all or a part of a sequence selected from the group consisting of the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5'-TGATATCAA-3' sequence, the 5'-TATAAATA-3' sequence, the 5'-TATAAGA-3' sequence, the 5'-AATAAA-3' sequence, the 5'-TTTATA-3' sequence, the 5'-CATAAAAA-3' sequence, the 5'-TATACA-3' sequence, the 5'-TTTAAGA-3' sequence, the 5'-GATAAAG-3' sequence, the 5'-TATAACA-3' sequence, the 5'-TCTTATCTT-3' sequence, the 5'-TTGTACTTT-3' sequence, the 5'-CATATAA-3' sequence, the 5'-TATAAAT-3' sequence, the 5'-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence.

Here, the target sequence may be a 10 to 25-nt contiguous sequence located to 5' end or 3' end of a sequence selected from the group consisting of the 5'-TATA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3'sequence, the 5'-TTTTAAAA-3' sequence, the 5'-

TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5’-

TGATATCAA-3' sequence, the 5’-TATAAATA-3' sequence, the 5’-TATAAGA-3' sequence, the 5’-AATAAA-3' sequence, the 5’-TTTATA-3' sequence, the 5’-CATAAAAA-3' sequence, the 5’-TATACA-3' sequence, the 5’-TTTAAGA-3' sequence, the 5’-GATAAAG-3' sequence, the 5’-TATAACA-3' sequence, the 5’-TCTTATCTT-3' sequence, the 5’-TTGTACTTT-3'sequence, the 5’-CATATAA-3' sequence, the 5’-TATAAAT-3' sequence, the 5’-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence.

The target sequence may be a 10 to 25-nt contiguous sequence located in an enhancer region of the duplicate gene.

The target sequence may be a 10 to 25-nt contiguous sequence located adjacent to an enhancer region of the duplicate gene.

The target sequence may be a 10 to 25-nt contiguous sequence located adjacent to 5’ end and/or 3’ end of PAM (proto-spacer-adjacent motif) sequence in a nucleic acid sequence of the transcriptional regulatory region of the duplicate gene.

Here, the PAM sequence may be determined according to the CRISPR enzyme.

The CRISPR enzyme may be a Cas9 protein or a Cpf1 protein.

Here, the Cas9 protein may be one or more Cas9 proteins selected from the group consisting of a *Streptococcus pyogenes*-derived Cas9 protein, a *Campylobacter jejuni*-derived Cas9 protein, a *Streptococcus thermophilus*-derived Cas9 protein, a *Staphylococcus aureus*-derived Cas9 protein and a *Neisseria meningitidis*-derived Cas9 protein.

The duplicate gene may be one or more genes selected from the group

consisting of a PMP22 gene, a PLP1 gene, a MECP2 gene, a SOX3 gene, a RAI1 gene,

[Description of Drawings]

FIG. 1 is a set of results illustrating an indel frequency (%) due to SpCas9-sgRNA-mediated gene manipulation, and illustrates an indel frequency of each of (a) TATA-box and (b) Enhancer, into which a target site of sgRNA is divided.

FIG. 2 is a set of results illustrating an indel frequency (%) due to CjCas9-sgRNA-mediated gene manipulation, and illustrates an indel frequency of each of (a) TATA-box and (b) Enhancer, into which a target site of sgRNA is divided.

FIG. 3 illustrates gene manipulation effects by SpCas9-sgRNA targeting regulatory elements of a human PMP22 gene in Schwann-like cells.

Fig 4 illustrates Frameshift mutation ratios induced by the SpCas9-sgRNAs targeting CDS of human PMP22.

Fig 5 illustrates Deletions of a small part of human PMP22 by the treatments of the dual sgRNAs. The WT sequence including sequences for Sox10 and Egr2 is a SEQ ID NO: 297, and the mutant sequence deleted a part thereof is a SEQ ID NO: 298. The target sequences for the Enh-Sp5 and the Enh-Sp16 are SEQ ID NOs: 299 and 300, respectively. In addition, the WT sequence including TATA-boxes is a SEQ ID NO: 301, and the mutant sequence deleted a part thereof is a SEQ ID NO: 302. The target sequences for the TATA-Sp12 and the TATA-Sp14 are SEQ ID NOs: 303 and 304, respectively.

FIG. 6 is a graph illustrating a decrease in mRNA expression of human PMP22 by SpCas9-sgRNA in human Schwann-like cells.

FIG. 7 is a graph illustrating effective and specific expression decreases of PMP22 by SpCas9-sgRNA at each target site of a human PMP22 gene in human

primary Schwann cells, and (a) illustrates indel frequency measurement results by SpCas9-sgRNA at each target site, (b) illustrates relative mRNA expression comparison results of PMP22, which are measured by qRT-PCR with or without a treatment of a myelination signal factor and an RNP complex for each target site (n = 3, One-way ANOVA and Tukey post-hoc tests: * $p < 0.05$), and (c) illustrates indel

frequency measurement results by SpCas9-sgRNA targeting distal enhancer sites (distal enhancer regions) B and C.

FIG. 8 is a graph illustrating effective and specific expression decreases of PMP22 through CRISPR-Cas9 targeting a TATA-box site of a human PMP22 gene in vitro, and (a) illustrates a target sequence targeting a promoter region of a human PMP22 position (the sequences of a part of the promoter region are SEQ ID NOs: 305 (top strand) and 306 (bottom strand)), and the leftmost graph, the middle graph, and the rightmost graph in (b) illustrate indel frequency measurement results using targeted deep sequencing in human primary Schwann cells, TATA-box 1 mutation frequency measurement results (n = 3) among the total indel frequencies, and relative mRNA expression comparison results of PMP22, which are measured by qRT-PCR with or without a treatment of a myelination signal factor and an RNP complex in human primary Schwann cells (n = 3, One-way ANOVA and Tukey post-hoc tests: * $p < 0.05$), respectively.

FIG. 9 illustrates indel frequencies by PMP22-TATA RNP in off-targets and on-targets found through an in silico off-target analysis by target deep sequencing in human primary Schwann cells, and (a) is a graph illustrating indel frequencies, (b) illustrates indel patterns with a high frequency (the local sequence of WT is a SEQ ID NO: 340, and the local sequences including indels are SEQ ID NOs: 308 to 312 (Indel -1 to -4 order)), and (c) shows off-target sites found through an in silico off-target analysis (the on-target sequence is a SEQ ID NO: 313, and the off-target sequences are SEQ ID NOs: 314 to 331 (Off1 to Off18 order)).

FIG. 10 is a set of results illustrating sites cleaved by PMP22-TATA RNP in a

human's entire genome, and (a) illustrates a Genome-wide Circos plot, (b) illustrates off-target sites appearing by the Digenome-seq among off-target sites found through an in silico off-target analysis, and (c) is a graph illustrating indel frequencies in off-target sites. In (b) and (c), the on-target sequence is a SEQ ID NO: 313, and the off-target sequences are SEQ ID NOs: 332 to 340 (Off1 to Off9 order).

FIG. 11 schematically illustrates a therapeutic approach using PMP22-TATA

RNA therapy in C22 mice.

FIG. 12 is a set of results illustrating the alleviation of a disease phenotype through expression inhibition of PMP22 by CRISPR/Cas9 in CMT1A mice, and (a) is a graph illustrating indel frequencies using targeted deep sequencing in a sciatic nerve treated with mRosa26 or a PMP22-TATA RNP complex (n = 3), (b) is a TATA-box 1 mutation frequency measurement result (n = 3) among the total indel frequencies, and (c) is a graph comparing the relative amounts of mRNA expressed of PMP22 using qRT-PCR from the sciatic nerve treated with mRosa26 or a PMP22-TATA RNP complex.

FIG. 13 is a set of results illustrating off-target sites and indel frequencies of PMP22-TATA sgRNA in a mouse genome by an in silico analysis, and (a) illustrates off-target sites and (b) is a graph illustrating an indel frequency at each off-target site. In (a) and (b), the on-target sequence is a SEQ ID NO: 313, and the off-target sequences are SEQ ID NOs: 341 to 348 (Off1 to Off8 order).

FIG. 14 is a set of results illustrating the alleviation of a disease phenotype through expression inhibition of PMP22 by CRISPR/Cas9 in CMT1A mice, and (a) is a set of images of a semithin section of the sciatic nerve tissue treated with mRosa26 or a PMP22-TATA RNP complex, and the upper graph and the low graph in (b) are a scatter plot illustrating that the g-ratio is increased in mice treated with PMP22-TATA RNP and a graph illustrating that the diameter of the myelinated axon is increased in mice treated with PMP22-TATA RNP, respectively.

FIG. 15 is a set of results illustrating electrophysiological changes through expression inhibition of PMP22 by CRISPR/Cas9 in CMT1A mice, and (a) is a graph

illustrating the change in distal latency (DL), (b) is a graph illustrating the change in motor nerve conduction velocity (NCV), and (c) is a graph illustrating the change in

5

distinguished by the numbers represented in target sequences on the graphs.

