

[DESCRIPTION]

[Title of Invention] METHOD FOR DETECTING SPECIFIC DNA
IN SAMPLE

[Technical Field]

5 The present invention relates to a method for
detecting a specific DNA in a sample by utilizing a
CRISPR-Cas3 system and a single-stranded probe DNA.

[Background Art]

10 The genome editing technologies are technologies
to specifically cleave genomic DNA sequences in cells
of animals and plants and to rewritten the sequences
into any desired sequences freely by utilizing the
inherent repair mechanisms. The utilization of the
genome editing technologies has been expanded not only
15 to the bioscience research but also to cultivar
improvements of agricultural crops and livestock,
regenerative medicines, and genome editing treatments
all over the world.

20 The CRISPR-Cas systems which bacteria and archaea
have are classified into Class 1 that cleaves target
sequences using a complex of multiple proteins and Class
2 that cleaves target sequences using a single protein.
The CRISPR-Cas9 system, the CRISPR-Cas12 (Cpf1) system,
the CRISPR-Cas13 system, and the like, which have been
25 developed as genome editing tools so far, are all
classified into Class 2.

Under such circumstances, the present inventors have found that the CRISPR-Cas3 system, which is a type I CRISPR-Cas system belonging to Class 1, can be used as a genome editing tool for eukaryotic cells (PTL 1).

5 However, the detailed molecular mechanism on how the CRISPR-Cas3 system cleaves DNAs and introduces mutations has not been revealed yet.

[Citation List]

[Patent Literature]

10 [PTL 1] International Publication No. WO2018/225858

[Summary of Invention]

[Technical Problem]

15 An object of the present invention is to clarify the molecular mechanism of DNA cleavage by the CRISPR-Cas3 system and to provide a method for detecting a target DNA using the molecular mechanism. A further object of the present invention is to provide a kit for use in the detection method.

[Solution to Problem]

20 The present inventors have made earnest studies to achieve the above objects, and have finally found that when the CRISPR-Cas3 system recognizes and binds to a target DNA in a sample, the CRISPR-Cas3 system cleaves single-stranded DNAs (ssDNAs) present
25 therearound at random. Moreover, the present inventors have found that based on this finding, by

mixing a single-stranded probe DNA whose cleavage can be detected in a reaction system containing a CRISPR-Cas3 system and a sample from which to detect a target DNA, it is possible to detect the target DNA in the sample using, as indication, a signal generated by the cleavage of the single-stranded probe DNA, and have completed the present invention.

Specifically, the present invention relates to a method for detecting a target DNA utilizing random cleavage of a single-stranded DNA by a CRISPR-Cas3 system, and a kit for use in the detection method, and more specifically provides the following.

(1) A method for detecting a specific DNA in a sample, comprising the steps of:

(a) contacting the sample with a CRISPR-Cas3 system targeting the specific DNA and a single-stranded probe DNA; and

(b) detecting cleavage of the single-stranded probe DNA by the CRISPR-Cas3 system that occurs in a case where the specific DNA is present in the sample.

(2) A kit for detecting a specific DNA in a sample by the method according to (1), comprising:

(a) a CRISPR-Cas3 system targeting the specific DNA; and

(b) a single-stranded probe DNA.

[Advantageous Effects of Invention]

The present invention makes it possible to detect a target DNA in a sample by utilizing a CRISPR-Cas3 system and a single-stranded probe DNA. The method of the present invention can detect a target DNA in a sample by using, as indication, a signal generated by the cleavage of the single-stranded probe DNA, and is simple and easy (see Fig. 10). Although only in the case where the CRISPR-Cas3 system correctly recognizes and binds to the target DNA, the CRISPR-Cas3 system cleaves the single-stranded probe DNA, the specificity in the detection of the target DNA is a single base level, and is significantly high. In addition, the method of the present invention can detect the target DNA even when only a trace amount of the target DNA is present in the sample. For example, by combining the method with a gene amplification method such as the RPA (recombinase polymerase amplification) method or the LAMP (loop-mediated isothermal amplification) method, it is possible to detect a target DNA even in a case where only 1 copy to 10 copies of the target DNA are present in the sample, thus achieving a high sensitivity.

The method of the present invention can be utilized as tools for next-generation diagnoses such as, for example, identification of viruses using various body fluids (urine, saliva, serum, plasma, whole blood, and the like) as samples in a short time, application

to liquid biopsy for detecting a trace amount of DNA and mutation thereof in bloods, or detection of a trace amount of pathogenic DNA contained in metagenome depending on the type of the pathogenic DNA, or the like.

As similar methods, DETECTR and Cas14-DETECTR using CRISPR-Cas12a and CRISPR-Cas14a, which belong to Class 2, have been reported (Chen et al., Science. 2018 Apr 27;360(6387): 436-439, Harrington et al., Science. 2018 Nov 16;362(6416): 839-842). While target sequences in these systems have 21 bases and 20 bases, respectively, the CRISPR-Cas3 system, which is a type I CRISPR-Cas system belonging to Class 1, has a longer target sequence of 27 bases. Therefore, the CRISPR-Cas3 system can detect target DNAs with higher specificity than the conventional systems.

[Brief Description of Drawings]

[Fig. 1] Fig. 1 is a diagram showing that a CRISPR-Cas3 system non-specifically cleaves the single-stranded DNA in the presence of a hEMX1 target sequence. The upper part of the diagram shows compositions of samples, and the lower part of the diagram shows the results of capillary electrophoresis. Non-specific single-stranded DNA cleavage activity was observed only in samples 4 and 5 in which DNA fragments containing appropriate PAM sequence and target sequence were

present (arrows).

[Fig. 2] Fig. 2 is a diagram showing that the CRISPR-Cas3 system does not non-specifically cleave the double-stranded DNA even in the presence of the hEMX1 target sequence. The upper part of the diagram shows compositions of samples, and the lower part of the diagram shows the results of capillary electrophoresis. Non-specific double-stranded DNA cleavage activity was not observed also in samples 4 and 5 in which DNA fragments containing appropriate PAM sequence and target sequence were present (arrows).

[Fig. 3] Fig. 3 is a diagram showing results of detecting that the CRISPR-Cas3 system non-specifically cleaved single-stranded DNA in the presence of hEMX1 target sequence by using a single-stranded DNA probe (DNaseAlert™) which emits fluorescence upon cleavage. The upper part of the diagram shows compositions of samples, and the lower part of the diagram shows the results of detecting fluorescence. Non-specific nuclease activity (fluorescence) with time was detected only in samples 4 and 5 in which DNA fragments containing appropriate PAM sequence and target sequence were present.

[Fig. 4] Fig. 4 is a diagram showing results of detecting that the CRISPR-Cas3 system non-specifically cleaved single-stranded DNA in the presence of mTyr

target sequence by using a single-stranded DNA probe (DNaseAlert™) which emits fluorescence upon cleavage. The upper part of the diagram shows compositions of samples, and the lower part of the diagram shows the results of detecting fluorescence. Non-specific nuclease activity (fluorescence) with time was detected only in samples 4 and 5 in which DNA fragments containing appropriate PAM sequence and target sequence were present.

[Fig. 5A] Fig. 5A is a diagram showing results of detecting that the CRISPR-Cas3 system non-specifically cleaved single-stranded DNA in the presence of hEMX1 target sequence by using a single-stranded DNA probe (Taqman™ probe) which emits fluorescence upon cleavage. The upper part of the diagram shows compositions of samples, and the lower part of the diagram shows the results of detecting fluorescence. Non-specific single-stranded DNA cleavage activity (fluorescence) with time was detected only in samples 2 and 3 in which DNA fragments containing appropriate PAM sequence and target sequence were present.

[Fig. 5B] Fig. 5B is a graph representing the degree of increase in fluorescence intensity in Fig. 5A as the intensity of the activity.

[Fig. 6] Fig. 6 is a diagram showing results of detecting that the CRISPR-Cas3 system non-specifically

cleaved single-stranded DNA in the presence of hEMX1 target sequence or mTyr target sequence by using a probe (Taqman™ probe) which emits fluorescence upon cleavage. In the present experiment, single-stranded DNA was used as the target sequence. The upper part of the diagram shows compositions of samples, and the lower part of the diagram represents the degree of increase in fluorescence intensity as the intensity of the activity. It was possible to induce non-specific single-stranded DNA cleavage as in the case of double-stranded DNA fragment even when the target sequence recognized by the CRISPR-Cas3 was single-stranded DNA fragment. N=3.

