CLAIMS

[Claim 1]

A method for designing a fusion protein containing a functional region, and a binding region fused thereto that can bind ~~in~~to a DNA base~~-selective manner~~ or ~~a~~specific DNA base sequence~~-specific manner, which contains~~, the method comprising

designing the binding region to contain one or more PPR motifs having a structure of the following formula 1:

~~[Formula 1]~~

(Helix A)-X-(Helix B)-L (Formula 1)

(wherein, in the formula 1:

Helix A is a part that can form an α-helix structure;

X does not exist, or is a part consisting of 1 to 9 amino acids;

Helix B is a part that can form an α-helix structure; and

L is a part consisting of 2 to 7 amino acids),

wherein,

under the following definitions:

the first amino acid of Helix A is referred to as No. 1 amino acid (No. 1 A.A.), the fourth amino acid as No. 4 amino acid (No. 4 A.A.), and

- when a next PPR motif (M_{n+1}) contiguously exists on the C-terminus side of the PPR motif (M_n) (when there is no amino acid insertion between the PPR motifs), the -2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n);

~~motif (M_n) (when there is no amino acid insertion between the PPR motifs), the -2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n);~~

- when a non-PPR motif consisting of 1 to 20 amino acids exists between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the amino acid locating upstream of the first amino acid of the next PPR motif (M_{n+1}) by 2 positions, i.e., the -2nd amino acid; or

- when any next PPR motif (M_{n+1}) does not exist on the C-terminus side of the PPR motif (M_n), or 21 or more amino acids constituting a non-PPR motif exist between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the 2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n)

is referred to as No. "ii" (-2) amino acid (No. "ii" (-2) A.A.), ~~one PPR motif (M_n) contained in the protein is~~

wherein a ~~PPR motif having a specific~~

combination of ~~amino acids corresponding to a target DNA base or target DNA base sequence as~~ the three amino acids of No. 1 A.A., No. 4 A.A., and

No. "ii" (-2) A.A~~.~~ of one PPR motif (M_n) contained in the binding region is

corresponding to the DNA base or DNA base sequence.

[Claim 2]

The ~~protein~~method according to claim 1~~,~~ wherein the combination of the three amino acids ~~of No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.~~ is ~~a combination corresponding to a target DNA base or target DNA base sequence, and the combination of amino acids is determined~~designated according to any ~~one~~ of the following definitions:

(1-1) when No. 4 A.A. is glycine (G), No. 1 A.A. may be an arbitrary amino acid, and No. "ii" (-2) A.A. is aspartic acid (D), asparagine (N), or serine (S);

(1-2) when No. 4 A.A. is isoleucine (I), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-3) when No. 4 A.A. is leucine (L), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-4) when No. 4 A.A. is methionine (M), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-5) when No. 4 A.A. is asparagine (N), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-6) when No. 4 A.A. is proline (P), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-7) when No. 4 A.A. is serine (S), each of No. 1 A.A. and No. "ii" (-2) A.A. may be ~~an arbitrary amino acid;~~ an arbitrary amino acid;

(1-8) when No. 4 A.A. is threonine (T), each of No. 1 A~~,~~.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid; and

(1-9) when No. 4 A.A. is valine (V), each of No. 1 A.A. and No. "ii" (-2) A.A. may ~~be an arbitrary amino acid.~~ be an arbitrary amino acid.

[Claim 3]

The ~~protein~~method according to claim 1 or 2, wherein the combination of the three amino acids ~~of No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.~~ is ~~a combination corresponding to a target DNA base or target DNA base sequence, and the combination of amino acids is determined~~designated according to any ~~one~~ of the following definitions:

(2-1) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are an arbitrary amino acid, glycine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-2) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and aspartic acid, respectively, the PPR motif

selectively binds to G;

(2-3) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-4) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-5) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, glycine, and serine, respectively, the PPR motif selectively binds to A, and next binds to C;

(2-6) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-7) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and asparagine, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-8) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-9) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and aspartic acid, respectively, the PPR motif selectively binds to C;

(2-10) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and lysine, respectively, the PPR motif selectively binds to T;

