27_2_c.370_377delACCCCACC | 131 |
| 27_28 probe ① (Wt-T) | 22188 |
| 27_28 probe ② (Wt-G) | 7791 |
| 29_1_c.410G>A | 41279 |
| 29_2_c.410G>A | 5540 |
| 30_1_c.411_442insAC | 41848 |
| 30_2_c.411_442insAC | 41 |

Table 9Data obtained by correcting data of Table 8

| | Radian (before correction) | Coefficient of correction C | Radian (after correction) |
|---|---|---|---|
| Site of mutation detection 1 | 1.17 | 1.35 | 1.59 |
| Site of mutation detection 2 | 1.14 | 1.36 | 1.56 |
| Site of mutation detection 3 | 1.41 | 1.09 | 1.54 |
| Site of mutation detection 4 | 1.42 | 1.10 | 1.57 |
| Site of mutation detection 5 | 1.09 | 1.37 | 1.49 |
| Site of mutation detection 6 | 1.38 | 1.18 | 1.62 |
| Site of mutation detection 7 | 1.56 | 1.00 | 1.57 |
| Site of mutation detection 8 | 1.52 | 1.03 | 1.56 |
| Site of mutation detection 9 | 1.57 | 1.00 | 1.57 |
| Site of mutation detection 10 | 1.50 | 1.05 | 1.57 |
| Site of mutation detection 11 | 1.35 | 1.19 | 1.60 |
| Site of mutation detection 12 | 1.00 | 1.35 | 1.36 |
| Site of mutation detection 13 | 1.56 | 1.01 | 1.57 |
| Site of mutation detection 14 | 1.49 | 1.05 | 1.56 |

| | | | |
|---|---|---|---|
| Site of mutation detection 15 | 1.55 | 1.01 | 1.57 |
| Site of mutation detection 16 | 1.27 | 1.24 | 1.58 |
| Site of mutation detection 17 | 1.56 | 1.01 | 1.57 |
| Site of mutation detection 18 | 1.56 | 1.01 | 1.58 |
| Site of mutation detection 19 | 1.42 | 1.10 | 1.56 |
| Site of mutation detection 20 | 0.97 | 1.57 | 1.52 |
| Site of mutation detection 21 | 1.56 | 1.01 | 1.57 |
| Site of mutation detection 22 | 1.57 | 1.00 | 1.57 |
| Site of mutation detection 23 | 1.23 | 1.28 | 1.58 |
| Site of mutation detection 24 | 1.44 | 1.11 | 1.59 |
| Site of mutation detection 25 | 1.57 | 1.00 | 1.57 |

INDUSTRIAL APPLICABILITY

According to the present invention, high-quality probes, a microarray having the same probes, and a method for evaluating the probes are provided.

SEQUENCE LISTING FREE TEXT

SEQ ID NOs:1 to 18: Probes

SEQ ID NOs:21 to 24: Primers

SEQ ID NOs:25 to 66: Probes

SEQUENCE LISTING

CLAIMS

1. A probe for detecting a polynucleotide sequence having one or more polymorphisms,

wherein the probe is hybridized to the sequence, and satisfies at least any one of the following requirements:

(1) the sequence contains one or more non-complementary bases at either end;

(2) the portion corresponding to the polymorphisms that are not targeted for detection, among the plural polymorphisms contained in the sequence, contains universal bases; and

(3) the polymorphism that is targeted for detection is located at a position six or fewer bases away from any one terminus of the probe.

2. A probe that is hybridized to a polynucleotide sequence having one or more polymorphisms, the polynucleotide sequence being a sequence in which the sum of the contents of guanine and cytosine in the sequence is 63% or more, and satisfies the following requirements (1) and/or (2):

(1) the sequence contains one or more non-complementary bases at either end; and

(2) the portion corresponding to the polymorphisms that are not targeted for detection, among the plural polymorphisms contained in the sequence, contains universal bases.

3. A probe that is hybridized to a polynucleotide sequence having one or more polymorphisms, the polynucleotide sequence being a sequence in which the sum of

the contents of guanine and cytosine in the sequence is 45% or less,

wherein the polymorphism that is intended for detection is located at a position six or fewer bases away from any one terminus of the probe.

4. The probe according to any one of claims 1 to 3, wherein the polynucleotide sequence having one or more polymorphisms is a human β-globin gene sequence.

5. The probe according to claim 2, wherein the sum of the contents of guanine and cytosine is 63% or more, and the polynucleotide sequence having one or more polymorphisms comprises a sequence set forth in SEQ ID NO:3, 4, 7, 8, 17 or 18.