[Fig. 7] Fig. 7 is a diagram showing result of evaluating the same experiment as in Fig. 5 with various PAM sequences (64 types). The upper part of the diagram shows the results of measuring the intensity of the activity, and the lower part of the diagram visualizes the directivity of PAM sequence by using the Sequence logo. N=3.

[Fig. 8] Fig. 8 is a diagram showing results of evaluating the detection limit concentration of the CRISPR-Cas3. DNA fragments containing target sequences (the upper part of the diagram is hEMX1, and the lower part of the diagram is mTyr) were diluted through limiting dilution, and the same experiment as in Fig. 5 was conducted. Signals were detected significantly

in the presence of 1×10^{10} copies or more. $N=3$.

[Fig. 9] Fig. 9 is a diagram showing results of detecting 1 copy of the DNA fragment in the same experiment as in Fig. 5. In order to detect 1 copy of the DNA fragment in samples, DNA amplification by the RPA method was used. $N=3$.

[Fig. 10] Fig. 10 is a diagram showing the overview of the detection method of the present invention utilizing the non-specific single-stranded DNA cleavage activity by the CRISPR-Cas3 system.

[Fig. 11] Fig. 11 is a diagram showing results of conducting the same experiment as in Fig. 9 by using mixed genome solutions containing the target DNA fragment and mouse genomic DNA.

[Fig. 12] Fig. 12 is a diagram showing results of conducting the same experiment as in Fig. 9 by using genomes of the novel coronavirus amplified by the RT-LAMP method.

[Fig. 13A] Fig. 13A is a diagram showing the principle of detection of the lateral flow assay utilized in the detection of the novel coronavirus.

[Fig. 13B] Fig. 13B is a diagram showing results of detecting the novel coronavirus by using the lateral flow assay of Fig. 13A.

[Fig. 14] Fig. 14 is a diagram showing results of conducting the lateral flow assay on samples of serially

diluted virus RNA in the same manner as in Fig. 13B.

[Fig. 15] Fig. 15 is a diagram showing flow of the method (CONAN method) of the present invention for detecting an RNA virus with the lateral flow assay.

5 [Fig. 16] Fig. 16 is a diagram showing results of detecting Xofluza-resistant mutation sites of the type A influenza viruses (H1N1 and H3N2) by using the Taqman™ probe. The genes with wild-type mutation sites were detected.

10 [Fig. 17] Fig. 17 is a diagram showing results of detecting Tamiflu-resistant mutation sites of the type A influenza viruses (H1N1 and H3N2) by using the Taqman™ probe. The genes with mutated-type mutation site were detected.

15 [Description of Embodiments]

<Method for Detecting Specific DNA in Sample>

The present invention provides a method for detecting a specific DNA in a sample utilizing the CRISPR-Cas3 system and a single-stranded probe DNA.

20 The detection method of the present invention comprises the steps of: (a) contacting the sample with a CRISPR-Cas3 system targeting the specific DNA and a single-stranded probe DNA; and (b) detecting cleavage of the single-stranded probe DNA by the CRISPR-Cas3
25 system that occurs in a case where the specific DNA is present in the sample.

-Target DNA-

In the detection method of the present invention, the specific DNA to be detected (hereinafter, referred to as "target DNA") is not particularly limited as long as the target DNA serves as a target of the CRISPR-Cas3 system. The target DNA may be a double-stranded DNA or may be a single-stranded DNA (see Fig. 6).

The source of the target DNA is not particularly limited, and the target DNA may be a DNA derived from the natural world or may be an artificial DNA. The DNA derived from the natural world includes, for example, DNAs derived from viruses (for example, animal viruses, plant viruses, and bacterial viruses), DNAs derived from prokaryotic organisms (for example, bacteria, actinomycetes, prokaryotic algae, and archaea), and DNAs derived from eukaryotic organisms (for example, animals, plants, eukaryotic algae, fungi, and protozoa).

An example of one preferable embodiment of the detection method of the present invention is diagnosis of viral infections. The virus from which the target DNA to be detected is derived is not particularly limited, and may be a DNA virus or may be an RNA virus. In the case of an RNA virus, the target DNA can be obtained by conducting reverse transcription on the RNA virus. Specific viruses include, for example, retroviruses, picornaviruses, caliciviruses,

togaviruses, flaviviruses, coronaviruses (for example, novel coronavirus "SARS-CoV2"), rhabdoviruses, filoviruses, paramyxoviruses, orthomyxoviruses, bungaviruses, arenaviruses, reoviruses, birnaviruses, 5 hepadnaviruses, parvoviruses, papovaviruses, adenovirus, herpesviruses, poxviruses, iridoviruses, geminiviruses, phycodnaviruses, and the like, but the virus is not limited to these.

An example of another preferable embodiment of the 10 detection method of the present invention is diagnosis of pathogenic infections and protozoa infections. The pathogenic bacteria and the protozoa from which the target DNA to be detected is derived include, for example, shigella, Salmonella Typhi, Salmonella, 15 pathogenic Escherichia coli, Vibrio cholerae, Vibrios, Campylobacter, Clostridium welchii, Bacillus cereus, Yersinia, Staphylococcus aureus, Entamoeba histolytica, Leishmania, malaria, Trypanosoma, Toxoplasma gondii, Giardia lamblia, Trichomonas, Balantidium, Coccidia, 20 and the like, but the pathogenic bacteria or the protozoa are not limited to these.

An example of another preferable embodiment of the detection method of the present invention is diagnosis of mutation and polymorphism of genes. For example, it 25 is known that the mutation of viral genes can be utilized as indication for drug resistance and the like

of the virus. For example, I38T mutation (ATA→AGA) of the RNA polymerase gene (hereinafter, referred to as "PA gene") of type A influenza virus H1N1 and I38T mutation (ATA→ACA) of the PA gene of type A influenza virus H3N2 relate to the resistance to Xofluza, I222R mutation (ATA→AGA), H274Y mutation (CAC→TAC), and N294S mutation (AAC→AGC) of the neuraminidase gene (hereinafter, referred to as "NA gene") of type A influenza virus H1N1 as well as E119V mutation (GAA→GTA), R292K mutation (AGA→AAA), and N294S mutation (AAC→AGC) of the NA gene of type A influenza virus H3N2 relate to the resistance to Tamiflu and Relenza (Pizzorno et al., Front. Immunol., 2019 March 19, doi.10.3389/fimmu.2019.00531). Hence, it is possible to evaluate the resistances to virus treatment drugs by detecting mutations of these viral genes by using the method of the present invention.

In addition, mutation and polymorphism (for example, single-nucleotide polymorphism) of genes of human and the like can be utilized, for example, as indication for the risk of development of diseases, the efficacies and the side effects of drugs, and the like.

In the present invention, in a case where 10^{10} copies or more of the target DNA are present in a sample, it is possible to favorably detect the target DNA (see Fig. 8). However, even in a case where only 1 copy to

10 copies of the target DNA are contained in the initial sample, it is possible to detect the target DNA by amplifying the DNA in the sample (see Figs. 9, 11, and 12).

5 The DNA in the sample can be amplified before the sample is contacted with the CRISPR-Cas3 system or during the contact. In addition, the amplification of DNA may be amplification specific to the target DNA (or a specific region thereof) or may be amplification oriented to the target DNA (or a specific region thereof).

10 The methods for amplifying DNA are known to a person skilled in the art, and various methods can be used. The methods include, for example, various PCR (polymerase chain reaction) methods (the PCR method, 15 the RT-PCR method, the quantitative PCR method, and the like), the RPA (recombinase polymerase amplification) method, various LAMP (loop-mediated isothermal Amplification) methods (the LAMP method, the RT-RAMP method, the quantitative LAMP method, and the like), 20 the HDA (heliCase-dependent Amplification) method, the SDA (strand displacement amplification) method, the NASBA (nucleic acid sequence-based amplification) method, the TMA (transcription mediated amplification) method, 25 the NEAR (nicking enzyme amplification reaction) method, the MDA (multiple displacement

amplification) method, the RAM (Ramification) method, and the like; however, the method is not limited to these.

-Test Sample-

5 As the sample in the detection method of the present invention, any desired sample for which the presence of a target DNA is desired to be detected (hereinafter, referred to as a "test sample") can be used. The test sample includes, for example, samples
10 containing body fluids (urine, saliva, serum, plasma, whole blood, and the like), body tissues, cells, cell lysates, and purified or synthesized DNAs.

 In the case where infections of viruses, pathogenic bacteria, or protozoa are diagnosed, the
15 test samples may be derived from human patients, animals, plants, or the like suspected to have these infections, for example.