(2-11) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T;

(2-12) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-13) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., ~~are~~ are isoleucine, methionine, and aspartic acid, respectively, the PPR motif

selectively binds to T, and next binds to C;

(2-14) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to C and T;

(2-15) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-16) when the three amino acids, No. 1 A~~,~~.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-17) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glycine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-18) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-19) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are threonine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-20) when the three amino acids, No. 1 A.A., No. 4 A.A., and No~~,~~. "ii" (-2) A.A. are valine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-21) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are tyrosine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-22) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-23) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-24) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are serine, asparagine, and asparagine, respectively, the PPR motif selectively ~~binds to C;~~

binds to C;

(2-25) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-26) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-27) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-28) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-29) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-30) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii~~.~~"" (-2) A.A., are an arbitrary amino acid, asparagine, and tryptophan, respectively, the PPR motif selectively binds to C, and next binds to T;

(2-31) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and tryptophan, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-32) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T;

(2-33) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-34) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-35) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are tyrosine, proline, and aspartic acid, respectively, the PPR motif selectively ~~binds to T;~~

~~(2-36) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;~~

binds to T;

(2-36) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-37) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., ~~are~~ are an arbitrary amino acid, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-38) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-39) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-40) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-41) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-42) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-43) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-44) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-45) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-46) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-47) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and an arbitrary amino acid, respectively, the PPR motif binds with A, C, and T, but does not bind to G;

(2-48) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, valine, and aspartic acid, respectively, the PPR motif selectively binds to C, and next binds to A;

(2-49) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and glycine, respectively, the PPR motif selectively binds to C; and

~~binds to C; and~~

(2-50) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and threonine, respectively, the PPR motif selectively binds to T.

~~an arbitrary amino acid, valine, and threonine, respectively, the PPR motif selectively binds to T.~~

[Claim 4]

The ~~protein~~method according to any one of claims 1 to 3, ~~which contains 2 to 30 of~~wherein the ~~PPR motifs (M_n) defined in claim 1~~ functional region is a DNA-cleaving enzyme, or a nuclease domain thereof, or a transcription control domain, and the fusion protein functions as a target sequence-specific DNA-cleaving enzyme or transcription control factor.

[Claim 5]

The ~~protein~~method according to ~~any one of claims 1 to 3, which contains 5 to 25 of the PPR motifs (M_n) defined in~~ claim ~~1.~~4, wherein the DNA-cleaving enzyme is the nuclease domain of FokI (SEQ ID NO: 6).

[Claim 6]

~~The~~A method for editing a genome, characterized by utilizing the fusion protein designed by the method according ~~to~~ any one of claims 1 to ~~3, which contains 9 to 15 of the PPR motifs (M_n) defined~~5 (excluding a method for performing in ~~claim 1.~~a human individual).

[Claim 7]

~~The PPR protein according to claim 6, which consists of a sequence selected from the amino acid sequence of SEQ ID NO: 1 containing 9 PPR motifs, the amino acid sequence of SEQ ID NO: 2 containing 11 PPR motifs, the amino acid sequence of SEQ ID NO: 3 containing 15 PPR motifs, the amino acid sequence of SEQ ID NO: 4 containing 10 PPR motifs, and the amino acid sequence of SEQ ID NO: 5 containing 11 PPR motifs.~~

A method for producing a genome, the method comprising editing a genome by the method according to claim 6, and producing the edited genome (excluding a method for performing in a human individual).

[Claim 8]

A method for ~~identifying~~producing a ~~DNA base or DNA base sequence that serves as a target of a DNA-binding protein~~cell containing ~~one or more (preferably 2 to 30) PPR motifs (M_n) defined in claim 1, wherein:~~

an edited genome, the ~~DNA base or DNA base sequence is identified~~method comprising editing a genome by ~~determining presence or absence of a DNA base corresponding to a combination of~~ the ~~three amino acids of No. I A.A., No. 4 A.A.,~~method according to claim 6, and ~~No. "ii" (-2) A.A. of~~producing a cell containing the ~~PPR motif on the basis of any one of the definitions (1-1) to (1-9) mentioned in claim 2, and (2-1) to (2-50) mentioned in claim 3.~~edited genome (excluding a method for performing in a human individual).