6. The probe according to claim 3, wherein the sum of the contents of guanine and cytosine is 45% or less, and the polynucleotide sequence having one or more polymorphisms comprises a sequence set forth in SEQ ID NO:11 or 12.

7. A microarray comprising at least one of the sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17 and 18.

8. A probe group for detecting mutations in a β-globin gene, the probe group comprising genes having the sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17, 18 and 25 to 66.

9. A microarray comprising at least one of the sequences set forth in SEQ ID NOs:3, 4, 7, 8, 11, 12, 17, 18 and 25 to 66.

10. A β-thalassemia detection kit, comprising at least the probe according to any one of claims 1 to 9.

11. A β-thalassemia detection kit, comprising the microarray according to claim 7 or 9.

12. A kit for β-globin gene mutation detection, the kit comprising:

(a) (i) an oligonucleotide primer having the sequence set forth in SEQ ID NO:21 and an oligonucleotide primer having the sequence set forth in SEQ ID NO:22, and/or (ii) an oligonucleotide primer having the sequence set forth in SEQ ID NO:23 and an oligonucleotide primer having the sequence set forth in SEQ ID NO:24; and

(b) the microarray according to claim 7 or 9.

13. A method for evaluating a microarray probe pair for polymorphism detection, the method comprising the following steps:

(1) plotting the fluorescence coordinates obtained by hybridizing a control nucleic acid for first polymorphism with a probe pair for polymorphism detection, in a fluorescence coordinate system which includes a Y-axis representing the signal intensity obtainable when the probe for first polymorphism detection is hybridized, and an X-axis representing the signal intensity obtainable when the probe for second polymorphism detection is hybridized;

(2) defining a value which is inversely proportional to the gradient of a straight line that passes through the intersection O between the Y-axis and the X-axis and the fluorescence coordinates plotted in step (1), as a correction value C; and

(3) a step of carrying out steps (1) and (2) on plural probe pairs for polymorphism detection, comparing the correction values C between the various probes, and determining a probe pair having the minimum correction value C as probes appropriate for first polymorphism detection.

14. The method according to claim 13, wherein the fluorescence coordinate system has the Y-axis and the X-axis perpendicularly intersecting each other.

15. The method according to claim 13 or 14, wherein

step (1) involves obtaining two or more points of fluorescence coordinates by performing hybridization between the control nucleic acid and the probe two or more times, and determining a representative value M for the various coordinates; and

in step (2), the straight line is a median straight line that passes through the intersection O and the representative value M.

16. The method according to claim 14 or 15, wherein

step (1) further involves a process of selecting a straight line having a difference in the gradient with the median straight line among plural straight lines that pass through the intersection O and the two or more points of fluorescence coordinates, and designating this as an error straight line; and

step (2) includes:

(a) a process of determining

correction value $C = \pi/2 \div \alpha$

from the angle $\alpha$ (radian) between the median straight line and the X-axis; and

(b) a process of determining

correction error angle $\theta'$ (radian) = $\theta$ (radian) × correction value C

from an error angle $\theta$ (radian) which is an angle formed by the median straight line and the error straight line.

17. The method according to claims 13 to 16, further comprising the following steps:

(4) plotting the fluorescence coordinates obtained by hybridizing a control nucleic acid for second polymorphism with a probe pair for second polymorphism detection;

(5) defining a value which is proportional to the gradient of a straight line that passes through the intersection O and the fluorescence coordinates plotted in step (4), as a correction value $C_2$; and

(6) performing steps (4) and (5) on plural probe pairs for second polymorphism detection, comparing the correction values $C_2$ between the various probes, and determining a probe pair having the minimum correction value $C_2$ as a probe appropriate for second polymorphism detection.

18. The method according to claim 17, wherein

step (4) involves obtaining two or more points of fluorescence coordinates by performing hybridization between the control nucleic acid for second polymorphism and the probe for second polymorphism detection two or more times, and determining a representative value $M_2$ for the various coordinates; and

in step (5), the straight line is a second median straight line that passes through the intersection O and the representative value $M_2$.

19. The method according to claim 18, wherein

step (4) further involves a process of selecting a straight line having a difference in the gradient with the second median straight line among plural straight lines that pass through the intersection O and the two or more points of fluorescence coordinates, and designating this as an error straight line; and

step (5) includes:

(c) a process of determining

correction value $C_2 = \pi/2 \div \beta$

from the angle $\beta$ (radian) between the second median straight line and the Y-axis; and

(d) a process of determining

second correction error angle $\theta_2'$ (radian) = $\theta_2$ (radian) × correction value $C_2$

from a second error angle $\theta_2$ (radian) which is an angle formed by the second median straight line and the error straight line.

20. A genotype discrimination display program utilizing the method according to any one of claims 13 to 19.