 The animal from which the test sample is derived includes, for example, mammals, fishes, birds, reptiles,
20 amphibians, and insects. The mammals are a concept that encompasses human and non-human mammals. Examples of non-human mammals include cloven-hoofed animals such as cattle, boars, pigs, sheep, and goats, odd-toed ungulates such as horses, rodents such as mice, rats,
25 guinea pigs, hamsters, and squirrels, lagomorphs such as rabbits, carnivores such as dogs, cats, and ferrets,

and the like. The non-human mammals may be livestock or companion animals (pets), or may be wild animals. The animal cells include, for example, cells constituting individuals of animals, cells constituting organs and tissues excised from animals, cultured cells derived from tissues of animals, and the like. Specifically, the animal cells include, for example, germ cells such as oocytes and sperm; embryonic cells of embryos at various stages (for example, 1-cell embryos, 2-cell embryos, 4-cell embryos, 8-cell embryos, 16-cell embryos, morula embryos, and the like); stem cells such as induced pluripotent stem (iPS) cells and embryonic stem (ES) cells; somatic cells such as fibroblasts, hematopoietic cells, neurons, muscle cells, bone cells, liver cells, pancreatic cells, brain cells, and kidney cells.

The plant from which the test samples are derived includes, for example, grains, oil crops, forage crops, fruits, and vegetables, and specifically include, for example, rice, corn, banana, peanut, sunflower, tomato, turnip rape, tobacco, wheat, barley, potato, soybean, cotton, carnation, and the like. The plant cells include, for example, cells constituting individuals of plants, cells constituting organs and tissues separated from plants, cultured cells derived from tissues of plants, and the like. The organs and tissues of plants

include, for example, leaves, stems, shoot apices (growing points), roots, tubers, calli, propagation material (for example, seeds, tuberous roots, tubers, and the like).

5 -CRISPR-Cas3 System-

10 The Class 1 CRISPR-Cas systems are classified into type I and type III, and type I is further classified into six types of type I-A, type I-B, type I-C, type I-D, type I-E, and type I-F as well as type I-G, which is
15 a subtype of type I-B, depending on the types of proteins constituting the Cascade (hereinafter, simply referred to as "Cascade" or "Cascade protein") (for example, see [van der Oost J et al., Nature Reviews Microbiologym, 2014 12(7): 479-492], [Jackson RN et al.,
20 Current Opinion in Structural Biology, 2014 24: 106-114])).

20 The type I CRISPR-Cas systems have the function of cleaving DNA by using cooperation of Cas3 (protein having nuclease activity and helicase activity), Cascade, and crRNA. Since Cas3 is used as a nuclease, these are referred to as the "CRISPR-Cas3 system" in the present invention.

25 Use of the CRISPR-Cas3 system of the present invention makes it possible to obtain, for example, the following advantages.

First, the crRNA used in the CRISPR-Cas3 system

generally recognizes a target sequence of 32 to 37 bases
(Ming Li et al., Nucleic Acids Res. 2017 May 5; 45(8):
4642-4654). On the other hand, the crRNA used in the
CRISPR-Cas9 system generally recognizes a target
5 sequence of 18 to 24 bases. Therefore, it is considered
that the CRISPR-Cas3 system can recognize target
sequences more accurately than the CRISPR-Cas9 system.

In addition, the PAM sequence of the CRISPR-Cas9
system, which is the Class 2 type II system, is "NGG (N
10 is an arbitrary base)" adjacent to the 3' side of the
target sequence. In addition, the PAM sequence of the
CRISPR -Cpf1 system, which is the Class 2 type V system,
is "AA" adjacent to the 5' side of the target sequence.
On the other hand, the PAM sequence of the CRISPR-Cas3
15 system of the present invention is "AAG" adjacent to
the 5' side of the target sequence or a nucleic acid
sequence similar to that (for example, "AGG", "GAG",
"TAC", "ATG", "TAG", or the like). Thus, using the
CRISPR-Cas3 system of the present invention makes it
20 possible to target DNA regions which cannot be
recognized by the conventional methods.

The CRISPR-Cas3 system of the present invention
encompasses all the six subtypes of type I. That is,
although proteins constituting the CRISPR-Cas3 system
25 may differ slightly in constitution and the like
depending on the subtypes (for example, proteins

constituting the Cascade are different), the present invention encompasses all of these proteins. In addition, the CRISPR-Cas3 system of the present invention is independent from the types of bacteria as the source.

The type I-E CRISPR-Cas3 system, which is common among the type I CRISPR-Cas3 systems, cleaves DNA by making a crRNA cooperate with Cas3 and Cascade (Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6, and Cas7).

The type I-A system has Cas8a1, Csa5 (Cas11), Cas5, Cas6, and Cas7 as constituent elements of the Cascade, the type I-B has Cas8b1, Cas5, Cas6, and Cas7 as constituent elements of the Cascade, the type I-C has Cas8c, Cas5, and Cas7 as constituent elements of the Cascade, the type I-D has Cas10 , Csc1 (Cas5), Cas6, and Csc2 (Cas7) as constituent elements of the Cascade, the type I-F has Csy1 (Cas8f), Csy2 (Cas5), Cas6, and Csy3 (Cas7) as constituent elements of the Cascade, and the type I-G system has Cst1 (Cas8a1), Cas5, Cas6, and Cst2 (Cas7) as constituent elements of the Cascade. In the present invention, Cas3 and Cascade are collectively referred to as the "Cas protein group".

Hereinafter, the type I-E CRISPR-Cas3 system will be described as a representative example. For the other types of CRISPR-Cas3 systems, the Cascade constituting the systems may be interpreted as appropriate.

-Cas Protein Group, Polynucleotide Encoding the protein group, Vector for Expressing the Polynucleotide-

In the detection method of the present invention,
5 the Cas protein group can be used in the form of a protein, in the form of a polynucleotide encoding the protein, or in the form of an expression vector containing the polynucleotide.

Preferable embodiments of the respective proteins
10 constituting the Cas protein group used in the present invention are the following proteins of wild type E. coli:

Cas3; a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 1 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 7)
15

Cse1 (Cas8); a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 2 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 8)

Cse2 (Cas11); a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 3 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 9)
20

Cas5; a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 4 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 10)

Cas6; a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 5 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 6)
25

acid sequence set forth in SEQ ID NO: 11)

Cas7; a protein composed of a nucleic acid sequence set forth in SEQ ID NO: 6 (a protein encoded by a nucleic acid sequence set forth in SEQ ID NO: 12).

5 Another preferable embodiment of each above protein constituting the Cas protein group used in the present invention is a protein composed of an amino acid sequence obtained by substituting, deleting, adding, and/or inserting one or multiple amino acids in
10 the amino acid sequence of the each protein. Here, "multiple" is normally 50 amino acids or less, preferably 30 amino acids or less, more preferably 20 amino acids or less, and particularly preferably 10 amino acids or less (for example, 5 amino acids or less,
15 3 amino acids or less, 2 amino acids or less, or 1 amino acid).

 Another preferable embodiment of each above protein constituting the Cas protein group used in the present invention is a protein composed of an amino
20 acid sequence having a high identity with the amino acid sequence of the each protein. The high identity means a sequence identity of, for example, 80% or more, preferably 85% or more, more preferably 90% or more (for example, 91% or more, 92% or more, 93% or more, or
25 94% or more), and further preferably 95% or more (for example, 96% or more, 97% or more, 98% or more, or 99%

or more). The sequence identity can be determined by using BLAST (Basic Local Alignment Search Tool at the National Center for Biological Information of the United States) or the like (for example, by using parameters of the default, that is, the initial setting).

Another preferable embodiment of each protein constituting the Cas protein group used in the present invention is a protein encoded by a polynucleotide that hybridizes under stringent conditions to a polynucleotide composed of a nucleic acid sequence complementary to the nucleic acid sequence of each above protein.

Here, "stringent conditions" mean conditions under which two polynucleotide chains form a double-stranded polynucleotide specific to a nucleic acid sequence but do not form a non-specific double-stranded polynucleotide. The phrase "hybridizes under stringent conditions" may be rephrased to conditions under which a polynucleotide can hybridize within a temperature range from a melting temperature (T_m value) of nucleic acids having high sequence identity (for example, hybrids that completely match) to a temperature 15°C lower, preferably a temperature 10°C lower, and more preferably a temperature 5°C lower, than the melting temperature.

The following is an example of the stringent

conditions. In a buffer solution (pH 7.2) composed of 0.25M Na_2HPO_4 , 7% SDS, 1 mM EDTA; and 1×Denhardt's solution, two types of polynucleotides are hybridized at 60 to 68°C (preferably 65°C, and more preferably 68°C) for 16 to 24 hours. Thereafter, in a buffer solution (pH 7.2) composed of 20 mM Na_2HPO_4 , 1% SDS, and 1 mM EDTA, washing for 15 minutes is conducted at 60 to 68°C (preferably 65°C, and more preferably 68°C).