[Claim 9]

A method ~~for identifying~~of using a ~~PPR~~ protein ~~containing one or more (preferably 2 to 30) PPR motifs (M_n) defined in claim 1 that can bind~~for binding to a ~~target~~ DNA base or ~~target DNA having a~~ specific DNA base sequence, wherein~~:~~ the protein contains one or 2 to 30 PPR motifs having a structure of the ~~PPR protein~~following formula 1:

(Helix A)-X-(Helix B)-L (Formula 1)

(wherein, in the formula 1:

Helix A is ~~identified by determining presence~~a part that can form an α-helix structure;

X does not exist, or ~~absence~~is a part consisting of ~~a~~1 to 9 amino acids;

Helix B is a part that can form an α-helix structure; and

L is a part consisting of 2 to 7 amino acids),

wherein,

under the following definitions:

the first amino acid of Helix A is referred to as No. 1 amino acid (No. 1 A.A.), the fourth amino acid as No. 4 amino acid (No. 4 A.A.), and

- when a next PPR motif (M_{n+1}) contiguously exists on the C-terminus side of the PPR motif (M_n) (when there is no amino acid insertion between the PPR motifs), the -2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n);

- when a non-PPR motif consisting of 1 to 20 amino acids exists between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the amino acid locating upstream of the first amino acid of the next PPR motif (M_{n+1}) by 2 positions, i.e., the -2nd amino acid; or

- when any next PPR motif (M_{n+1}) does not exist on the C-terminus side of the PPR motif (M_n), or 21 or more amino acids constituting a non-PPR motif exist between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the 2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n)

is referred to as No. "ii" (-2) amino acid (No. "ii" (-2) A.A.),

wherein a combination of the three amino acids of No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.~~corresponding to the target DNA base or a specific base constituting the target DNA on the basis~~ of ~~any~~ one ~~of the definitions (1-1) to (1-9) mentioned~~PPR motif (M_n) contained in ~~claim 2, and (2-1) to~~the protein is designated according to any of the following definitions:

~~(2-50) mentioned in claim 3.~~

(1-1) when No. 4 A.A. is glycine (G), No. 1 A.A. may be an arbitrary amino acid, and No. "ii" (-2) A.A. is aspartic acid (D), asparagine (N), or serine (S); (1-2) when No. 4 A.A. is isoleucine (I), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-3) when No. 4 A.A. is leucine (L), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-4) when No. 4 A.A. is methionine (M), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-5) when No. 4 A.A. is asparagine (N), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-6) when No. 4 A.A. is proline (P), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-7) when No. 4 A.A. is serine (S), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid;

(1-8) when No. 4 A.A. is threonine (T), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid; and

(1-9) when No. 4 A.A. is valine (V), each of No. 1 A.A. and No. "ii" (-2) A.A. may be an arbitrary amino acid.

[Claim 10]

~~A~~The method ~~for controlling a function of DNA, which uses the protein~~ according to claim ~~1.~~9, wherein the combination of the three amino acids is designated according to any of the following definitions:

(2-1) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are an arbitrary amino acid, glycine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-2) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-3) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-4) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-5) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are

an arbitrary amino acid, glycine, and serine, respectively, the PPR motif selectively binds to A, and next binds to C;

(2-6) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-7) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and asparagine, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-8) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-9) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and aspartic acid, respectively, the PPR motif selectively binds to C;

(2-10) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and lysine, respectively, the PPR motif selectively binds to T;

(2-11) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T;

(2-12) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-13) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, methionine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-14) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to C and T;

(2-15) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-16) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.,

are phenylalanine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-17) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glycine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-18) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-19) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are threonine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-20) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are valine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-21) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are tyrosine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-22) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-23) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-24) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are serine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-25) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-26) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-27) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.,

are valine, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-28) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-29) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-30) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and tryptophan, respectively, the PPR motif selectively binds to C, and next binds to T;

(2-31) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and tryptophan, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-32) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T;