FIG. 20 is a set of graphs showing screening results of CjCas9-sgRNAs targeting the TATA box region of mPlp1. (a) shows the indel frequency (%) confirmed in NIH-3T3 cells, and (b) shows the indel frequency (%) confirmed in N20.1 cells. Here, the used sgRNAs were mPlp1-TATA-Cj-01 to mPlp1-TATA-Cj-04, and distinguished by the numbers represented in target sequences on the graphs.

FIG. 21 is a set of graphs showing screening results of SpCas9-sgRNAs targeting an enhancer (wMN1 enhancer) region of mPlp1. (a) shows the indel frequency (%) confirmed in NIH-3T3 cells, and (b) shows the indel frequency (%) confirmed in N20.1 cells. Here, the used sgRNAs were mPlp1-wMN1-Sp-01 to mPlp1-wMN1-Sp-36, and distinguished by the numbers represented in target sequences on the graphs.

FIG. 22 is a graph showing a screening result of CjCas9-sgRNAs targeting an enhancer (wMN1 enhancer) region of mPlp1. (a) shows the indel frequency (%) confirmed in NIH-3T3 cells, and (b) shows the indel frequency (%) confirmed in N20.1 cells. Here, the used sgRNAs were mPlp1-wMN1-Cj-01 to mPlp1-wMN1-Cj-28, and distinguished by the numbers represented in target sequences on the graph.

FIG. 23 is a set of graphs showing the mRNA expression levels of Plp according to SpCas9-sgRNA and CjCas9-sgRNA targeting the TATA box and enhancer (wMN1 enhancer) regions of mPlp1. (a) shows the mRNA expression level of Plp according to SpCas9-sgRNA, and here, mPlp1-TATA-Sp-01 targeting the TATA box region and mPlp1-wMN1-Sp-07 + mPlp1-wMN1-Sp-27 and mPlp1-wMN1-Sp-08 + mPlp1-wMN1-Sp-27 targeting the enhancer were used as sgRNAs. (b) shows the

TATA(A/T)A(A/T)-3’ sequence present in a promoter and/or a core promoter. Alternatively, the TATA box may be a region including a 5’-TATA(A/T)A(A/T)(A/G)-3’ sequence present in a promoter and/or a core promoter.

For example, the TATA box may be a region including one or more sequences selected form the group consisting of 5’-CATAAAA-3’ sequence sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5’-TGATATCAA-3' sequence, the 5’-TATAAATA-3' sequence, the 5’-TATAAGA-3' sequence, the 5’-AATAAA-3' sequence, the 5’-TTTATA-3' sequence, the 5’-CATAAAAA-3' sequence, the 5’-TATACA-3' sequence, the 5’-TTTAAGA-3' sequence, the 5’-GATAAAG-3' sequence, the 5’-TATAACA-3' sequence, the 5’-TCTTATCTT-3' sequence, the 5’-TTGTACTTT-3' sequence, the 5’-CATATAA-3' sequence, the 5’-TATAAAT-3' sequence, the 5’-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence, which are present in a promoter and/or a core promoter.

For example, the TATA box may be a region in which a TATA-binding protein (TBP) present in a promoter and/or a core promoter binds.

Here, the proximal promoter may include a region 1 to 300-bp upstream of the

TSS, a CpG site and/or a specific transcription factor-binding site.

The enhancer may include an enhancer-box (E-box).

domain of *Streptococcus pyogenes*, *Campylobacter jejuni*, *Streptococcus thermophilus*, *Staphylococcus aureus* or *Neisseria meningitides*, or a first complementary domain derived therefrom.

For example, when the first complementary domain is the first complementary domain of *Streptococcus pyogenes* or a first complementary domain derived therefrom, the first complementary domain may be 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264) or a nucleotide sequence having partial, that is, at least 50% or more, or complete homology with 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264). Here, the first complementary domain may further include $(X)_n$, resulting in 5'-GUUUUAGAGCUA$(X)_n$-3' (SEQ ID NO: 264). The X may be selected from the group consisting of bases A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 5 to 15. Here, the $(X)_n$ may be n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G.

In another embodiment, when the first complementary domain is the first complementary domain of *Campylobacter jejuni* or a first complementary domain derived therefrom, the first complementary domain may be 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265), or a nucleotide sequence having partial, that is, at least 50% or more, or complete homology with 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265). Here, the first complementary domain may further include $(X)_n$, resulting in 5'-GUUUUAGUCCCUUUUUAAAUUUCUU$(X)_n$-3' (SEQ ID NO: 265). The X may be selected from the group consisting of nucleotides A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 5 to 15. Here, the $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G.

For example, when the second complementary domain is a second complementary domain of *Streptococcus pyogenes* or a second complementary domain derived therefrom, the second complementary domain may be 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266) (a nucleotide sequence forming a double strand with the first complementary domain is underlined). Here, the second complementary domain may further include $(X)_n$ and/or $(X)_m$, resulting in 5'-$(X)_n$ UAGCAAGUUAAAAU$(X)_m$-3' (SEQ ID NO: 266). The X may be selected from the group consisting of nucleotides A, T, U and G, and each of the n and m may represent the number of nucleotides, in which the n may be an integer of 1 to 15, and the m may be an integer of 1 to 6. Here, the $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G. In addition, $(X)_m$ may represent m repeats of the same nucleotide, or a mixture of m nucleotides of A, T, U and G.

In another example, when the second complementary domain is the second complementary domain of *Campylobacter jejuni* or a second complementary domain derived therefrom, the second complementary domain may be 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267) (a nucleotide sequence forming a double strand with the first complementary domain is underlined). Here, the second complementary domain may further include $(X)_n$ and/or $(X)_m$, resulting in 5'-$(X)_n$AAGAAAUUUAAAAAGGGACUAAAAU$(X)_m$-3' (SEQ ID NO:

267). The X may be selected from the group consisting of nucleotides A, T, U and G, and each of the n and m may

represent the number of nucleotides, in which the n may be an integer of 1 to 15, and the m may be an integer of 1 to 6. Here, $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G. In addition, $(X)_m$ may represent m repeats of the same nucleotide, or a mixture of m nucleotides of A, T, U and G.

In another embodiment, the second complementary domain may have partial, that is, at least 50% or more, or complete homology with a second complementary domain of *Parcubacteria bacterium* (GWC2011_GWC2_44_17), *Lachnospiraceae bacterium* (MC2017), *Butyrivibrio proteoclasiicus*, *Peregrinibacteria bacterium* (GW2011_GWA_33_10), *Acidaminococcus sp.* (BV3L6), *Porphyromonas macacae*, *Lachnospiraceae bacterium* (ND2006), *Porphyromonas crevioricanis*, *Prevotella disiens*, *Moraxella bovoculi* (237), *Smiihella sp.* (SC_KO8D17), *Leptospira inadai*, *Lachnospiraceae bacterium* (MA2020), *Francisella novicida* (U112), *Candidatus Methanoplasma termitum* or *Eubacterium eligens*, or a second complementary domain derived therefrom.

For example, when the second complementary domain is a second complementary domain of *Parcubacteria bacterium* or a second complementary domain derived therefrom, the second complementary domain may be 5'-AAAUUUCUACU-3' (SEQ ID NO: 268), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-AAAUUUCUACU-3' (SEQ ID NO: 268) (a nucleotide sequence forming a double strand with the first complementary domain is underlined). Here, the second complementary domain may further include $(X)_n$ and/or $(X)_m$, resulting in 5'-$(X)_n$AAAUUUCUACU$(X)_m$-3' (SEQ ID NO: 268). The X may be selected from the group consisting of

may be derived from a proximal domain contained in the species existing in nature, or may have partial or complete homology with the proximal domain contained in the species existing in nature.

In an exemplary embodiment, the proximal domain may have partial, that is, at least 50% or more, or complete homology with a proximal domain of *Streptococcus pyogenes*, *Campylobacter jejuni*, *Streptococcus thermophilus*, *Staphylococcus aureus* or *Neisseria meningitides*, or a proximal domain derived therefrom.

For example, when the proximal domain is a proximal domain of *Streptococcus pyogenes* or a proximal domain derived therefrom, the proximal domain may be 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269). Here, the proximal domain may further include $(X)_n$, resulting in 5'-AAGGCUAGUCCG$(X)_n$-3' (SEQ ID NO: 269). The X may be selected from the group consisting of nucleotides A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 1 to 15. Here, the $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G.

In yet another example, when the proximal domain is a proximal domain of *Campylobacter jejuni* or a proximal domain derived therefrom, the proximal domain may be 5'-AAAGAGUUUGC-3' (SEQ ID NO: 270), or a nucleotide sequence having at least 50% or more homology with 5'-AAAGAGUUUGC-3' (SEQ ID NO: 270). Here, the proximal domain may further include $(X)_n$, resulting in 5'-AAAGAGUUUGC$(X)_n$-3' (SEQ ID NO: 270). The X may be selected from the group consisting of nucleotides A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 1 to 40. Here, the $(X)_n$ may represent

n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G.

Here, the tail domain is a domain which is able to be selectively added to the 3' end of single-stranded gRNA or the first or second strand of double-stranded gRNA.