The following method is another example. First, in a hybridization solution containing 25% formamide (under more stringent conditions, 50% formamide), 4×SSC (sodium chloride/sodium citrate), 50 mM Hepes (pH 7.0), 10×Denhardt's solution, and 20 µg/mL denatured salmon sperm DNA, prehybridization is conducted at 42°C overnight, then a labeled probe is added, and the temperature is maintained at 42°C overnight to hybridization two types of polynucleotides.

Next, washing is conducted under any of the following conditions. Normal conditions; washing is conducted at around 37°C using 1×SSC and 0.1% SDS as a cleaning liquid. Stringent conditions; washing is conducted at around 42°C using 0.5×SSC and 0.1% SDS as a cleaning liquid. Further stringent conditions; washing is conducted at around 65°C using 0.2×SSC and 0.1% SDS as a cleaning liquid.

In this way, the more stringent the washing

conditions for hybridization are, the higher the specificity of the hybridization is. Note that the above combinations of conditions on SSC, SDS, and the temperatures are mere examples. Stringency similar to the above can be achieved by combining the above-described elements to determine the stringency of hybridization, or other elements (for example, the concentration of the probe, the length of the probe, reaction time of the hybridization, and the like) as appropriate. This is described in, for example, [Joseph Sambrook & David W. Russell, Molecular cloning: a laboratory manual 3rd Ed., New York: Cold Spring Harbor Laboratory Press, 2001] and the like.

When each of the above proteins forms a complex with another molecule constituting the CRISPR-Cas3 system to recognize a target DNA, the protein brings non-specific cleavage activity of a single-stranded DNA to the complex (see Examples).

To each of the proteins constituting the above Cas protein groups, functional molecules may be further added, as necessary. The functional molecule includes, for example, nuclear localization signals for accelerating the transition into the nucleus of the eukaryotic cell, tags for facilitating the purification, reporter proteins for facilitating the detection, and the like, but is not limited to these. These functional

molecules can be added, for example, to the N terminal side and/or the C terminal of each protein constituting the Cas protein group.

The nuclear localization signals include, for example, PKKKRKV (SEQ ID NO: 13), KRTADGSEFESPKKKRKVE (SEQ ID NO: 14), and the like. The tags include, for example, a polyhistidine-tag, a FLAG-tag, glutathione-S-transferase (GST), and the like. The reporter proteins include, for example, fluorescent proteins such as a green fluorescent protein (GFP), chemiluminescent proteins such as luciferase, and the like.

In the present invention, a polynucleotide encoding the Cas protein group can be used, and preferable embodiments thereof are as follows:

Cas3; a polynucleotide composed of a nucleic acid sequence set forth in SEQ ID NO: 7

Cse1 (Cas8); a polynucleotide composed of a nucleic acid sequence set forth in SEQ ID NO: 8

Cse2 (Cas11); a polynucleotide composed of a nucleic acid sequence set forth in SEQ ID NO: 9

Cas5; a polynucleotide composed of a nucleic acid sequence set forth in SEQ ID NO: 10

Cas6; a polynucleotide composed of a nucleic acid sequence set forth in SEQ ID NO: 11

Cas7; a polynucleotide composed of a nucleic acid

sequence set forth in SEQ ID NO: 12

These are polynucleotides encoding the wild type Cas protein group of E. coli, but artificially modified polynucleotides can be used depending on the purpose.

5 The artificial modification of polynucleotides includes, for example, modification to a nucleic acid sequence suitable for expression in a host cell or an in vitro expression system (for example, optimization of codons).

In the present invention, expression vectors for

10 expressing the Cas protein group can be used. As the expression vector, various publicly-known vectors can be used. The expression vectors include, for example, phage vectors, plasmid vectors, virus vectors, retrovirus vectors, chromosome vectors, episome vectors,

15 and virus-derived vectors (bacterial plasmids, bacteriophages, yeast episomes, and the like), yeast chromosomal elements, and viruses (baculoviruses, papova viruses, vaccinia viruses, adenovirus, triplex viruses, pseudorabies viruses, herpesviruses,

20 lentiviruses, retroviruses, and the like), and vectors derived from combinations of these (cosmids, phagemids, and the like).

Preferably, the expression vector further contains sites for transcription initiation and transcription

25 termination as well as a ribosome binding site in the transcription region. The coding site of the mature

transcript in the vector contains the transcription initiation codon AUG at the beginning of the polypeptide to be translated and an appropriately located termination codon at the end of the polypeptide.

5 In the present invention, the expression vector for expressing the Cas protein group may contain a promoter sequence. The promoter sequence may be appropriately selected depending on the location of the gene expression (for example, type of the host). In
10 addition, the expression vector may contain a sequence for enhancing transcription from DNA, for example, an enhancer sequence. The enhancer includes, for example, the SV40 enhancer (which is arranged at 100 to 270 bp downstream of the replication origin), the early
15 promoter enhancer of the cytomegalovirus, the polyoma enhancer and the adenovirus enhancer arranged downstream of the replication origin. In addition, the expression vector may contain a sequence for stabilizing transcribed RNA, for example, a poly(A)
20 addition sequence (polyadenylated sequence, polyA). Examples of the poly(A) addition sequence include poly(A) addition sequences derived from the growth hormone gene, poly(A) addition sequences derived from the bovine growth hormone gene, poly(A) addition
25 sequences derived from the human growth hormone gene, poly(A) addition sequences derived from SV40 virus, and

poly(A) addition sequences derived from the β -globin gene of human or rabbit.

It is possible to design the polynucleotides encoding the Cas protein groups to be mounted on a single (the same) vector, and it is also possible to design all or part of the polynucleotides encoding the respective Cas protein groups to be mounted on separate vectors. For example, it is possible to employ a design that the polynucleotides encoding the Cascade protein are mounted on a single (the same) vector and the polynucleotides encoding Cas3 are mounted on different vectors.

In addition, it is possible to use an expression vector which contains multiple nucleic acid sequences encoding the Cas protein groups and which has nucleic acid sequences that are inserted between the multiple nucleic acid sequences and that encode amino acid sequences (2A peptides and the like) to be cleaved by intracellular proteases. When polynucleotides having such nucleic acid sequences are transcribed and translated, polypeptide chains linked together in the cell are expressed. Thereafter, due to the action of intracellular proteases, the Cas protein groups are separated, become separate proteins, and then form complexes to function. This makes it possible to regulate the amount ratio of Cas protein groups to be

expressed.

The expression vector used in the present invention can be prepared in accordance with a publicly-known method. Such a method includes the method
5 described in the implementation manuals attached to kits for preparing vectors as well as methods described in various handbooks. For example, [Joseph Sambrook & David W. Russell, Molecular cloning: a laboratory manual 3rd Ed., New York: Cold Spring Harbor Laboratory
10 Press, 2001] is a comprehensive handbook.

-crRNA, Polynucleotide Encoding the crRNA, or Expression Vector Containing the Polynucleotide-

The CRISPR-Cas3 system used in the present invention contains crRNA, a polynucleotide encoding the
15 crRNA, or an expression vector containing the polynucleotide, for targeting to a specific DNA.

crRNA is RNA that forms part of the CRISPR-Cas system, and has a nucleic acid sequence complementary to the target sequence in the target DNA. The CRISPR-Cas3 system makes it possible for the crRNA to
20 specifically recognize the target sequence and to cleave the sequence.

In the case where the functional complex of the CRISPR-Cas3 system is formed in cells, it is preferable
25 to use a pre-crRNA as the crRNA (WO 2018/225858 A).

The pre-crRNA typically has a structure of "leader

sequence-repeated sequence-spacer sequence-repeated
sequence (LRSR structure)" or "repeated sequence-spacer
sequence-repeated sequence (RSR structure)". The
leader sequence is an AT-rich sequence and functions as
5 a promoter to express the pre-crRNA. The repeated
sequence is a sequence repeating with a spacer sequence
in between, and the spacer sequence is a sequence
designed as a sequence complementary to the target
sequence in the target DNA (originally, it is a sequence
10 derived from a foreign DNA incorporated in the course
of adaptation). The pre-crRNA becomes a mature crRNA
when cleaved by proteins constituting the Cascade (for
example, Cas6 for type I-A, B, D to E, and Cas5 for
type I-C).