(2-33) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-34) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-35) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are tyrosine, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-36) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-37) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-38) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.,

are phenylalanine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-39) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-40) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-41) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-42) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-43) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-44) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-45) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-46) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-47) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and an arbitrary amino acid, respectively, the PPR motif binds with A, C, and T, but does not bind to G;

(2-48) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, valine, and aspartic acid, respectively, the PPR motif selectively binds to C, and next binds to A;

(2-49) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A.,

are an arbitrary amino acid, valine, and glycine, respectively, the PPR motif selectively binds to C; and

(2-50) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and threonine, respectively, the PPR motif selectively binds to T.

[Claim 11]

A ~~complex consisting of~~ nucleic acid encoding a ~~region comprising~~protein, wherein the protein ~~according~~contains one or 2 to ~~claim 1,~~30 of PPR motifs having a structure of the following formula 1:

~~and a functional region bound together.~~

(Helix A)-X-(Helix B)-L (Formula 1)

(wherein, in the formula 1:

Helix A is a part that can form an α-helix structure;

X does not exist, or is a part consisting of 1 to 9 amino acids;

Helix B is a part that can form an α-helix structure; and

L is a part consisting of 2 to 7 amino acids),

wherein,

under the following definitions:

the first amino acid of Helix A is referred to as No. 1 amino acid (No. 1 A.A.), the fourth amino acid as No. 4 amino acid (No. 4 A.A.), and

- when a next PPR motif (M_{n+1}) contiguously exists on the C-terminus side of the PPR motif (M_n) (when there is no amino acid insertion between the PPR motifs), the -2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n);

- when a non-PPR motif consisting of 1 to 20 amino acids exists between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the amino acid locating upstream of the first amino acid of the next PPR motif (M_{n+1}) by 2 positions, i.e., the -2nd amino acid; or

- when any next PPR motif (M_{n+1}) does not exist on the C-terminus side of the PPR motif (M_n), or 21 or more amino acids constituting a non-PPR motif exist between the PPR motif (M_n) and the next PPR motif (M_{n+1}) on the C-terminus side, the 2nd amino acid counted from the end (C-terminus side) of the amino acids constituting the PPR motif (M_n)

is referred to as No. "ii" (-2) amino acid (No. "ii" (-2) A.A.),

one PPR motif (M_n) contained in the protein is a PPR motif having a specific

combination of the three amino acids of No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A corresponding to a target DNA base or target DNA base sequence, and

wherein the combination of three amino acids of No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A is any of the following definitions:

(2-1) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are an arbitrary amino acid, glycine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-2) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-3) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-4) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glutamic acid, glycine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-5) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, glycine, and serine, respectively, the PPR motif selectively binds to A, and next binds to C;

(2-6) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-7) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, isoleucine, and asparagine, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-8) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T and C;

(2-9) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and aspartic acid, respectively, the PPR motif selectively binds to C;

(2-10) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, leucine, and lysine, respectively, the PPR motif

selectively binds to T;

(2-11) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to T;

(2-12) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, methionine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-13) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, methionine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-14) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to C and T;

(2-15) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-16) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-17) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are glycine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-18) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-19) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are threonine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-20) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are valine, asparagine, and aspartic acid, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-21) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A. are tyrosine, asparagine, and aspartic acid, respectively, the PPR motif selectively

binds to T, and next binds to C;

(2-22) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-23) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-24) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are serine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-25) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and asparagine, respectively, the PPR motif selectively binds to C;

(2-26) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-27) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and serine, respectively, the PPR motif selectively binds to C;

(2-28) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-29) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, asparagine, and threonine, respectively, the PPR motif selectively binds to C;

(2-30) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, asparagine, and tryptophan, respectively, the PPR motif selectively binds to C, and next binds to T;

(2-31) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, asparagine, and tryptophan, respectively, the PPR motif selectively binds to T, and next binds to C;

(2-32) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and an arbitrary amino acid, respectively,

the PPR motif selectively binds to T;

(2-33) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-34) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-35) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are tyrosine, proline, and aspartic acid, respectively, the PPR motif selectively binds to T;