The tail domain may have homology with a natural tail domain, or may be derived from the natural tail domain. In addition, the tail domain may have a difference in nucleotide sequence according to a species existing in nature, may be derived from a tail domain contained in a species existing in nature, or may have partial or complete homology with a tail domain contained in a species existing in nature.

In one exemplary embodiment, the tail domain may have partial, that is, at least 50% or more, or complete homology with a tail domain of *Streptococcus pyogenes*, *Campylobacter jejuni*, *Streptococcus thermophilus*, *Staphylococcus aureus* or *Neisseria meningitides* or a tail domain derived therefrom.

For example, when the tail domain is a tail domain of *Streptococcus pyogenes* or a tail domain derived therefrom, the tail domain may be 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271). Here, the tail domain may further include $(X)_n$, resulting in 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC$(X)_n$-3' (SEQ ID NO: 271). The X may be selected from the group consisting of nucleotides A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 1 to 15. Here, the $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides such as A, T,

U and G.

In another example, when the tail domain is a tail domain of *Campylobacter jejuni* or a tail domain derived therefrom, the tail domain may be 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' (SEQ ID NO: 272), or a nucleotide sequence having partial, that is, at least 50% or more homology with 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' (SEQ ID NO: 272). Here, the tail domain may further include $(X)_n$, resulting in 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU$(X)_n$-3' (SEQ ID NO: 272). The X may be selected from the group consisting of nucleotides A, T, U and G, and the n may represent the number of nucleotides, which is an integer of 1 to 15. Here, the $(X)_n$ may represent n repeats of the same nucleotide, or a mixture of n nucleotides of A, T, U and G.

In another embodiment, the tail domain may include a 1 to 10-nt sequence at the 3' end involved in an *in vitro* or *in vivo* transcription method.

For example, when a T7 promoter is used in *in vitro* transcription of gRNA, the tail domain may be an arbitrary nucleotide sequence present at the 3' end of a DNA template. In addition, when a U6 promoter is used in *in vivo* transcription, the tail domain may be UUUUUU, when an H1 promoter is used in transcription, the tail domain may be UUUU, and when a pol-III promoter is used, the tail domain may include several uracil nucleotides or alternative nucleotides.

The gRNA may include a plurality of domains as described above, and therefore, the length of the nucleic acid sequence may be regulated according to a

of either strand of a double strand of a nucleic acid in a transcriptional regulatory region of a target gene, and a nucleotide sequence region which may be changed according to a target sequence on a nucleic acid in a transcriptional regulatory region of a target gene.

Here, the $(Q)_m$ is a nucleotide sequence including a first complementary domain, which is able to form a complementary bond with the second complementary domain of the second strand. The $(Q)_m$ may be a sequence having partial or complete homology with the first complementary domain of a species existing in nature, and the nucleotide sequence of the first complementary domain may be changed according to the species of origin. The Q may be each independently selected from the group consisting of A, U, C and G, and the m may be the number of nucleotides, which is an integer of 5 to 35.

For example, when the first complementary domain has partial or complete homology with a first complementary domain of *Streptococcus pyogenes* or a *Streptococcus pyogenes*-derived first complementary domain, the $(Q)_m$ may be 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264), or a nucleotide sequence having at least 50% or more homology with 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264).

In another example, when the first complementary domain has partial or complete homology with a first complementary domain of *Campylobacter jejuni* or a *Campylobacter jejuni*-derived first complementary domain, the $(Q)_m$ may be 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265), or a nucleotide sequence having at least 50% or more homology with 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265).

In still another example, when the first complementary domain has partial or

complete homology with a first complementary domain of *Streptococcus thermophilus* or a *Streptococcus thermophilus*-derived first complementary domain, the $(Q)_m$ may be 5'-GUUUUAGAGCUGUGUUGUUUCG-3' (SEQ ID NO: 273), or a nucleotide sequence having at least 50% or more homology with 5'-GUUUUAGAGCUGUGUUGUUUCG-3' (SEQ ID NO: 273).

In addition, each of the $(X)_a$, $(X)_b$ and $(X)_c$ is selectively an additional nucleotide sequence, where the X may be each independently selected from the group consisting of A, U, C and G, and each of the a, b and c may be the number of nucleotides, which is 0 or an integer of 1 to 20.

In one exemplary embodiment, the second strand may be 5'-$(Z)_h$-$(P)_k$-3'; or 5'-$(X)_d$-$(Z)_h$-$(X)_e$-$(P)_k$-$(X)_f$-3'.

In another embodiment, the second strand may be 5'-$(Z)_h$-$(P)_k$-$(F)_i$-3'; or 5'-$(X)_d$-$(Z)_h$-$(X)_e$-$(P)_k$-$(X)_f$-$(F)_i$-3'.

Here, the $(Z)_h$ is a nucleotide sequence including a second complementary domain, which is able to form a complementary bond with the first complementary domain of the first strand. The $(Z)_h$ may be a sequence having partial or complete homology with the second complementary domain of a species existing in nature, and the nucleotide sequence of the second complementary domain may be modified according to the species of origin. The Z may be each independently selected from the group consisting of A, U, C and G, and the h may be the number of nucleotides, which is an integer of 5 to 50.

For example, when the second complementary domain has partial or complete homology with a second complementary domain of *Streptococcus pyogenes* or a

second complementary domain derived therefrom, the $(Z)_h$ may be 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266), or a nucleotide sequence having at least 50% or more homology with 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266).

In another example, when the second complementary domain has partial or complete homology with a second complementary domain of *Campylobacter jejuni* or a second complementary domain derived therefrom, the $(Z)_h$ may be 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267), or a nucleotide sequence having at least 50% or more homology with 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267).

In still another example, when the second complementary domain has partial or complete homology with a second complementary domain of *Streptococcus thermophilus* or a second complementary domain derived therefrom, the $(Z)_h$ may be 5'-CGAAACAACACAGCGAGUUAAAAU-3' (SEQ ID NO: 274), or a nucleotide sequence having at least 50% or more homology with 5'-CGAAACAACACAGCGAGUUAAAAU-3' (SEQ ID NO: 274).

The $(P)_k$ is a nucleotide sequence including a proximal domain, which may have partial or complete homology with a proximal domain of a species existing in nature, and the nucleotide sequence of the proximal domain may be modified according to the species of origin. The P may be each independently selected from the group consisting of A, U, C and G, and the k may be the number of nucleotides, which is an integer of 1 to 20.

For example, when the proximal domain has partial or complete homology with a proximal domain of *Streptococcus pyogenes* or a proximal domain derived

therefrom, the $(P)_k$ may be 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269), or a nucleotide sequence having at least 50% or more homology with 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269).

5

In another example, when the proximal domain has partial or complete homology with a proximal domain of *Campylobacter jejuni* or a proximal domain derived therefrom, the $(P)_k$ may be 5'-AAAGAGUUUGC-3' (SEQ ID NO: 270), or a nucleotide sequence having at least 50% or more homology with 5'-AAAGAGUUUGC-3' (SEQ ID NO: 270).

In still another example, when the proximal domain has partial or complete homology with a proximal domain of *Streptococcus thermophilus* or a proximal domain derived therefrom, the $(P)_k$ may be 5'-AAGGCUUAGUCCG-3' (SEQ ID NO: 275), or a nucleotide sequence having at least 50% or more homology with 5'-AAGGCUUAGUCCG-3' (SEQ ID NO: 275).

The $(F)_i$ may be a nucleotide sequence including a tail domain, and having partial or complete homology with a tail domain of a species existing in nature, and the nucleotide sequence of the tail domain may be modified according to the species of origin. The F may be each independently selected from the group consisting of A, U, C and G, and the i may be the number of nucleotides, which is an integer of 1 to 50.

For example, when the tail domain has partial or complete homology with a tail domain of *Streptococcus pyogenes* or a tail domain derived therefrom, the $(F)_i$ may be 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271), or a nucleotide sequence having at least 50% or more homology with 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271).

In another example, when the tail domain has partial or complete homology with a tail domain of *Campylobacter jejuni* or a tail domain derived therefrom, the $(F)_i$ may be 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' (SEQ ID NO: 272), or a nucleotide sequence having at least 50% or more homology with 5'-

GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3’ (SEQ ID NO: 272).

In still another example, when the tail domain has partial or complete homology with a tail domain of *Streptococcus thermophilus* or a tail domain derived therefrom, the $(F)_i$ may be 5’-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3’ (SEQ ID NO: 276), or a nucleotide sequence having at least 50% or more homology with 5’-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3’ (SEQ ID NO: 276).

In addition, the $(F)_i$ may include a sequence of 1 to 10 nucleotides at the 3’ end involved in an *in vitro* or *in vivo* transcription method.

For example, when a T7 promoter is used in *in vitro* transcription of gRNA, the tail domain may be an arbitrary nucleotide sequence present at the 3’ end of a DNA template. In addition, when a U6 promoter is used in *in vivo* transcription, the tail domain may be UUUUUU, when an H1 promoter is used in transcription, the tail domain may be UUUU, and when a pol-III promoter is used, the tail domain may include several uracil nucleotides or alternative nucleotides.