15 Typically, the chain length of the leader sequence
is 86 bases, and the chain length of the repeated
sequence is 29 bases. The chain length of the spacer
sequence is, for example, 10 to 60 bases, preferably 20
to 50 bases, more preferably 25 to 40 bases, and
20 typically 32 to 37 bases. Hence, in the case of the
LRSR structure, the chain length of the pre-crRNA used
in the present invention is, for example, 154 to 204
bases, preferably 164 to 194 bases, more preferably 169
to 184 bases, and typically , 176 to 181 bases. In
25 addition, in the case of the RSR structure, the chain
length of the pre-crRNA is, for example, 68 to 118 bases,

preferably 78 to 108 bases, more preferably 83 to 98 bases, and typically , 90 to 95 bases.

In order to make the complex of the CRISPR-Cas3 system be formed and function in cells, it is considered that the process by which the repeated sequences of the pre-crRNA is cleaved by the proteins constituting the Cascade is important. Thus, it should be understood that the repeated sequence may be shorter or longer than the above chain length as long as such cleavage takes place. That is, it can be said that the pre-crRNA is crRNA formed by adding sequences sufficient for cleavage by proteins constituting the Cascade to both ends of the mature crRNA described below.

On the other hand, the mature crRNA generated by cleavage of the pre-crRNA has a structure of "5' handle sequence-spacer sequence-3' handle sequence". Typically, the 5' handle sequence is composed of 8 bases from the 22nd to 29th of the repeated sequence and is held in Cas5. In addition, typically, the 3' handle sequence is composed of 21 bases from the 1st to 21st of the repeated sequence, forms a stem loop structure with the 6th to 21st bases, and is held in Cas6. Hence, the chain length of the mature crRNA is normally 61 to 66 bases. However, depending on the type of the CRISPR-Cas3 system, there are mature crRNAs that do not have 3' handle sequences. In this case, the chain length is

21 bases shorter.

Note that the sequence of RNA may be designed as appropriate depending on the target sequence. In addition, RNA can be synthesized by using any method known in the art.

-Single-stranded Probe DNA-

The single-stranded probe DNA used in the detection method of the present invention is not particularly limited as long as the cleavage of the single-stranded probe DNA can be detected. The single-stranded probe DNA may be linear or may be cyclic, but is preferably linear (see Fig. 1). Nucleic acids constituting the single-stranded probe DNA may contain one or more modifications (for example, base modification, backbone modification, or sugar modification).

One preferable embodiment of the label bound to the single-stranded probe DNA is a fluorescent dye/quencher pair. In this embodiment, in a state where the fluorescent dye/quencher pair is bound and close to the single-stranded probe DNA, a signal from the fluorescent dye is reduced or eliminated. When the single-stranded probe DNA is cleaved and the fluorescent dye is separated from the quencher, sufficient signals from the fluorescent dye are detected. Hence, it is possible to detect the target

DNA in a test sample by using, as indication, the fluorescence signal generated from the single-stranded probe DNA cleaved by the CRISPR-Cas3 system that has recognized and bound to the target DNA.

5 The binding of the fluorescent dye and quencher to the single-stranded probe DNA is not particularly limited as long as they are close enough to cause quenching and when the single-stranded probe DNA is cleaved, they are separated. For example, it is possible to bind the fluorescent dye to one of the terminals of the single-stranded probe DNA and the quencher to the other terminal. In order to enhance the quenching action in the probe, another quencher may be bound. In this case, a double-quencher probe in which an additional quencher (internal quencher) is bound to the inside of the single-stranded probe DNA may be used.

20 The fluorescent dye used in the fluorescent dye/quencher pair includes, for example, ATTO dyes, cyanine dyes (for example, Cy3 and Cy5), tetramethylrhodamine (for example, TRITC), carboxyfluorescein (FAM), tetrachlorofluorescein (TET), hexachlorofluorescein (HEX), Texas Red, Yakima Yellow, and the like, and specific examples of the quencher include, for example, dark quencher, BHQ (Black Hole Quencher), IBFQ (Iowa Blak FQ), IBRQ (Iowa Blak RQ),

Eclipse, and the like, but the fluorescent dye and the quencher are not limited to these. As the internal quencher, for example, ZEN and TAO can be favorably used. The use of a fluorescent dye/quencher pair for
5 detecting cleavage of probe DNA is known to a person skilled in the art (for example, see International Publication No. WO2019/104058 (WO 2019/104058 A)).

Another preferable embodiment of the label bound to the single-stranded probe DNA is a donor/acceptor
10 pair for fluorescence resonance energy transfer (FRET). In this embodiment, in a state where the donor/acceptor pair is bound and close to the single-stranded probe DNA, excitation of the donor causes excitation and luminescence of the acceptor (that is, a FRET signal is
15 generated). When the single-stranded probe DNA is cleaved and the donor and the acceptor are separated, the FRET signal is reduced or eliminated. Hence, it is possible to detect the target DNA in the test sample by using, as indication, reduction or elimination of the
20 FRET signal in the single-stranded probe DNA cleaved by the CRISPR-Cas3 system that has recognized and bound to the target DNA.

The binding of the donor and the acceptor for FRET to the single-stranded probe DNA is not particularly
25 limited as long as they are close enough to cause FRET and when the single-stranded probe DNA is cleaved, they

are separated. For example, it is possible to bind the donor to one of the terminals of the single-stranded probe DNA and the acceptor to the other terminal.

The donor/acceptor pair for FRET includes, for example, BFP/eGFP, BFP/YFP, BFP/DsRed2, CFP/YFP, CFP/DsRed2, Midoriishi-Cyan/Kurabira-Orange, eGFP/DsRed, eGFP/Rhod-2, FITC/TRITC, FITC/Rhod-2, FITC/Cy3, Alexa488/Alexa546, Alexa488/Alexa555, Alexa488/Cy3, YFP/TRITC, YFP/Cy3, Cy3/Cy5, Cy3/Cy5.5, fluorescein/tetramethylrhodamine, IAEDANS/FITC, IAEDANS/5-(iodoacetamide)fluorescein, fluorescein/fluorescein, EDANS/dansyl, tryptophan/dansyl, tryptophan/pyrene, dansyl/fluorescein, naphthalene/dansyl, pyrene/coumarin, phycoerythrin/Cy5, and the like, but is not limited to these. The use of a donor/acceptor pair for detecting cleavage of probe DNA is known to a person skilled in the art (for example, see WO 2019/104058 A).

In addition the detection method of the present invention using a single-stranded probe DNA can be applied to immunochromatography (lateral flow assay). In this embodiment, for example, a single-stranded probe DNA to which a first tag and a second tag are bound is used, and an antibody for a labeled second tag (hereinafter, referred to as "labeled antibody") is

used as a labeled body for capturing the single-stranded probe DNA to which the tags are bound. A molecule binding to the first tag and a molecule binding to the labeled antibody are fixed respectively to a first line of a strip for immunochromatography (line upstream in the strip) and a second line thereof (line downstream in the strip).

In this immunochromatography, in the case where the target DNA is not present in the sample, a complex of the single-stranded probe DNA and the labeled antibody is formed by means of the second tag of the single-stranded probe DNA (not cleaved), and the complex is captured by a molecule binding to the first tag fixed to the first line by means of the first tag of the single-stranded probe DNA (not cleaved), and thus a signal generated from the labeled antibody is detected on the first line.

On the other hand, in the case where the target DNA is present in the sample, the first tag and the second tag are separated by the cleavage of the single-stranded probe DNA. A complex of the second tag thus separated and the labeled antibody is formed, and the complex is captured on the second line by a molecule binding to the labeled antibody fixed to the second line, and thus a signal generated from the labeled antibody is detected on the second line. The complex

of the second tag and the labeled antibody is not holding the first tag and is not captured on the first line, and thus no signal is detected on the first line.

The binding of the first tag and the second tag to the single-stranded probe DNA is not particularly limited as long as the tags are separated when the single-stranded probe DNA is cleaved. For example, it is possible to bind the first tag to one of the terminals of the single-stranded probe DNA and the second tag to the other terminal.

The first tag and the molecule binding thereto include, for example, a combination of streptavidin and biotin. The second tag includes, for example, FITC. The label in the labeled antibody includes, for example, gold particles. The molecule binding to the labeled antibody includes, for example, protein A and a secondary antibody (an antibody binding to the labeled antibody). The principle of the detection of the cleavage of the single-stranded probe DNA using immunochromatography is known to a person skilled in the art (Gootenberg et al., Science 2018; 360: 439-444). As the test strip in the immunochromatography, a commercial product (for example, HybriDetect (Milenia Biotec GmbH)) can also be used.

