



(12)

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

(11)

**EP 3 889 266 A1**

(43) Date of publication:

**06.10.2021 Bulletin 2021/40**

(51) Int Cl.:

**C12N 15/63 (2006.01)****C12N 15/11 (2006.01)**(21) Application number: **20884222.9**

(86) International application number:

**PCT/CN2020/120633**(22) Date of filing: **13.10.2020**

(87) International publication number:

**WO 2021/088601 (14.05.2021 Gazette 2021/19)**

(84) Designated Contracting States:

**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB  
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO  
PL PT RO RS SE SI SK SM TR**

Designated Extension States:

**BA ME**

Designated Validation States:

**KH MA MD TN**(30) Priority: **07.11.2019 CN 201911081617****15.08.2020 CN 202010821877****16.09.2020 CN 202010974151**(71) Applicant: **Qingdao KingAgroot Chemical Compound Co., Ltd.****Qingdao, Shandong 266000 (CN)**

- **MO, Sudong**  
**Qingdao, Shandong 266000 (CN)**
- **WANG, Jiyao**  
**Qingdao, Shandong 266000 (CN)**
- **LI, Yucai**  
**Qingdao, Shandong 266000 (CN)**
- **QI, Wei**  
**Qingdao, Shandong 266000 (CN)**
- **LI, Huarong**  
**Qingdao, Shandong 266000 (CN)**
- **CHEN, Bo**  
**Qingdao, Shandong 266000 (CN)**

(74) Representative: **Finnegan Europe LLP**  
**1 London Bridge**  
**London SE1 9BG (GB)**

(72) Inventors:

- **JIANG, Linjian**  
**Qingdao, Shandong 266000 (CN)**

---

**(54) METHOD FOR GENERATING NEW MUTATIONS IN ORGANISMS, AND APPLICATION THEREOF**

---

(57) The present invention pertains to the technical field of genetic engineering, and specifically relates to a method for generating a site-specific mutation in an organism in the absence of an artificial DNA template and a use thereof. The method comprises the following steps: sequentially generating two or more DNA breaks at a specific site in a genome of an organism and spontaneously repairing them respectively, wherein a later DNA break is generated based on a new sequence generated

from a previous DNA break repair. In the present invention, a new target is designed based on a sequence formed by a new repair event generated by sequential editing, and thus, mutations may be sequentially formed for multiple times at a specific site in the genome, which greatly enrich the types of repair events after DNA breaks, and realize new base substitution, deletion and insertion mutations that cannot be obtained in a single gene editing.

**Description****Priority information**

5 [0001] The present application claims the benefits of the Chinese invention patent application No. 201911081617.X filed on November 7, 2019, the Chinese invention patent application No. 202010821877.2 filed on August 15, 2020, and the Chinese invention patent application No. 202010974151.2 filed on September 16, 2020. The above applications are incorporated herein by reference in their entireties.

**10 Technical Field**

[0002] The present invention pertains to the technical field of genetic engineering, and specifically relates to a method for generating a site-specific mutation in an organism in the absence of an artificial DNA template and a use thereof.

**15 Background Art**

[0003] The genetic engineering technology for modifying genome of organisms has been widely used in industrial and agricultural production, such as genetically modified microorganisms commonly used in the pharmaceutical and chemical fields, and genetically modified crops with insect-resistance and herbicide-resistance in the agricultural field. With the 20 advent of site-specific nucleases, by introducing a targeted fragmentation into the genome of recipient organism and causing spontaneous repair, it has been possible to achieve site-specific editing of genome and more precise modification of genome.

[0004] Gene editing tools mainly include three types of sequence-specific nuclease (SSN): zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR) associated Cas system (CRISPR/Cas system). Sequence-specific nucleases are programmable nucleases that can generate DNA double-strand breaks (DSBs) at specific sites in genome. DNA double-strand breaks activate the endogenous DNA repair pathway to repair DNA damage in a cell, but the repair process easily leads to changes in the DNA sequence at target sites, thereby achieving the introduction of mutations at interesting sites. This technology enables biologists to accurately target a target gene and edit it. Among them, both ZFN and TALEN need to design 30 specific recognition protein modules for the target sequence, thereby having low throughput and complex operations. However, Cas protein is universal in the CRISPR/Cas system, in which a guide RNA (gRNA) can be formed by a specific CRISPR-RNA (crRNA) designed for a target site alone or in conjunction with transactivating RNA (tracrRNA), or a single 35 guide RNA (sgRNA) alone is enough, the crRNA and tracrRNA together or sgRNA alone can be assembled with the Cas protein to form a ribonucleoprotein complex (RNP), the target sequence is identified on the basis of protospacer adjacent motif (PAM) in genome, thereby realizing site-specific editing. And thus, it has become a main gene editing tool because of its simple operation, wide application range and high throughput.