In addition, the $(X)_d$, $(X)_e$ and $(X)_f$ may be nucleotide sequences selectively added, where the X may be each independently selected from the group consisting of A, U, C and G, and each of the d, e and f may be the number of nucleotides, which is 0 or an integer of 1 to 20.

Single-stranded gRNA

Single-stranded gRNA may be classified into a first single-stranded gRNA and a second single-stranded gRNA.

region of a target gene, and a nucleotide sequence region capable of being changed according to a target sequence on a transcriptional regulatory region of a target gene.

The $(Q)_m$ includes a nucleotide sequence including the first complementary domain, which is able to form a complementary bond with a second complementary domain. The $(Q)_m$ may be a sequence having partial or complete homology with a first complementary domain of a species existing in nature, and the nucleotide sequence of the first complementary domain may be changed according to the species of origin. The Q may be each independently selected from the group consisting of A, U, C and G, and the m may be the number of nucleotides, which is an integer of 5 to 35.

For example, when the first complementary domain has partial or complete homology with a first complementary domain of *Streptococcus pyogenes* or a first complementary domain derived therefrom, the $(Q)_m$ may be 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264), or a nucleotide sequence having at least 50% or more homology with 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 264).

In another example, when the first complementary domain has partial or complete homology with a first complementary domain of *Campylobacter jejuni* or a first complementary domain derived therefrom, the $(Q)_m$ may be 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265), or a nucleotide sequence having at least 50% or more homology with 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' (SEQ ID NO: 265).

In still another example, when the first complementary domain has partial or complete homology with a first complementary domain of *Streptococcus thermophilus* or a first complementary domain derived therefrom, the $(Q)_m$ may be 5'-GUUUUAGAGCUGUGUUGUUUCG-3' (SEQ ID NO: 273), or a nucleotide sequence having at least 50%

or more homology with 5'-GUUUUAGAGCUGUGUUGUUUCG-3' (SEQ ID NO: 273).

In addition, the $(L)_j$ is a nucleotide sequence including the linker domain, and connecting the first complementary domain with the second complementary domain, thereby producing single-stranded gRNA. Here, the L may be each independently selected from the group consisting of A, U, C and G, and the j may be the number of nucleotides, which is an integer of 1 to 30.

The $(Z)_h$ is a nucleotide sequence including the second complementary domain, and includes a nucleotide sequence capable of complementary binding with the first complementary domain. The $(Z)_h$ may be a sequence having partial or complete homology with the second complementary domain of a species existing in nature, and the nucleotide sequence of the second complementary domain may be changed according to the species of origin. The Z may be each independently selected from the group consisting of A, U, C and G, and the h is the number of nucleotides, which may be an integer of 5 to 50.

For example, when the second complementary domain has partial or complete homology with a second complementary domain of *Streptococcus pyogenes* or a second complementary domain derived therefrom, the $(Z)_h$ may be 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266), or a nucleotide sequence having at least 50% or more homology with 5'-UAGCAAGUUAAAAU-3' (SEQ ID NO: 266).

In another example, when the second complementary domain has partial or complete homology with a second complementary domain of *Campylobacter jejuni* or a second complementary domain derived therefrom, the $(Z)_h$ may be 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267), or a nucleotide sequence having at

least 50% or more homology with 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' (SEQ ID NO: 267).

In still another example, when the second complementary domain has partial or complete homology with a second complementary domain of *Streptococcus thermophilus* or a second complementary domain derived therefrom, the $(Z)_h$ may be 5'-CGAAACAACACAGCGAGUUAAAAU-3' (SEQ ID NO: 274), or a nucleotide sequence having at least 50% or more homology with 5'-CGAAACAACACAGCGAGUUAAAAU-3' (SEQ ID NO: 274).

The $(P)_k$ is a nucleotide sequence including a proximal domain, which may have partial or complete homology with a proximal domain of a species existing in nature, and the nucleotide sequence of the proximal domain may be modified according to the species of origin. The P may be each independently selected from the group consisting of A, U, C and G, and the k may be the number of nucleotides, which is an integer of 1 to 20.

For example, when the proximal domain has partial or complete homology with a proximal domain of *Streptococcus pyogenes* or a proximal domain derived therefrom, the $(P)_k$ may be 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269), or a nucleotide sequence having at least 50% or more homology with 5'-AAGGCUAGUCCG-3' (SEQ ID NO: 269).

In another example, when the proximal domain has partial or complete homology with a proximal domain of *Campylobacter jejuni* or a proximal domain derived therefrom, the $(P)_k$ may be 5'-AAAGAGUUUGC-3' (SEQ ID NO: 270), or a nucleotide sequence having at least 50% or more homology with 5'-

AAAGAGUUUGC-3’ (SEQ ID NO: 270).

In still another example, when the proximal domain has partial or complete homology with a proximal domain of *Streptococcus thermophilus* or a proximal domain derived therefrom, the $(P)_k$ may be 5’-AAGGCUUAGUCCG-3’ (SEQ ID NO: 275), or a

nucleotide sequence having at least 50% or more homology with 5'-AAGGCUUAGUCCG-3' (SEQ ID NO: 275).

The $(F)_i$ may be a nucleotide sequence including a tail domain, and having partial or complete homology with a tail domain of a species existing in nature, and the nucleotide sequence of the tail domain may be modified according to the species of origin. The F may be each independently selected from the group consisting of A, U, C and G, and the i may be the number of nucleotides, which is an integer of 1 to 50.

For example, when the tail domain has partial or complete homology with a tail domain of *Streptococcus pyogenes* or a tail domain derived therefrom, the $(F)_i$ may be 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271), or a nucleotide sequence having at least 50% or more homology with 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 271).

In another example, when the tail domain has partial or complete homology with a tail domain of *Campylobacter jejuni* or a tail domain derived therefrom, the $(F)_i$ may be 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' (SEQ ID NO: 272), or a nucleotide sequence having at least 50% or more homology with 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' (SEQ ID NO: 272).

In still another example, when the tail domain has partial or complete homology with a tail domain of *Streptococcus thermophilus* or a tail domain derived therefrom, the $(F)_i$ may be 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3' (SEQ ID NO: 276), or a nucleotide sequence having at least 50% or more homology with 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3' (SEQ ID NO: 276).

The $(Z)_h$ is a nucleotide sequence including a second complementary domain, and includes a nucleotide sequence capable of complementary binding with a second complementary domain. The $(Z)_h$ may be a sequence having partial or complete homology with the second complementary domain of a species existing in nature, and the nucleotide sequence of the second complementary domain may be modified according to the species of origin. The Z may be each independently selected from the group consisting of A, U, C and G, and the h may be the number of nucleotides, which is an integer of 5 to 50.

For example, when the second complementary domain has partial or complete homology with a second complementary domain of *Parcubacteria bacterium* or a *Parcubacteria bacterium*-derived second complementary domain, the $(Z)_h$ may be 5'-AAAUUUCUACU-3' (SEQ ID NO: 268), or a nucleotide sequence having at least 50% or more homology with 5'-AAAUUUCUACU-3' (SEQ ID NO: 268).

In addition, the $(L)_j$ is a nucleotide sequence including the linker domain, which connects the first complementary domain with the second complementary domain. Here, the L may be each independently selected from the group consisting of A, U, C and G, and the j may be the number of nucleotides, which is an integer of 1 to 30.

In addition, each of the $(X)_a$, $(X)_b$ and $(X)_c$ is selectively an additional nucleotide sequence, where the X may be each independently selected from the group consisting of A, U, C and G, and the a, b and c may be the number of nucleotides, which is 0 or an integer of 1 to 20.

In one exemplary embodiment of the specification, the guide nucleic acid may be gRNA complementarily binding to a target sequence in the transcriptional regulatory region of a duplicate gene.

The "target sequence" is a nucleotide sequence present in the transcriptional regulatory region of a target gene, and specifically, a partial nucleotide sequence in a target region in the transcriptional regulatory region of a target gene, and here, the "target region" is a region that can be modified by a guide nucleic acid-editor protein in the transcriptional regulatory region of a target gene.

Hereinafter, the target sequence may be used to refer to both of two types of nucleotide sequence information. For example, in the case of a target gene, the target sequence may refer to the nucleotide sequence information of a transcribed strand of target gene DNA, or the nucleotide sequence information of a non-transcribed strand.

For example, the target sequence may refer to a partial nucleotide sequence (transcribed strand), that is, 5'-ATCATTGGCAGACTAGTTCG-3' (SEQ ID NO: 277), in the target region of target gene A, and a nucleotide sequence complementary thereto (non-transcribed strand), that is, 5'-CGAACTAGTCTGCCAATGAT-3' (SEQ ID NO: 278).

The target sequence may be a 5 to 50-nt sequence.

In one exemplary embodiment, the target sequence may be a 16-nt sequence, a 17-nt sequence, a 18-nt sequence, a 19-nt sequence, a 20-nt sequence, a 21-nt sequence, a 22-nt sequence, a 23-nt sequence, a 24-nt sequence or a 25-nt sequence.