In the detection of the cleavage of the single-stranded probe DNA, various methods other than that

described above can be used (for example, see WO 2019/104058 A).

-Contact of Test Sample, CRISPR-Cas3 System, and Single-stranded Probe DNA-

5 The test sample, the CRISPR-Cas3 system, and the single-stranded probe DNA in the detection method of the present invention can be contacted by mixing the test sample, the CRISPR-Cas3 system, and the single-stranded probe DNA, for example. In the case where the
10 test sample contains cells, a further operation for introducing the CRISPR-Cas3 system and the single-stranded probe DNA into the cells in the test sample can be included.

 In the contacting of the test sample and the
15 CRISPR-Cas3 system, it is possible to individually separate the crRNA and the Cas protein group and contact these with the test sample, or to prepare a complex of the crRNA and the Cas protein group in advance and contact the complex with the test sample. In the case
20 where the CRISPR-Cas3 system is introduced into cells in the test sample, the crRNA and/or the Cas protein group may be expressed in the cells.

 As the method for introducing the molecules constituting the CRISPR-Cas3 system into cells in the
25 form of polynucleotide or an expression vector containing the polynucleotide, a publicly-known method

such as an electroporation method, a calcium phosphate method, a liposome method, a DEAE dextran method, a microinjection method, cationic lipid-mediated transfection, electroporation, transduction, or infection using virus vectors can be used, for example. As the method for introducing molecules constituting the CRISPR-Cas3 system into cells in the form of proteins (or, a ribonucleoprotein complex), a publicly-known method such as microinjection, electroporation, or cationic lipid-mediated transfection can be used, for example. Such methods are described in many standard laboratory manuals such as "Leonard G. Davis et al., Basic methods in molecular biology, New York: Elsevier, 1986".

<Kit for Detecting Specific DNA in Sample>

A kit for detecting a specific DNA in a sample in the present invention comprises: (a) a CRISPR-Cas3 system targeting the specific DNA; and (b) a single-stranded probe DNA. The constituent elements of the kit of the present invention may be in a form in which they are separate from one another or in a form in which all or some of these are mixed.

In addition, the kit of the present invention may further contain, for example, a reference sample (each concentration), a control sample, a sample diluent, a reaction buffer solution, a cleaning liquid, and the

like depending of the embodiment of the detection method.
In addition, the kit of the present invention may
further contain, for example, a molecule binding to a
label (for example, a tag) (for example, a labeled
5 antibody binding to the tag), a reagent necessary for
detecting the label , and the like, depending on the
type of the label in the single-stranded probe DNA.
Moreover, the kit of the present invention may contain
an instruction manual.

10 [Examples]

The present invention will be described in further
detail based on Examples, but the present invention is
not limited to the following Examples.

A. Materials and Methods

15 (1) Preparation of Cas3 and Cascade Protein

In the present detection system, type I-E CRISPR
derived from E. coli was used. The Cas3 proteins were
purified using insect cells Sf9. A Cas3 sequence with
a His-tagged nuclear localization signal (bpNLS) was
20 incorporated into a Baculovirus gene expression vector,
and then baculoviruses were prepared using E. coli
strain DH10bac. The baculoviruses thus prepared were
caused to infect the insect cells Sf9 to express the
Cas3 proteins. Thereafter, all the proteins were
25 collected, and the Cas3 proteins were purified using
nickel column and gel filtration purification. In the

preparation of Cascade complexes with NLS containing crRNA, three types of His-tagged Cas11 expression plasmids (pCDFDuet-1), Cas5, Cas6, Cas7, Cas8, and Cas11 expression plasmids (pRSFDuet-1), and target crRNA expression plasmids (pACYCDuet-1) were prepared, all of which were introduced into E. coli BL21 and expressed. After all the proteins were collected, the Cascade complexes were purified using nickel column and gel filtration purification.

Note that as the target sequences, a sequence in human EMX1 gene (TGGCGCATTGCCACGNNNCAGGCCAATGGGGAGGACATCGATGTCACCTCCAA TGACTAG/SEQ ID NO: 15) and a sequence in mouse Tyr gene (GCATTACTATGTGTCNNNGGACACACTGCTTGGGGGCTCTGAAATATGGAGGG ACATTGA/SEQ ID NO: 16) were used. "NNN" was the PAM sequence used in each experiment. In addition, the sequences of the Cas protein groups are shown in SEQ ID NOs: 1 to 12 and the sequence of the nuclear localization signal (bpNLS) is shown in SEQ ID NO: 14.

(2) Measurement of Non-specific Cleavage Using Electrophoresis

Into a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 μM CoCl₂, and 2.5 mM ATP), Cas3 protein (5 ng/μL), Cascade complex (25 ng/μL), double-stranded DNA fragment containing the target sequence (2 ng/μL), and donor DNA for checking non-specific

cleavage (4.5 ng/ μ L) were mixed. This reaction solution was incubated at 37°C for 1 hour, and electrophoresis was conducted using a capillary electrophoresis apparatus "MultiNa" (Shimadzu Corporation). As the PAM sequences of the double-stranded DNA fragment, in addition to "AAG" and "ATG" which were recognizable by the CRISPR-Cas3 system, "CCA" which was not recognizable by the CRISPR-Cas3 system was used, and each was studied. As the donor DNAs, about 1 kb of linear single-stranded DNA, about 7 kb of cyclic single-stranded DNA (M13mp18 DNA), and about 3 kb of cyclic double-stranded DNA (pBlueScript) were used, and the cleavage activity of each by the CRISPR-Cas3 system was studied.

(3) Measurement of Nuclease Activity Using DNaseAlert™

In order to detect non-specific cleavages by using fluorescent probes, measurement of non-specific cleavages was conducted by using DNaseAlert™ Kit (IDT, Inc.). Into a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 μ M CoCl₂, and 2.5 mM ATP), Cas3 protein (20 ng/ μ L), Cascade complex (32 ng/ μ L), double-stranded DNA fragment containing target sequence (2 ng/ μ L) were mixed. Into a substrate tube of DNaseAlert, 5 μ L of Nuclease free water and 10 $\times$ Alert Buffer were added to prepare a substrate solution. Then, 20 μ L of

the sample and 5 μ L of the substrate solution (fluorescent probe solution) were mixed, and the intensity of HEX signals was observed with time using a real-time PCR apparatus at 37°C every 30 seconds.

5 (4) Measurement of Nuclease Activity Using Taqman™ Probe

In order to detect non-specific cleavages by using a fluorescent probe of single-stranded DNA, non-specific cleavage were measured by using Taqman™ probe
10 (sequence information: HEX-AAGGTCGGA-ZEN-GTCAACGGATTTGGTC-IBFQ/SEQ ID NO: 17; Thermo Fisher Science). Into a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 μ M CoCl₂, and 2.5 mM ATP), Cas3 protein (20 ng/ μ L), Cascade complex (32 ng/ μ L),
15 double-stranded DNA fragment or single-stranded DNA fragment containing the target sequence (2 ng/ μ L) were mixed. After 0.5 μ L of Taqman™ probe was prepared using 5 μ L of Nuclease free water and 10×Alert Buffer, 20 μ L of the sample and 5 μ L of the substrate solution
20 were mixed, and the intensity of HEX signals was observed with time using the real-time PCR apparatus (CFX Connect, Bio-Rad Laboratories, Inc.) at 37°C every 30 seconds.

(5) Evaluation of Target Specificity of PAM Sequence

25 Samples were prepared by changing the PAM sequences (3 bases) of the double-stranded DNA

fragments to all sequence patterns (64 types), and the non-specific cleavage activity using Taqman™ probe was measured in the same manner as in (4). The speeds of increase in fluorescence signal 1 minute after the start of the reaction were calculated, and the calculated values were compared to evaluate the intensity of the activity from differences in target recognition of the PAM sequences.

(6) Measurement of detection limit concentration

(i) Measurement in Double-stranded DNA Fragment Solution

The concentration of the double-stranded DNA fragment was adjusted through serial dilution, and the intensity of the activity in each concentration was measured by using Taqman™ probe in the same manner as in (4). In addition, the RPA method using TwistAmp Basic kit (TwistDx Inc.) was conducted on the diluted DNA sample under incubation conditions of 37°C and 20 minutes to amplify DNA in the sample, and then the intensity of the activity was measured again.

(ii) Measurement in Genome-mixed Solution

To a DNA solution obtained by diluting the concentration of the double-stranded DNA fragment through serial dilution, 20 ng of a mouse genome was mixed as a non-specific DNA to prepare a genome-mixed solution. Then, the RPA method using TwistAmp Basic

kit (TwistDx Inc.) was conducted under incubation conditions of 37°C and 20 minutes to amplify DNA in the sample. Then, the intensity of the activity in each concentration was measured using Taqman™ probe in the same manner as in (4).