[0005] Sequence-specific nuclease can produce DNA double-strand breaks at specific sites in the genome. These DNA double-strand breaks can be repaired into a variety of different repair types, which are mainly base insertions or deletions. For example, the two most common types of CRISPR/Cas9 editing events are inserting a base at the break 40 or deleting a base at the break (Shen et al. 2018. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature. DOI: 10.1038/s41586-018-0686-x). The insertion or deletion of bases in coding region will cause frameshift mutations, leading to loss of gene function. Therefore, the main purpose of the above gene editing tools is still to perform gene knockout.

[0006] It has been considered all the time that the use of sequence-specific nuclease alone cannot achieve mutations 45 in type of base substitution. To this end, the prior art proposes three solutions: 1) adding an exogenous DNA fragment as a repair template to initiate a homologous recombination repair pathway; 2) fusing deaminase with Cas9 to sequentially develop single-base editing tools for C to T and A to G; 3) fusing reverse transcriptase with Cas9, using pegRNA to guide the synthesis and substitution of small DNA strand. However, the editing efficiency of these three solutions is significantly lower than the efficiency of gene knockout, and the simultaneously introduced exogenous DNA fragment 50 and reverse transcriptase may readily cause concerns about biological safety. The off-target effect of single-base editing also restricts its potential application in cell therapy. Especially for long-term plant breeding projects, how to improve base substitution efficiency at target site while reducing the regulatory authorities' concerns about biosafety is a problem that needs to be resolved in the application of gene editing technology.

[0007] In summary, there are urgent technical needs in the fields of cell therapy and biological breeding for site-specific 55 base substitution by using only targeted knockout of sequence-specific nuclease without introducing a foreign DNA fragment, especially through the non-transgenic transient editing system to efficiently complete site-specific base substitution editing.

**Summary of Invention**

[0008] The invention provides a method for generating a site-specific mutation in an organism only by generating double-strand breaks on a genome and without providing an artificial DNA template, and use of the method.

5 [0009] The technical solutions adopted by the present invention are as follows:

A method for generating a new mutation in an organism, which comprises the following steps: sequentially generating two or more DNA breaks at a specific site in a genome of the organism and spontaneously repairing them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair.

10 [0010] In a specific embodiment, the "DNA break" is achieved by delivering a nuclease with targeting property into a cell of an organism to contact with a specific site of genomic DNA.

[0011] In a specific embodiment, the "nuclease with targeting property" is a ZFN, TALEN or CRISPR/Cas system.

[0012] In a specific embodiment, the "sequentially generating two or more DNA breaks at a specific site" refers to that based on a new sequence generated from a previous DNA break repair event caused by ZFN or TALEN editing, a new ZFN or TALEN protein is designed to cut the site again.

15 [0013] In another specific embodiment, "sequentially generating two or more DNA breaks at a specific site" refers to that based on a new sequence generated from a previous DNA break repair event caused by a CRISPR/Cas system, a new target RNA is designed to cut the site again. For example, a second cutting is performed at the site again by designing a new target RNA on the basis of a new sequence generated from a first break repair event of Cas9 editing. In a similar way, a third cutting is performed at the site by designing a new target RNA on the basis of a new sequence 20 generated from a second break repair event and so on, as shown in Figure 1.

[0014] In a specific embodiment, the "two or more DNA breaks" are generated by sequentially delivering different targeted nucleases into recipient cells of different generations, wherein a mutant cell that has completed the previous editing is used as a recipient to receive the delivery of the targeted nuclease for the later editing, thereby performing second editing to generate site-specific mutation. This method is preferably used for ZFN and TALEN editing systems.

25 [0015] In another specific embodiment, the "two or more DNA breaks" are generated by delivering different targeted nucleases for different targets into a same recipient cell. This method is preferably used for CRISPR/Cas editing system.

[0016] In a specific embodiment, the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.

30 [0017] In another specific embodiment, the "two or more DNA breaks" are generated when RNP complexes formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences. For example, the PAM sequence recognized by Cas9 from *Streptococcus pyogenes* is "NGG" or "NAG" (Jinek et al., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity", Science 2012, 337:816- 821), the PAM sequence recognized by Cas9 of *Staphylococcus aureus* is "NNGRRT" or "NNGRR(N)", the PAM sequence recognized by *Neisseria meningitidis* Cas9 is NNNNGATT, and the PAM sequence recognized by *Streptococcus thermophilus* Cas9 is NNAGAAW. In this way, the editable window of a DNA molecule is larger.

[0018] In a specific embodiment, the targeted nuclease is any CRISPR/Cas nuclease capable of achieving genome editing.

40 [0019] In a specific embodiment, the targeted nuclease is in a form of DNA.

[0020] In another specific embodiment, the targeted nuclease is in a form of mRNA or protein instead of DNA. The protein form is preferred.