The target sequence includes a guide nucleic acid-binding sequence or a guide nucleic acid-non binding sequence.

ATCATTGGCAGACTAGTTCG-3’ (SEQ ID NO: 277), of a target region in the transcriptional regulatory region of the target gene A, and a nucleotide sequence, that is, 5’-CGAACTAGTCTGCCAATGAT-3’ (SEQ ID NO: 278), which is complementary thereto, are used as target sequences, the guide nucleic acid-binding sequence may be one of the two target sequences, that is, 5’-ATCATTGGCAGACTAGTTCG-3’ (SEQ ID NO: 277) or 5’-CGAACTAGTCTGCCAATGAT-3’ (SEQ ID NO: 278). Here, when the guide nucleic acid-binding sequence is 5’-ATCATTGGCAGACTAGTTCG-3’ (SEQ ID NO: 277), the guide nucleic acid-non binding sequence may be 5’-CGAACTAGTCTGCCAATGAT-3’ (SEQ ID NO: 278), or when the guide nucleic acid-binding sequence is 5’-CGAACTAGTCTGCCAATGAT-3’ (SEQ ID NO: 278), the guide nucleic acid-non binding sequence may be 5’-ATCATTGGCAGACTAGTTCG-3’ (SEQ ID NO: 277).

The guide nucleic acid-binding sequence may be one of the target sequences, that is, a nucleotide sequence which is the same as a transcribed strand and a nucleotide sequence which is the same as a non-transcribed strand. Here, the guide nucleic acid-non binding sequence may be a nucleotide sequence other than the guide nucleic acid-binding sequence of the target sequences, that is, one selected from a nucleotide sequence which is the same as a transcribed strand and a nucleotide sequence which is the same as a non-transcribed strand.

The guide nucleic acid-binding sequence may have the same length as the target sequence.

The guide nucleic acid-non binding sequence may have the same length as the target sequence or the guide nucleic acid-binding sequence.

The guide nucleic acid-binding sequence may be a 5 to 50-nt sequence.

In one exemplary embodiment, the guide nucleic acid-binding sequence may

located in a region including or near the transcription factor-binding region of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence located in a region including or near the TATA box of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence including the entire or a part of the 5’-TATA(A/T)A(A/T)-3’ sequence, which is present in the core promoter region of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence including the entire or a part of the 5’-TATA(A/T)A(A/T)(A/G)-3’ sequence, which is present in the core promoter region of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence including the entirety or a part of one or more sequences selected from the group consisting of the 5’-CATAAAA-3’ sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5’-TGATATCAA-3' sequence, the 5’-TATAAATA-3' sequence, the 5’-TATAAGA-3' sequence, the 5’-AATAAA-3' sequence, the 5’-TTTATA-3' sequence, the 5’-CATAAAAA-3' sequence, the 5’-TATACA-3' sequence, the 5’-TTTAAGA-3' sequence, the 5’-GATAAAG-3' sequence, the 5’-TATAACA-3' sequence, the 5’-

TCTTATCTT-3' sequence, the 5’-TTGTACTTT-3' sequence, the 5’-CATATAA-3' sequence, the 5’-TATAAAT-3' sequence, the 5’-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence, which are present in the core promoter region of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence, which includes the entire or a part of a TATA-binding protein (TBP)-binding nucleic acid sequence, which is present in the core promoter region of a duplicate gene.

In another example, the target sequence may be a 10 to 25-nt contiguous sequence located in the proximal promoter region of a duplicate gene.

For example, the target sequence may be a 10 to 25-nt contiguous sequence located in a 1 to 300 bp upstream region of the TSS of a duplicate gene.

In still another example, the target sequence may be a 10 to 25-nt contiguous sequence located in the distal promoter region of a duplicate gene.

The target sequence disclosed in the specification may be a 10 to 35-nt contiguous sequence located in the enhancer region of a duplicate gene.

Here, the duplicate gene may be one or more genes selected from the group consisting of a PMP22 gene, a PLP1 gene, an MECP2 gene, a SOX3 gene, an RAI1 gene, a TBX1 gene, an ELN gene, a JAGGED1 gene, an NSD1 gene, an MMP23 gene, an LMB1 gene, a SNCA gene and an APP gene.

Alternatively, the duplicate gene may be one or more genes selected from the group consisting of an MYC gene, an ERBB2 (HER2) gene, a CCND1(Cyclin D1) gene, an FGFR1 gene, an FGFR2 gene, a HRAS gene, a KRAS gene, an MYB gene,

[Table 2]

Target sequences of human PMP22 gene for CjCas9

gRNA	No.	Target (5' to 3')	SEQ ID NO
hPMP22-TATA-Cj#1	1	GCCCTCTGAATCTCCAGTCAAT	SEQ ID NO: 32
hPMP22-TATA-Cj#2	2	AATCTCCAGTCAATTCCAACAC	SEQ ID NO: 33
hPMP22-TATA-Cj#3	3	AATTAGGCAATTCTTGTAAAGC	SEQ ID NO: 34
hPMP22-TATA-Cj#4	4	TTAGGCAATTCTTGTAAAGCAT	SEQ ID NO: 35
hPMP22-TATA-Cj#5	5	AAAGCATAGGCACACATCACCC	SEQ ID NO: 36
hPMP22-TATA-Cj#6	6	GCCTGGTCTGGCTTCAGTTACA	SEQ ID NO: 37
hPMP22-TATA-Cj#7	7	GTGTCCAACTTTGTTTGCTTTC	SEQ ID NO: 38
hPMP22-TATA-Cj#8	8	GTATTCTGGAAAGCAAACAAAG	SEQ ID NO: 39
hPMP22-TATA-Cj#9	9	CAGTCTTGGCATCACAGGCTTC	SEQ ID NO: 40
hPMP22-TATA-Cj#10	10	GGACCTCTTGGCTATTACACAG	SEQ ID NO: 41
hPMP22-TATA-Cj#11	11	GGAGCCAGTGGGACCTCTTGGC	SEQ ID NO: 42
hPMP22-Enh-Cj#1	12	TAAATCACAGAGGCAAAGAGTT	SEQ ID NO: 43
hPMP22-Enh-Cj#2	13	TTGCATAGTGCTAGACTGTTTT	SEQ ID NO: 44
hPMP22-Enh-Cj#3	14	GGGTCATGTGTTTTGAAAACAG	SEQ ID NO: 45
hPMP22-Enh-Cj#4	15	CCCAAACCTCCCAACCCACAAC	SEQ ID NO: 46
hPMP22-Enh-Cj#5	16	ACTCAGCTATTTCTGGAATGAC	SEQ ID NO: 47
hPMP22-Enh-Cj#6	17	TCATCGCCTTTGTGAGCTCCAT	SEQ ID NO: 48
hPMP22-Enh-Cj#7	18	CAGACACAGGCTTTGCTCTAGC	SEQ ID NO: 49
hPMP22-Enh-Cj#8	19	CAAAGCCTGTGTCTGGCCACTA	SEQ ID NO: 50
hPMP22-Enh-Cj#9	20	AGCAGTTTGTGCCCACTAGTGG	SEQ ID NO: 51
hPMP22-Enh-Cj#10	21	ATGTCAAGGTATTCCAGCTAAC	SEQ ID NO: 52
hPMP22-Enh-Cj#11	22	GAATAACTGTATCAAAGTTAGC	SEQ ID NO: 53
hPMP22-Enh-Cj#12	23	TTCCTAATTAAGAGGCTTTGTG	SEQ ID NO: 54
hPMP22-Enh-Cj#13	24	GAGCTAGTTTGTCAGGGTCTAG	SEQ ID NO: 55

[Table 3]

Target sequences of human PLP1 gene for SpCas9

sgRNA	No.	Target with PAM(5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
hPLP1-TATA-Sp-01	1	GACTTTGGGAGCTAATATCTAGG	SEQ ID NO: 56	+	1	0	0	-
hPLP1-wMN1-Sp-01	2	CCCTTTCATCTTCCCATTCGTGG	SEQ ID NO: 57	+	1	0	0	Up