(2-36) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-37) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-38) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-39) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, serine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-40) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and an arbitrary amino acid, respectively, the PPR motif selectively binds to A and G;

(2-41) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-42) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and aspartic acid, respectively, the PPR motif selectively binds to G;

(2-43) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, threonine, and asparagine, respectively, the PPR

motif selectively binds to A;

(2-44) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are phenylalanine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-45) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-46) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are valine, threonine, and asparagine, respectively, the PPR motif selectively binds to A;

(2-47) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and an arbitrary amino acid, respectively, the PPR motif binds with A, C, and T, but does not bind to G;

(2-48) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are isoleucine, valine, and aspartic acid, respectively, the PPR motif selectively binds to C, and next binds to A;

(2-49) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and glycine, respectively, the PPR motif selectively binds to C; and

(2-50) when the three amino acids, No. 1 A.A., No. 4 A.A., and No. "ii" (-2) A.A., are an arbitrary amino acid, valine, and threonine, respectively, the PPR motif selectively binds to T

(excluding a protein consisting of any sequence selected from the amino acid sequence of SEQ ID NO:1, the amino acid sequence of SEQ ID NO:2, the amino acid sequence of SEQ ID NO: 3, the amino acid sequence of SEQ ID NO:4, and the amino acid sequence of SEQ ID NO: 5).

[Claim 12]

The ~~complex~~nucleic acid according to claim 11, wherein the ~~functional region~~PPR motif is ~~fused to~~

any selected from the group consisting of

the 9 PPR motifs contained in protein ~~according to claim 1 on the C-terminus side~~consisting of the amino acid sequence of SEQ ID NO: 1,

the 11 PPR motifs contained in protein consisting of the amino acid sequence of SEQ ID NO: 2,

the 15 PPR motifs contained in protein consisting of the amino acid sequence of SEQ ID NO: 3,

the 10 PPR motifs contained in protein consisting of the amino acid sequence of SEQ ID NO: 4, and

the 11 PPR motifs contained in protein consisting of the amino acid sequence of SEQ ID NO: 5.

[Claim 13]

~~The complex according to claim 11 or 12, wherein the functional region is a DNA-cleaving enzyme, or a nuclease domain thereof,~~A nucleic acid encoding a protein that can bind in a DNA base-selective manner or a ~~transcription control domain,~~ ~~and the complex functions as a target~~DNA base sequence-specific ~~DNA-cleaving enzyme or~~manner, the protein consisting of the amino acid sequence of NOs. 230-541 of SEQ ID NO.:1, NOs. 234-621 of SEQ ID NO.:2, NOs. 106-632 of SEQ ID NO.:3, NOs. 106-632 of SEQ ID NO.:4, or NOs. 256-624 of SEQ ID NO.:5 including substitution of one or more amino acids according to any of definitions (2-1) - (2-50) defined in claim 11.

~~transcription control factor.~~

[Claim 14]

~~The complex~~A vector containing the nucleic acid according to ~~claim 13, wherein the DNA-cleaving enzyme is the nuclease domain~~any one of ~~FokI (SEQ ID NO: 6).~~claims 11 to 13.

[Claim 15]

~~A method for modifying a genetic substance of~~ A cell ~~comprising~~in which the ~~following steps:~~

~~preparing a cell containing a DNA having a target sequence; and~~

~~introducing the complex~~vector according to claim ~~11 into the cell so that the region of the complex consisting of the protein binds to the DNA having a target sequence, and therefore the functional region modifies the DNA having a target sequence~~14 has been introduced.

[Claim 16]

~~A method for identifying, recognizing, or targeting a DNA base or DNA having a specific base sequence by using a PPR protein containing one or more PPR motifs.~~

~~[Claim 17]~~

~~The method according to claim 16, wherein the protein contains one or more PPR motifs in which three amino acids among the amino acids constituting the motif constitute a specific combination of amino acids.~~

~~[Claim 18]~~

~~The method according to claim 16 or 17, wherein the protein contains one or more PPR motifs (M_n) defined in claim 1.~~The cell according to claim 15, the cell being E. coli.