(7) Detection of Novel Coronavirus RNA

A SARS-specific Cascade complex for use in detection was prepared, targeting the N gene of the novel coronavirus (SARS-CoV2). As the novel coronavirus RNA, a novel coronavirus RNA obtained by purifying the viruses, which was proliferated using the VeroE6/TMPRSS2 cell line, by using the QIAamp Viral RNA Mini Kit (QIAGEN) was used. After the virus RNA solution was serially diluted, the RT-LAMP method using the WarmStart LAMP Kit (NEB Inc.) was conducted under incubation conditions of 62°C and 30 minutes to amplify the virus genome. Into 18 µl of a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 µM CoCl₂, 2.5 mM ATP, 400 nM Cas3 protein, 100 nM Cascade complex, and 250 nM Taqman™ probe), 2 µl of the virus genome-amplified solution was mixed, and the intensity of the activity in each concentration was measured by using the real-time PCR. The primer set used in the target amplification was designed using the PrimerExplorerV5 (<https://primerexplorer.jp/lampv5/index.html>).

Note that the virus-specific target sequences are

as follows:

SARS-N1: aaggccaaactgtcactaagaaatctgctgctgag/SEQ ID

NO: 18

SARS-N2: aaggaactgattacaaacattggccgcaaattgca/SEQ ID

5 NO: 19.

In addition, the primers used in the RT-RAMP method are as follows:

<SARS-N1 Set>

SARS-N1-FIP:

10 GTTGGCCTTTACCAGACATTTTGGTGATGCTGCTCTTGCTT/SEQ ID NO:

20

SARS-N1-BIP:

TGCTGAGGCTTCTAAGAAGCCAGCTTGTGTTACATTGTATGC/SEQ ID NO:

21

15 SARS-N1-F3: ACTTCTCCTGCTAGAAATGG/SEQ ID NO: 22

SARS-N1-B3: GTTTGTTCTGGACCACGT/SEQ ID NO: 23

SARS-N1-LF: TTCAATCTGTCAAGCAGCAGCA/SEQ ID NO: 24

SARS-N1-LB: GGCAAAAACGTACTGCCACTA/SEQ ID NO: 25.

<SARS-N2 Set>

20 SARS-N2-FIP:

TCTGATTAGTTCCTGGTCCCCAAAGCATACAATGTAACACAAGC/SEQ ID

NO: 26

SARS-N2-BIP:

CGCATTGGCATGGAAGTCACTTTGATGGCACCTGTGTAG/SEQ ID NO: 27

25 SARS-N2-F3: GCAAAAACGTACTGCCAC/SEQ ID NO: 28

SARS-N2-B3: GAAATTTGGATCTTTGTCATCC/SEQ ID NO: 29

SARS-N2-LF: TGGACCACGTCTGCCGA/SEQ ID NO: 30

SARS-N2-LB: ACCTTCGGGAACGTGGTT/SEQ ID NO: 31.

(8) Detection Using Lateral Flow Test Strip

In order to visualize signal in lateral flow test strips, the synthesis of a signal probe labeled with FITC and biotin (FITC-GTCAACGGATTGGTC-BIO/SEQ ID NO: 32) was synthesized by outsourcing the synthesis to FASMAC. Into 18 μ l of a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 μ M CoCl₂, 2.5 mM ATP, 400 nM Cas3 protein, 200 nM Cascade complex, and 100 nM probe), 2 μ l of the above virus genome-amplified solution was mixed and incubated at 37°C for 10 minutes. After 50 μ l of water was added, one end (on the control line side) of a test strip of HybriDetect (Milenia Biotec GmbH) was immersed into the reaction buffer, and left to stand at room temperature for 2 minutes, followed by detection.

(9) Detection of Sequences of Drug-resistant Mutant Influenza Viruses

In order to detect single nucleotide mutations of drug-resistant mutant influenza viruses, target sequences of crRNAs was designed for mutated portions resistant to a baloxavir marboxil drug and a neuraminidase inhibitor, and a Cascade complex containing the crRNA was prepared. Then, 20 μ l of a mixture liquid of a double-stranded DNA fragment

solution (40 ng/μl) containing a wild-type sequence and a mutated sequence and a reaction buffer (5 mM HEPES-K pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10 μM CoCl₂, 2.5 mM ATP, 400 nM Cas3 protein, 100 nM Cascade complex, and 250 nM Taqman[™] probe) was prepared, and the intensity of the activity was measured using a real-time PCR apparatus (CFX Connect, Bio-Rad Laboratories, Inc.).

Note that the crRNA sequences are as follows.

<Xofluza-resistant Mutation>

For detecting the wild type; H1N1-Cascade-I38:
TGCAGCAAACCTTATTAGTTTCAATTTTGGGGT/SEQ ID NO: 33

For detecting the wild type; H3N2-Cascade-I38:
TGCACTCACTTGGAGGTGTGTTTCATGTATTC/SEQ ID NO: 34.

<Tamiflu-resistant Mutation>

For detecting the mutated type; H1N1-Cascade-I222:
ATTGAGAACACAAGAGTCTGAATGTGCATGTG/SEQ ID NO: 35

For detecting the mutated type; H1N1-Cascade-H274:
TAATTAGGGGCTTTCATTTGCGACTGATTTGAT/SEQ ID NO: 36

For detecting the mutated type; H3N2-Cascade-N294:
CTGTCTCTGCAGACACATCTGACACCAGGATA/SEQ ID NO: 37.

In addition, DNA fragment sequences for detecting drug-resistant mutations of influenza viruses are as follows (the underlines indicate bases corresponding to the mutation sites).

H1N1-PA-I38:

GGGAAGACCCCAAAATTGAACTAATAAGTTTGCTGCAATTTGCACACATTTGG

AAGTTTGTTCATGTATTCGGATTTC/SEQ ID NO: 38

H1N1-PA-I38T:

GGGAAGACCCCAAAATTGAACTAATAAGTTTGCTGCAACTTGCACACATTTGG

AAGTTTGTTCATGTATTCGGATTTC/SEQ ID NO: 39

5 H3N2-PA-I38:

GGGGAGGATCTGAAAATTGAAACCAACAAATTTGCAGCAATATGCACTCACTTG

GAGGTGTGTTTCATGTATTCAGATTT/SEQ ID NO: 40

H3N2-PA-I38T:

GGGGAGGATCTGAAAATTGAAACCAACAAATTTGCAGCAACATGCACTCACTTG

10 GAGGTGTGTTTCATGTATTCAGATTT/SEQ ID NO: 41

H1N1-NA-I222:

GGCATAATAACAGACACTATCAAGAGTTGGAGGAACAATATATTGAGAACACAA

GAGTCTGAATGTGCATGTGTAAATGG/SEQ ID NO: 42

H1N1-NA-I222R:

15 GGCATAATAACAGACACTATCAAGAGTTGGAGGAACAATAGATTGAGAACACAA

GAGTCTGAATGTGCATGTGTAAATGG/SEQ ID NO: 43

H1N1-NA-H274:

AAAGATAATCAAATCAGTCGAAATGAAAGCCCCTAATTATCACTATGAGGAATG

CTCCTGTTACCCTGATTCTAGTGAAA/SEQ ID NO: 44

20 H1N1-NA-H274Y:

AAAGATAATCAAATCAGTCGAAATGAAAGCCCCTAATTATTACTATGAGGAATG

CTCCTGTTACCCTGATTCTAGTGAAA/SEQ ID NO: 45

H3N2-PA-N294:

TATCCTCGATATCCTGGTGTCTGAGATGTGTCTGCAGAGACAACTGGAAAGGATCC

25 AACCGGCCCATCATAGATATAAACAT/SEQ ID NO: 46

H3N2-PA-N294S:

TATCCTCGATATCCTGGTGTCTGATGTGTCTGCAGAGACAGCTGGAAAGGATCC
AACCGGCCCATCATAGATATAAACAT/SEQ ID NO: 47.