hPLP1-wMN1-Sp-22	23	ATTTGACTATCCAGCAGGCTTGG	SEQ ID NO: 78	+	1	0	0	Up
hPLP1-wMN1-Sp-23	24	GTCCCGAAGTCTCTGGGGCCTGG	SEQ ID NO: 79	-	1	0	0	Up
hPLP1-wMN1-Sp-24	25	AAAACAGTCCCGAAGTCTCTGGG	SEQ ID NO: 80	-	1	0	0	Up
hPLP1-wMN1-Sp-25	26	GAAAACAGTCCCGAAGTCTCTGG	SEQ ID NO: 81	-	1	0	0	Up
hPLP1-wMN1-Sp-26	27	TATATACCACATTCAAGTGCTGG	SEQ ID NO: 82	-	1	0	0	Up
hPLP1-wMN1-Sp-27	28	TGGATATAACGAAGTTGTGTGGG	SEQ ID NO: 83	-	1	0	0	Down
hPLP1-wMN1-Sp-28	29	ATGGATATAACGAAGTTGTGTGG	SEQ ID NO: 84	-	1	0	0	Down
hPLP1-wMN1-Sp-29	30	ATATGTTTGTTCACCCCAACAGG	SEQ ID NO: 85	+	1	0	0	Down
hPLP1-wMN1-Sp-30	31	GAAAACTTGAAATCCTGTTGGGG	SEQ ID NO: 86	-	1	0	0	Down
hPLP1-wMN1-Sp-31	32	TAGACATTAGGAGAAACAGAAGG	SEQ ID NO: 87	-	1	0	0	Down
hPLP1-wMN1-Sp-32	33	CTAGCAGTGACATAGACATTAGG	SEQ ID NO: 88	-	1	0	0	Down
hPLP1-wMN1-Sp-33	34	AGCCACCTGACTTTGATGAAAGG	SEQ ID NO: 89	-	1	0	0	Down
hPLP1-wMN1-Sp-34	35	TGAGAAATGTTATTACTATATGG	SEQ ID NO: 90	+	1	0	0	Down
hPLP1-wMN1-Sp-35	36	AGACTGCGAGATGAGAGAGTTGG	SEQ ID NO: 91	-	1	0	0	Down
hPLP1-wMN1-Sp-36	37	CTCGCAGTCTGTACTTAGACTGG	SEQ ID NO: 92	+	1	0	0	Down
hPLP1-wMN1-Sp-37	38	AATGTCTCTTGAGAGAGCCAAGG	SEQ ID NO: 93	+	1	0	0	Down

[Table 4]

Target sequences of human PLP1 gene for CjCas9

sgRNA	**No.**	**Target with PAM (5' to 3')**	**SEQ ID NO**	**Direction**	**Mismatch 0**	**Mismatch 1**	**Mismatch 2**	**location**

hPLP1-wMN1-Cj-21	21	ATGCCAACTCTCTCATCTCGCAGTCTGTAC	SEQ ID NO: 114	+	1	0	0	Down
hPLP1-wMN1-Cj-22	22	GAGACATTCTCACATTTCCAGTCTAAGTAC	SEQ ID NO: 115	-	1	0	0	Down

[Table 5]

Target sequences of mouse Plp1 gene for SpCas9

sgRNA	**No.**	**Target with PAM (5' to 3')**	**SEQ ID NO**	**Direction**	**Mismatch 0**	**Mismatch 1**	**Mismatch 2**	**location**
mPlp1-TATA-Sp-01	1	TGTTTGGTAGTATAGTAAGTAGG	SEQ ID NO: 116	+	1	0	1	-
mPlp1-wMN1-Sp-01	2	GGTCTAGAAAAGATCAAGCCAGG	SEQ ID NO: 117	+	1	0	0	Up
mPlp1-wMN1-Sp-02	3	GCCAGGACTGTGACCTGATAAGG	SEQ ID NO: 118	+	1	0	0	Up
mPlp1-wMN1-Sp-03	4	TCACCTTCACACTTTAACCAAGG	SEQ ID NO: 119	+	1	0	0	Up
mPlp1-wMN1-Sp-04	5	CAAGGTTGAGACAATGTTCCAGG	SEQ ID NO: 120	+	1	0	0	Up
mPlp1-wMN1-Sp-05	6	CCAATTCATGTGCAAACATTTGG	SEQ ID NO: 121	-	1	0	0	Up
mPlp1-wMN1-Sp-06	7	CATCACAGTTTATACTTAGCTGG	SEQ ID NO: 122	+	1	0	0	Down
mPlp1-wMN1-Sp-07	8	ATCACAGTTTATACTTAGCTGGG	SEQ ID NO: 123	+	1	0	0	Down
mPlp1-wMN1-Sp-08	9	GGAATACCTCAGGCTCAACAGGG	SEQ ID NO: 124	-	1	0	0	Down
mPlp1-wMN1-Sp-09	10	TCTCTGTTTCGGAATACCTCAGG	SEQ ID NO: 125	-	1	0	0	Down
mPlp1-wMN1-Sp-10	11	CTGTCGACTACTTTGATGAAAGG	SEQ ID NO: 126	+	1	0	0	Down
mPlp1-wMN1-Sp-11	12	TGAACCAAGATGATTATTTGTGG	SEQ ID NO: 127	-	1	0	1	Up
mPlp1-wMN1-Sp-12	13	ATCTTGGTTCATAGAAATTTGGG	SEQ ID NO: 128	+	1	0	1	Up
mPlp1-wMN1-Sp-13	14	AGCCTTGCATGGCAGAGCTTGGG	SEQ ID NO: 129	-	1	0	1	Up
mPlp1-wMN1-Sp-14	15	ACACTTTAACCAAGGAAAGAGGG	SEQ ID NO: 130	+	1	0	1	Up

mPlp1-wMN1-Sp-35	36	TACAGATTGGTTACACTTGGGGG	SEQ ID NO: 151	+	1	0	1	Down
mPlp1-wMN1-Sp-36	37	ACAGATTGGTTACACTTGGGGGG	SEQ ID NO: 152	+	1	0	1	Down

[Table 6]

Target sequences of mouse Plp1 gene for CjCas9

sgRNA	No.	Target with PAM (5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
mPlp1-TATA-Cj-01	1	CTACTTACTATACTACCAAACACACCGCAC	SEQ ID NO: 153	-	1	0	0	-
mPlp1-TATA-Cj-02	2	AAAGCCTACTTACTATACTACCAAACACAC	SEQ ID NO: 154	-	1	0	0	-
mPlp1-TATA-Cj-03	3	CAAAAGCCTACTTACTATACTACCAAACAC	SEQ ID NO: 155	-	1	0	0	-
mPlp1-TATA-Cj-04	4	GGGTCTGAATCAAAAGCCTACTTACTATAC	SEQ ID NO: 156	-	1	0	0	-
mPlp1-wMN1-Cj-01	5	AGAGTGGGATTCTACAAGTCACCTTCACAC	SEQ ID NO: 157	+	1	0	0	Up
mPlp1-wMN1-Cj-02	6	GGAAAGAGGGGATCTGGTAGCATAAAGTAC	SEQ ID NO: 158	+	1	0	0	Up
mPlp1-wMN1-Cj-03	7	GGGATCTGGTAGCATAAAGTACAGCTACAC	SEQ ID NO: 159	+	1	0	0	Up
mPlp1-wMN1-Cj-04	8	ATCTGTCACTAGCGACAAGTGTAGCTGTAC	SEQ ID NO: 160	-	1	0	0	Up
mPlp1-wMN1-Cj-05	9	TCATGTGCAAACATTTGGAGGCCAAAATAC	SEQ ID NO: 161	-	1	0	0	Up
mPlp1-wMN1-Cj-06	10	GACATACAGAGAGGGGGCGGAGAGAAATAC	SEQ ID NO: 162	-	1	0	0	Up
mPlp1-wMN1-Cj-07	11	ATACTGACGCCATCACATCACAGTTTATAC	SEQ ID NO: 163	+	1	0	0	Down
mPlp1-wMN1-Cj-08	12	TAAAACTATAAGCTCTCTGTTTCGGAATAC	SEQ ID NO: 164	-	1	0	0	Down
mPlp1-wMN1-Cj-09	13	TCATCAAAGTAGTCGACAGTCAAAGCATAC	SEQ ID NO: 165	-	1	0	0	Down
mPlp1-wMN1-Cj-10	14	TGAATTCTAACAGGAAAACTCAGAACATAC	SEQ ID NO: 166	-	1	0	0	Down
mPlp1-wMN1-Cj-11	15	ACTGCTGCTACTACTTATGGTGACTAGTAC	SEQ ID NO: 167	-	1	0	0	Up

Here, the functional domain, peptide, polypeptide or protein may be a domain, peptide, polypeptide or protein having methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, RNA cleavage activity or nucleic acid binding activity, or a tag or reporter gene for isolation and purification of a protein (including a peptide), but the present invention is not limited thereto.

The functional domain, peptide, polypeptide or protein may be a deaminase.The tag includes a histidine (His) tag, a V5 tag, a FLAG tag, an influenza hemagglutinin (HA) tag, a Myc tag, a VSV-G tag and a thioredoxin (Trx) tag, and the reporter gene includes glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) β-galactosidase, β-glucoronidase, luciferase, autofluorescent proteins including the green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP) and blue fluorescent protein (BFP), but the present invention is not limited thereto.

In addition, the functional domain, peptide, polypeptide or protein may be a nuclear localization sequence or signal (NLS) or a nuclear export sequence or signal (NES).