45 [0021] In a specific embodiment, the method for delivering targeted nucleases into a cell is selected from, but not limited to: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) a microinjection; 5) a gene gun bombardment; 6) an *Agrobacterium*-mediated transformation method.

50 [0022] In the method, a new target is designed based on a new sequence generated from the previous DNA break repair, and thus, mutations can be sequentially formed for many times at a specific site in the genome, thereby exponentially enriching types of repair events after DNA breaks, and generating new types of base substitution, deletion and insertion mutations that cannot be obtained by a single gene editing, so that the method is suitable to be used as a tool to create new mutations. This method can be briefly described as a method of programmed sequential cutting/editing or successive cutting/editing.

55 [0023] In a specific embodiment, a new target is designed based on a new specific sequence predicted to be generated from previous break repair at a specific site of organism genome, sequential editing is then performed, and thus a final possible mutation at the site can be designed in advance to achieve an expected editing.

[0024] In another specific embodiment, a new target is designed based on a new sequence predicted to be generated from previous break repair at a specific site of organism genome, sequential editing is then performed, and in addition to an expected editing event, a variety of different mutations can be generated eventually at the site, so that the method

can be used as a tool to create various different mutations.

[0025] In another aspect, the present invention further provides a method for generating a new mutation in an organism, which comprises the following step: sequentially generating two or more DNA breaks at a specific site in a gene at the level of genome or chromosome of the organism, thereby achieving a precise base substitution, deletion or insertion.

5 [0026] In a specific embodiment, the "sequentially generating two or more DNA breaks at a specific site" refers to that a new target RNA designed based on a new sequence generated from a previous break repair event, and a cutting is performed again at the same site.

[0027] In a specific embodiment, the "DNA break" is achieved by a nuclease with targeting property.

10 [0028] The present invention further provides a new mutation obtained by the aforementioned method.

[0029] The present invention further provides a protein or biologically active fragment thereof that has the aforementioned new mutation.

[0030] The present invention further provides a nucleic acid, which comprises a nucleic acid sequence or complementary sequence thereof that encodes the protein or biologically active fragment thereof.

15 [0031] The present invention further provides a nucleic acid, which comprises:

(a) a nucleotide sequence encoding a target RNA, wherein

the target RNA comprises at least two target RNAs, in which a first target RNA targets a DNA to cause a break in the DNA, and a latter target RNA targets a sequence generated from a previous break repair event and generates a break again.

20 [0032] In a specific embodiment, the nucleic acid further comprises (b) a nucleotide sequence encoding a Cas polypeptide.

[0033] In a specific embodiment, the target RNA is a sgRNA or gRNA.

[0034] In a specific embodiment, the Cas polypeptide and the target RNA are present in an in vitro cell or an ex vivo cell.

25 [0035] The present invention further provides a recombinant expression vector, which comprises the aforementioned nucleic acid and a promoter operably linked thereto.

[0036] The present invention further provides an expression cassette, which comprises the aforementioned nucleic acid.

[0037] The present invention further provides a host cell, which comprises the aforementioned expression cassette.

30 [0038] The present invention further provides an organism that is regenerated by using the aforementioned host cell.

[0039] The present invention further provides a method for lysing a target DNA, which comprises contacting the target DNA with a complex, wherein the complex comprises:

35 (a) a Cas polypeptide; and

(b) at least two target RNAs, in which a first targets RNA targets the DNA to cause a break in the DNA, and a latter target RNA targets a sequence generated from a previous break repair event and generates a break again.

40 [0040] In a specific embodiment, the target RNA is a sgRNA or gRNA.

[0041] In a specific embodiment, the target DNA is present in a bacterial cell, eukaryotic cell, plant cell or animal cell.

[0042] In a specific embodiment, the target DNA is a chromosomal DNA.

[0043] In a specific embodiment, the Cas polypeptide and the target RNA are present in an in vitro cell or an ex vivo cell.

45 [0044] In a specific embodiment, the contacting comprises introducing the following into a cell: (a) the Cas polypeptide or a polynucleotide encoding the Cas polypeptide, and (b) the target RNA or a DNA polynucleotide encoding the target RNA.

[0045] The present invention further provides a composition, which comprises:

50 (a) a Cas polypeptide, or a polynucleotide encoding the Cas polypeptide; and

(b) at least two target RNAs, or DNA polynucleotides encoding the target RNAs, wherein a first RNA targets a DNA to cause a break in the DNA, and a latter target RNA targets a sequence generated from a previous break repair event and generates a break again.

[0046] In a specific embodiment, the target RNA is a sgRNA or gRNA.

[0047] In a specific embodiment, the Cas polypeptide and the target RNA are present in an in vitro cell or an ex vivo cell.