B. Results

As a result of measuring non-specific cleavages
5 by electrophoresis, the cleavage activity of the
single-stranded DNA was not detected without the DNA
fragments containing the target sequences in any of the
cases of the Cas3 protein only, the Cascade complex
only, or the combination of the Cas3 protein and the
10 Cascade complex. On the other hand, in the case where
the DNA fragment containing the PAM sequence (AAG or
ATG) which was recognizable by the CRISPR-Cas3 system
and the target sequence was mixed, the decomposition of
the donor single-stranded DNA was detected (arrows in
15 Fig. 1). On the other hand, in the case where the DNA
fragment in which the PAM sequence was changed to the
PAM sequence (CCA) which was not recognizable by the
CRISPR-Cas3 system was mixed, the decomposition of the
donor DNA was not detected. In the case where the
20 cyclic single-stranded DNA was targeted as well,
although the cleavage activity by the CRISPR-Cas3
system was observed, the cleavage activity was slightly
lower than the case where the linear single-stranded
DNA was targeted. On the other hand, as a result of
25 using the double-stranded DNA as the donor DNA, the
decomposition of the DNA was not detected even in the

presence of the DNA fragment containing the target sequence (arrows in Fig. 2). From the above-described result, it was revealed that the CRISPR-Cas3 system exhibited the cleavage activity of a non-specific single-stranded DNA only in the presence of of the target sequence.

Next, by utilizing the characteristics of this CRISPR-Cas3 system, it was studied whether or not it was possible to detect the presence or absence of the target sequence by using a fluorescent probe (DNaseAlert™; a probe in which a fluorescent dye "HEX" and a quencher "dark quencher" were bound to the respective ends). Fluorescence signals were measured with time by mixing a substrate (fluorescent probe) that emits fluorescence upon DNA decomposition (DNase). As a result, it was revealed that the fluorescence signals were strongly detected only in the case where the DNA fragment containing the PAM sequence (AAG or ATG) which was recognizable by the CRISPR-Cas3 system and the target sequence was mixed (Fig. 3). The same result was obtained also in the case where mTyr was used as the target sequence (Fig. 4). In order to more precisely detect the decomposition of the single-stranded DNA, studies were conducted using a fluorescent probe (Taqman™ probe; a double-quencher probe in which a fluorescent dye "HEX" and a quencher

"IBFQ" were bound to the respective ends and into which "ZEN" was introduced as the internal quencher). As a result, it was revealed that the fluorescence signals were detected only in the case where the CRISPR-Cas3 system was able to recognize the target sequence (Fig. 5).

Moreover, in order to check whether or not the CRISPR-Cas3 system exhibits non-specific single-stranded DNA cleavage activity even when target DNA is single-stranded, an experiment using a single-stranded DNA fragment as target DNA was conducted. As a result, it was revealed that the CRISPR-Cas3 system significantly exhibited non-specific single-stranded DNA cleavage activity even when the target DNA was single-stranded (target strand) (Fig. 6).

From the above-described results, it was found that the non-specific cleavage activity of the single-stranded DNA present therearound was induced by the CRISPR-Cas3 system recognizing the sequence of the target DNA (which may be a single-stranded DNA or may be a double-stranded DNA). In addition, it was found that it was possible to simply and easily as well as efficiently determine the activity of the CRISPR-Cas3 system by using a fluorescent probe.

Next, in order to study differences in single-stranded DNA cleavage activity due to differences in

target specificity of the CRISPR-Cas3 system, the PAM sequences in the target double-stranded DNA fragments were modified, and the activity was measured for all the PAM sequence patterns. As a result, it was revealed that while high intensity of the activity was exhibited with the conventionally reported PAM sequences (AAG, ATG, and the like), in the case where the first base or the second base was C and the case where the third base was T, almost no activity was exhibited (Fig. 7). Since the change of a single base on the PAM sequence affected the non-specific single-stranded DNA cleavage activity by the CRISPR-Cas3 system, it was found that based on the activity as indication, it was possible to detect target DNAs in samples with specificity on a single base level.

Next, in order to study the limit concentration of target DNA that was able to be detected, the target DNA fragment was diluted by limiting dilution and adjusted from 1 copy to 10^{12} copies, and the intensity of the activity in the target DNA fragment of each concentration was measured. As a result, significant signals were detected in the case where target DNAs of 10^{10} copies or more were used (Fig. 8). It has been reported that existing systems using Cas12a and Cas13 also exhibit detection sensitivities of around 10^{10} copies, and it was revealed that the CRISPR-Cas3 system

exhibited comparable detection sensitivity. Furthermore, DNAs in these diluted samples were amplified using the RPA method, and the intensity of the activity was measured again. As a result, even for a sample that initially contained only 1 copy of the target DNA, the non-specific single-stranded DNA cleavage activity was detected (Fig. 9). From the above, it was revealed that it was possible to detect an infinitesimal amount (for example, 1 copy) of a target DNA in a sample using the CRISPR-Cas3 system only by combining the CRISPR-Cas3 system with a simple and easy gene amplification method such as the RPA method.

In addition, in order to study the limit concentration of the target DNA in the presence of a large number of non-specific DNA sequences, mouse genomic DNA was mixed into a dilution of the target DNA fragment obtained through limiting dilution, and the intensity of the activity was measured. As a result, it was possible to detect the target DNA even with the sample containing only 1 copy of the target DNA like in Fig. 9 (Fig. 11). That is, it was revealed that it was possible to specifically detect 1 copy or more of the target DNA even when other DNAs were contained, by amplifying the target DNA by a method such as the RPA method.

In order to prove that the present invention can

be utilized for detecting viruses of emerging infectious diseases and the like in practice, two Cascade complexes (N1, N2) specific to the N genes of the novel coronaviruses (SARS-CoV2) were synthesized.

5 The RT-LAMP method was conducted for each region, and amplified virus genomes were mixed with Cas3 and Cascade complex protein. As a result, it was possible to detect the signal in the virus RNA sample containing 1 copy of N1 and 10 copies of N2. It was possible to detect
10 signals in all of the samples containing 1000 copies or more of both viruses (Fig. 12).

Next, in order to detect novel coronaviruses in lateral flow assay (Fig. 13A), virus genomes amplified by the RT-LAMP method were mixed with a signal probe
15 labeled with FITC and biotin, Cas3, and the Cascade complex protein, and one end (on the control line side) of a lateral flow test strip was immersed in the reaction system. As a result, the signal of the positive band in the lateral flow test strip was
20 intensified depending on the reaction time, and it was revealed that it was possible to sufficiently determine the presence or absence of the band when reaction was made at 37°C for 10 minutes (Fig. 13B).

Furthermore, as a result of conducting lateral
25 flow assay on samples of serially diluted virus RNA, it was revealed that sufficient determination was possible

in the case of using 100 copies of the sample (Fig. 14).

The overview of the assay in the present experiment is summarized in Fig. 15. From the above-described results, it was proved that it was possible to simply and easily detect the novel coronavirus (SARS-CoV2) by using the Cascade complex and the Cas3 protein.

Next, it was studied whether it was possible, by using the CRISPR-Cas3 system, to detect single nucleotide mutations which the influenza viruses resistant to baloxavir marboxil drugs (Xofluza and the like) and neuraminidase inhibitors (Tamiflu, Relenza, and the like) had. As a result of detecting those for I38T of the RNA polymerase (PA) gene of type A influenza viruses (H1N1 and H3N2), which was known as Xofluza-resistant mutations, significantly strong signals were successfully obtained in wild-type sequences for both H1N1 and H3N2 (Fig. 16).

Subsequently, as a result of detecting those for mutations (H1N1: I222R or H274Y and H3N2: N294S) of the neuraminidase (NA) gene which are known as mutations resistant to Tamiflu and Relenza of the type A influenza viruses, significantly strong signals were successfully obtained in mutated sequences for both H1N1 and H3N2 (Fig. 17).

From the above, drug-resistant mutations on a single base level were successfully detected by

designing the target sequence of crRNA in the CRISPR-Cas3 system for the mutation sites.

[Industrial Applicability]

5 The present invention makes it possible to simply
and easily as well as efficiently detect target DNAs in
samples by utilizing the CRISPR-Cas3 system and the
single-stranded probe DNA. The present invention
greatly contributes, for example, to diagnoses of
infections with viruses, pathogenic bacteria, protozoa,
10 diagnoses of mutation and polymorphism of genes related
to risks of diseases, efficacies of drugs, and the like,
and the like.

[Sequence Listing]

[CLAIMS]

[Claim 1]

A method for detecting a specific DNA in a sample comprising the steps of:

5 (a) contacting the sample with a CRISPR-Cas3 system targeting the specific DNA and a single-stranded probe DNA; and

 (b) detecting cleavage of the single-stranded probe DNA by the CRISPR-Cas3 system that occurs in a
10 case where the specific DNA is present in the sample.

[Claim 2]

A kit for detecting a specific DNA in a sample by the method according to claim 1, comprising:

 (a) a CRISPR-Cas3 system targeting the specific
15 DNA; and

 (b) a single-stranded probe DNA.