The NLS may be NLS of SV40 virus large T-antigen with an amino acid sequence PKKKRKV (SEQ ID NO: 279); NLS derived from nucleoplasmin (e.g., nucleoplasmin bipartite NLS with a sequence KRPAATKKAGQAKKKK (SEQ ID NO: 280); c-myc NLS with an amino acid sequence PAAKRVKLD (SEQ ID NO: 281) or RQRRNELKRSP (SEQ ID NO: 282); hRNPA1 M9 NLS with a sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 283); an importin-α-derived IBB domain sequence

RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 284); myoma T protein sequences VSRKRPRP (SEQ ID NO: 285) and PPKKARED (SEQ ID NO: 286); human p53 sequence PQPKKKPL (SEQ ID NO: 287); a mouse c-abl IV sequence SALIKKKKKMAP (SEQ ID NO: 288); influenza virus NS1 sequences DRLRR (SEQ ID NO: 289) and PKQKKRK (SEQ ID NO: 290); a hepatitis virus-δ antigen sequence RKLKKKIKKL (SEQ ID NO: 291); a mouse Mx1 protein sequence REKKKFLKRR (SEQ ID NO: 292); a human poly(ADP-ribose) polymerase sequence KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 293); or steroid hormone receptor (human) glucocorticoid sequence RKCLQAGMNLEARKTKK (SEQ ID NO: 294), but the present invention is not limited thereto.

The additional domain, peptide, polypeptide or protein may be a non-functional domain, peptide, polypeptide or protein that does not perform a specific function. Here, the non-functional domain, peptide, polypeptide or protein may be a domain, peptide, polypeptide or protein that does not affect the enzyme function.

The fusion protein may be a form in which the non-functional domain, peptide, polypeptide or protein is added to one or more of the amino end of an enzyme or the proximity thereof; the carboxyl end of the enzyme or the proximity thereof; the middle part of the enzyme; or a combination thereof.

The editor protein may be a natural enzyme or fusion protein.

The editor protein may be present in the form of a partially modified natural enzyme or fusion protein.

The editor protein may be an artificially produced enzyme or fusion protein, which does not exist in nature.

The editor protein may be present in the form of a partially modified artificial

expression control composition having the donor, and thus highly expressed.

Optionally, the donor may include an additional nucleotide sequence. Here, the additional nucleotide sequence may serve to increase the stability of the donor, the efficiency of insertion into a target, or homologous recombination efficiency.

For example, the additional nucleotide sequence may be an A and T nucleotide-rich nucleic acid sequence, that is, an A-T rich domain. For example, the additional nucleotide sequence may be a scaffold/matrix attachment region (SMAR).

The guide nucleic acid, editor protein or guide nucleic acid-editor protein complex disclosed in the specification may be delivered or introduced into a subject in various ways.

Here, the term "subject" refers to an organism into which a guide nucleic acid, editor protein or guide nucleic acid-editor protein complex is introduced, an organism in which a guide nucleic acid, editor protein or guide nucleic acid-editor protein complex operates, or a specimen or sample obtained from the organism.

The subject may be an organism including a target gene or chromosome of a guide nucleic acid-editor protein complex.

The organism may be an animal, animal tissue or an animal cell.

The organism may be a human, human tissue or a human cell.

The tissue may be eyeball, skin, liver, kidney, heart, lung, brain, muscle tissue, or blood.

The cell may be a fibroblast, a Schwann cell, a nerve cell, an oligodendrocyte, a myoblast, a glial cell, a macrophage, an immune cell, a hepatocyte, a retinal pigment epithelial cell, a cancer cell or a stem cell.

combination thereof. The NLS may be an NLS sequence derived from the following NLSs, but the present invention is not limited thereto: NLS of a SV40 virus large T-antigen having the amino acid sequence PKKKRKV (SEQ ID NO: 279); NLS from nucleoplasmin (e.g., nucleoplasmin bipartite NLS having the sequence KRPAATKKAGQAKKKK (SEQ ID NO: 280)); c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: 281) or RQRRNELKRSP (SEQ ID NO: 282); hRNPA1 M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 283); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 284) of the IBB domain from importin-α; the sequences VSRKRPRP (SEQ ID NO: 285) and PPKKARED (SEQ ID NO: 286) of a myoma T protein; the sequence PQPKKKPL (SEQ ID NO: 287) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO: 288) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO: 289) and PKQKKRK (SEQ ID NO: 290) of influenza virus NS1; the sequence RKLKKKIKKL (SEQ ID NO: 291) of a hepatitis delta virus antigen; the sequence REKKKFLKRR (SEQ ID NO: 292) of a mouse Mx1 protein; the sequence KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 293) of a human poly (ADP-ribose) polymerase; or the NLS sequence RKCLQAGMNLEARKTKK (SEQ ID NO: 294), derived from a sequence of a steroid hormone receptor (human) glucocorticoid.

In addition, the CRISPR enzyme mutant may include a split-type CRISPR enzyme prepared by dividing the CRISPR enzyme into two or more parts. The term "split" refers to functional or structural division of a protein or random division of a protein into two or more parts.

The split-type CRISPR enzyme may be a complete, incomplete or partially active enzyme or inactive enzyme.

For example, when the CRISPR enzyme is a SpCas9, the SpCas9 may be

Dejerine-Sottas disease (DSD), congenital hypomyelination neuropathy (CHN), Roussy-Levy syndrome (RLS), Pelizaeus-Merzbacher disease (PMD), MECP2 duplication syndrome, X-linked hypopituitarism (XLHP), Potocki-Lupski syndrome (PTLS), velocardiofacial syndrome (VCFS), Williams Beuren syndrome (WBS), Alagille syndrome (AS), growth retardation syndrome, premature closure cranial sutures, autosomal dominant leukodystrophy (ADLD), Parkinson's disease or Alzheimer's disease.

In addition, the gene duplication disease may be breast cancer, cervical cancer, colorectal cancer, esophageal cancer, gastric cancer, glioblastoma, head and neck cancer, hepatocellular cancer, neuroblastoma, ovarian cancer, a sarcoma or small cell lung cancer.

The expression control composition may be administered to a treatment subject with a gene duplication disease.

The treatment subject may include mammals including a human, a primate such as monkey, and a rodent such as a mouse and a rat.

The expression control composition may be administered to the treatment subject.

The administration may be performed by injection, transfusion, implantation or transplantation.

The administration may be performed via an administration route selected from intraneural, subretinal, subcutaneously, intradermal, intraocular, intravitreal, intratumoral, intranodal, intramedullary, intramuscular, intravenous, intralymphatic, and intraperitoneal routes.

A dose of the expression control composition (a pharmaceutically effective

GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGC-3' (SEQ ID NO: 296).

2. Construction and Synthesis of gRNA

sgRNA was packed into an AAV vector or synthesized with RNA. In order to insert the sgRNA into a viral vector, a DNA oligonucleotide corresponding to 20 to 22 base sequences of the sgRNA was designed and annealed, and ligated into a pRGEN-CAS9 (developed in-house) vector using a BsmBI site. Cas9 and the sgRNA including a variable target sequence at the 5' end were expressed through the CMV and U6 promoters, respectively.

Furthermore, for a delivery system by RNP, the sgRNA was transcribed by T7 RNA polymerase after a template was produced by annealing two partially complementary oligonucleotides produced by Phusion Taq-mediated polymerization. The transcribed sgRNA was purified and quantified using spectrometry.

3. Purification of Cas9 Protein

Codon-optimized Cas9 DNA sequences including NLS and HA epitopes were subcloned into a pET28 vector and expressed in BL21 (DE3) using IPTG under optimal culture conditions. The expressed Cas9 protein was purified using Ni-NTA agarose beads and dialyzed with an appropriate buffer. The activity of Cas9 was confirmed through an in vitro cleavage test using a well-known effective sgRNA.

4. Cell Culture

WHAT IS CLAIMED IS:

1 A method for regulating an expression of a duplicate gene in a eukaryotic cell, comprising introducing an expression control composition into the eukaryotic cell,

wherein the expression control composition comprises the following:

a guide RNA or nucleic acid encoding the same, wherein the guide RNA targets a target region present in a transcriptional regulatory region of the duplicate gene,

wherein the duplicate gene may be one or more genes selected from the group consisting of a PMP22 gene, a PLP1 gene, a MECP2 gene, a SOX3 gene, a RAI1 gene, a TBX1 gene, an ELN gene, a JAGGED1 gene, a NSD1 gene, a MMP23 gene, a LMB1 gene, a SNCA gene and an APP gene, a MYC gene, an ERBB2 (HER2) gene, a CCND1 (Cyclin D1) gene, a FGFR1 gene, a FGFR2 gene, a HRAS gene, a KRAS gene, a MYB gene, a MDM2 gene, a CCNE (Cyclin E) gene, and an AKT2 gene,

wherein the transcriptional regulatory region is a promoter or an enhancer,

wherein the target region has two strands having a guide nucleic acid -binding sequence and a guide nucleic acid non-binding sequence, which are sequences of 15 to 25 continuous nucleotides,

wherein the guide RNA comprises a guide sequence complementarily binding with the guide nucleic acid-binding sequence of the target region,

wherein the guide sequence is an RNA sequence having homology with the guide nucleic acid non-binding sequence of the target region,

wherein the guide sequence includes 0 to 5 nucleotides mismatched with the guide nucleic acid-binding sequence of the target region,

wherein the guide RNA comprises a nucleotide sequence including one or more

sequences selected from the group consisting of SEQ ID NOs: 264 to 276, SEQ ID NO: 295 and SEQ ID NO: 296, or a nucleotide sequence having 80% or more homologous therewith; and

a CRISPR enzyme or a nucleic acid encoding the same,

wherein the CRISPR enzyme is a Cas9 protein or a Cpf1 protein,

wherein the the Cas9 protein is at least one selected from the group consisting of a *Streptococcus pyogenes*-derived Cas9 protein, a *Campylobacter jejuni*-derived Cas9 protein and *Streptococcus thermophilus*-derived Cas9 protein,

whereby the expression of the duplicate gene is decreased in the eukaryotic cell

2. The method of claim 1,

wherein the promoter is a promoter including a TATA-box.

3. The method of claim 2,

wherein the guide nucleic acid-binding sequence or the guide nucleic acid-non binding sequence of the target region includes at least one TATA-box sequence selected from the group consisting of the 5'-TATA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5'-TGATATCAA-3' sequence, the 5'-TATAAATA-3' sequence, the 5'-TATAAGA-3' sequence, the 5'-AATAAA-3' sequence, the 5'-TTTATA-3' sequence, the 5'-CATAAAAA-3' sequence, the 5'-TATACA-3' sequence, the 5'-TTTAAGA-3' sequence, the 5'-GATAAAG-3' sequence, the 5'-TATAACA-3' sequence, the 5'-TCTTATCTT-3' sequence, the 5'-TTGTACTTT-3' sequence, the 5'-CATATAA-3' sequence, the 5'-TATAAAT-3' sequence, the 5'-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence.

4. The method of claim 3,

wherein the guide sequence is a RNA sequence having homology with the guide nucleic acid-non binding sequence including at least one TATA-box sequence selected from the group consisting of the 5'-TATA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5'-TGATATCAA-3' sequence, the 5'-TATAAATA-3' sequence, the 5'-TATAAGA-3' sequence, the 5'-AATAAA-3' sequence, the 5'-TTTATA-3' sequence, the 5'-CATAAAAA-3' sequence, the 5'-TATACA-3' sequence, the 5'-TTTAAGA-3' sequence, the 5'-GATAAAG-3' sequence, the 5'-TATAACA-3' sequence, the 5'-TCTTATCTT-3' sequence, the 5'-TTGTACTTT-3' sequence, the 5'-CATATAA-3' sequence, the 5'-TATAAAT-3' sequence, the 5'-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence.

5. The method of claim 1,

wherein the guide sequence is a RNA sequence having homology with one guide nucleic acid-non binding sequence selected from the group consisting of SEQ ID NOs: 1 to 29 and SEQ ID NOs: 32 to 115,

wherein the guide nucleic acid-non binding sequence do not include a PAM sequence.

6. The method of claim 1,

wherein the expression control composition includes the guide RNA and the CRISPR enzyme in a form of a guide RNA-CRISPR enzyme complex.

7. The method of claim 1,

wherein the expression control composition is in a form of a vector including the nucleic acid encoding the guide RNA and the CRISPR enzyme, respectively.

8. The method of claim 1,

wherein the introducing of the expression control composition is performed by one or more methods selected from electroporation, liposomes, plasmids, viral vectors,

nanoparticles and a protein translocation domain (PTD) fusion protein method.

9. A guide RNA for targeting a duplicate gene, or nucleic acid encoding the same,

wherein the guide RNA targets a target region present in a transcriptional regulatory region of the duplicate gene,

wherein the transcriptional regulatory region is a promoter or an enhancer,

wherein the target region has two strands having a guide nucleic acid-binding sequence and a guide nucleic acid-non binding sequence, which are sequences of 15 to 25 contiguous nucleotides,

wherein the guide RNA comprises a guide sequence complementarily binding with the guide nucleic acid-binding sequence of the target region,

wherein the guide sequence is a RNA sequence having homology with the guide nucleic acid-non binding sequence of the target region,

wherein the guide nucleic acid-non binding sequence is a sequence including one selected from the group consisting of the 5'-TATA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5'-TGATATCAA-3' sequence, the 5'-TATAAATA-3' sequence, the 5'-TATAAGA-3' sequence, the 5'-AATAAA-3' sequence, the 5'-TTTATA-3' sequence, the 5'-CATAAAAA-3' sequence, the 5'-TATACA-3' sequence, the 5'-TTTAAGA-3' sequence, the 5'-GATAAAG-3' sequence, the 5'-TATAACA-3' sequence, the 5'-TCTTATCTT-3' sequence, the 5'-TTGTACTTT-3' sequence, the 5'-CATATAA-3' sequence, the 5'-TATAAAT-3' sequence, the 5'-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and 5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence, or is one sequence selected from the group consisting of SEQ ID NOs: 1 to 29 and SEQ ID NOs: 32 to 115,

wherein the guide nucleic acid-non binding sequence do not include a PAM sequence,

wherein the guide sequence includes 0 to 5 nucleotides mismatched with the guide nucleic acid-binding sequence of the target region,

wherein the guide RNA comprises a nucleotide sequence including or more sequences selected from the group consisting of SEQ ID NOs: 264 to 276, SEQ ID NO: 295 and SEQ ID NO: 296, or a nucleotide sequence having 80% or more

homologous therewith.

10. The guide RNA of claim 9,

wherein the nucleic acid encoding the guide RNA is included in a vector.

11. The guide RNA of claim 10,

wherein the vector is a plasmid or a viral vector.

12. A composition for regulating an expression of a duplicate gene, comprising:

a guide RNA for targeting a duplicate gene, or nucleic acid encoding the same,

wherein the guide RNA targets a target region present in a transcriptional regulatory region of the duplicate gene,

wherein the transcriptional regulatory region is a promoter or an enhancer,

wherein the target region has two strands having a guide nucleic acid-binding sequence and a guide nucleic acid-non binding sequence, which are sequences of 15 to 25 contiguous nucleotides,

wherein the guide RNA comprises a guide sequence complementarily binding with the guide nucleic acid-binding sequence of the target region,

wherein the guide sequence is a RNA sequence having homology with the guide nucleic acid-non binding sequence of the target region,

wherein the guide nucleic acid-non binding sequence is a sequence including one selected from the group consisting of the 5'-TATA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-CATAAAA-3' sequence, the 5'-TATAA-3' sequence, the 5'-TATAAAA-3' sequence, the 5'-CATAAATA-3' sequence, the 5'-TATATAA-3' sequence, the 5'-TATATATATATATAA-3' (SEQ ID NO: 261) sequence, the 5'-TATATTATA-3' sequence, the 5'-TATAAA-3' sequence, the 5'-TATAAAATA-3' sequence, the 5'-TATATA-3' sequence, the 5'-GATTAAAAA-3' sequence, the 5'-TATAAAAA-3' sequence, the 5'-TTATAA-3' sequence, the 5'-TTTTAAAA-3' sequence, the 5'-TCTTTAAAA-3' sequence, the 5'-GACATTTAA-3' sequence, the 5'-TGATATCAA-3' sequence, the 5'-TATAAATA-3' sequence, the 5'-TATAAGA-3' sequence, the 5'-AATAAA-3' sequence, the 5'-TTTATA-3' sequence, the 5'-CATAAAAA-3' sequence, the 5'-TATACA-3' sequence, the 5'-TTTAAGA-3' sequence, the 5'-GATAAAG-3' sequence, the 5'-TATAACA-3' sequence, the 5'-TCTTATCTT-3' sequence, the 5'-TTGTACTTT-3' sequence, the 5'-CATATAA-3' sequence, the 5'-TATAAAT-3' sequence, the 5'-TATATATAAAAAAAA-3' (SEQ ID NO: 262) sequence and

5'-CATAAATAAAAAAAATTA-3' (SEQ ID NO: 263) sequence, or is one sequence selected from the group consisting of SEQ ID NOs: 1 to 29 and SEQ ID NOs: 32 to 115,

wherein the guide nucleic acid-non binding sequence do not include a PAM sequence,

wherein the guide sequence includes 0 to 5 nucleotides mismatched with the guide nucleic acid-binding sequence of the target region,

wherein the guide RNA comprises a nucleotide sequence including or more sequences selected from the group consisting of SEQ ID NOs: 264 to 276, SEQ ID NO: 295 and SEQ ID NO: 296, or a nucleotide sequence having 80% or more homologous therewith; and

a CRISPR enzyme or a nucleic acid encoding the same,

wherein the CRISPR enzyme is a Cas9 protein or a Cpf1 protein,

wherein the Cas9 protein is at least one selected from the group consisting of a *Streptococcus pyogenes*-derived Cas9 protein, a *Campylobacter jejuni*-derived Cas9 protein and a *Streptococcus thermophilus*-derived Cas9 protein.

13. The composition of claim 12,

wherein the composition includes the guide RNA and the CRISPR enzyme in a form of a guide RNA-CRISPR enzyme complex.

14. The composition of claim 12,

wherein the composition is in a form of a vector including the nucleic acid encoding the guide RNA and the CRISPR enzyme, respectively.

15. (New) The composition of claim 14,

wherein the vector is a plasmid or a viral vector.