55 [0048] The invention further provides a use of the composition in manufacture of a medicament for treatment of a disease.

[0049] The disease that can be treated with the composition of the present invention includes, but is not limited to, a disease caused by single gene mutation, such as genetic tyrosinemia type 1, phenylketonuria, progeria, sickle cell

disease, etc. A spontaneous cell repair is induced by delivering into a cell the Cas protein and the crRNA or sgRNA composition that is expected to repair a pathogenic mutation site to produce a normal functional protein, and thus a therapeutic effect is obtained.

[0050] The present invention further provides a kit, which comprises:

- 5 (a) a Cas polypeptide, or a nucleic acid comprising a nucleotide sequence encoding the Cas polypeptide; and
- (b) at least two target RNAs, or nucleic acids comprising nucleotide sequences encoding the target RNAs, wherein a first target RNA targets a DNA to cause a break in the DNA, and a latter target RNA targets a sequence generated from a previous break repair event and generates a break again; wherein (a) and (b) are in a same or separate containers.

[0051] In a specific embodiment, the target RNA is a sgRNA or gRNA.

[0052] In a specific embodiment, the target RNAs in (b) are in a same or separate containers.

15 [0053] The present invention further provides a method for screening editing events independent of exogenous transgenic markers, comprising the following steps:

- 20 1) two or more DNA breaks are sequentially generated in sequence at a specific site of a first target gene of a recipient cell and spontaneously repaired respectively, wherein a later DNA break is generated based on a new sequence generated from a repair of a previous DNA break;
- 25 2) certain editing events are generated after the specific site of the first target gene is sequentially cut and repaired, which can confer a mutant cell with a resistance to a certain selection pressure to produce a phenotypic selectable trait, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete organism that contains such editing events is isolated;
- 30 3) optionally, in addition to the first target gene, a targeted nuclease for at least one second target gene is used to edit another target site at the same time, and an editing event of the second target gene is enriched and screened synchronously through the screening of the selectable trait generated by mutations of the first target gene, and a cell, tissue, organ or complete organism that simultaneously contains the editing events of the first target gene and of the at least one second target gene is isolated.

[0054] In a specific embodiment, the "first target gene" is a gene locus encoding at least one phenotypic selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

35 [0055] In a specific embodiment, the "specific site of a first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs, which is capable of conferring the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.

[0056] In a specific embodiment, the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.

[0057] In a specific embodiment, the "certain selection pressure" may be an environmental pressure or a pressure resulted from an added compound; for example, the environmental pressure is high temperature, low temperature or hypoxia and the like; the pressure resulted from an added compound may be a pressure resulted from a salt ion concentration, antibiotic, cytotoxin, herbicide, etc.

45 [0058] In a specific embodiment, the "DNA break" is achieved by delivering a nuclease with a targeting property into a cell of an organism to contact with a specific site of genomic DNA.

[0059] In a specific embodiment, the "nuclease with a targeting property" is any CRISPR/Cas nuclease capable of performing genome editing.

50 [0060] In a specific embodiment, the feature, "two or more DNA breaks are sequentially generated in sequence at a specific site", refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.

[0061] In a specific embodiment, the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.

55 [0062] In another specific embodiment, the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences. In this way, the editable window of a DNA molecule is larger.

[0063] In a specific embodiment, the "second target gene" refers to another gene that is different in coding from the

first target gene.

[0064] In a specific embodiment, the "targeted nuclease for at least one second target gene" and the CRISPR/Cas nuclease used for generating DNA break at a specific site of the first target gene are the same.

[0065] In another specific embodiment, the "targeted nuclease for at least one second target gene" and the CRISPR/Cas nuclease used for generating DNA break at a specific site of the first target gene are different. In this way, there are more selectable editing sites on the second target gene.

[0066] In a specific embodiment, the targeted nuclease is in a form of DNA.

[0067] In another specific embodiment, the targeted nuclease is in a form of mRNA or protein instead of DNA. The protein form is preferred.

[0068] In a specific embodiment, the method for delivering targeted nuclease into cell is selected from, but not limited to: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) a microinjection; 5) a gene gun bombardment; or 6) an *Agrobacterium*-mediated transformation method.

[0069] The present invention further provides a method for non-transgenic transient editing of an organism genome, comprising the following steps:

1) a combination of at least two crRNA fragments or a combination of at least two sgRNA fragments is designed and synthesized for a specific site of a first target gene of a recipient cell, wherein the crRNA combination combined with tracrRNA or the sgRNA combination alone is capable of guiding a corresponding Cas protein to sequentially generate two or more DNA breaks at the specific site in the first target gene of the recipient cell and to spontaneously repair them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair;

2) an appropriate amount of CRISPR/Cas protein or corresponding mRNA thereof is mixed with the combination of crRNA fragments and tracrRNA fragment or with the combination of sgRNA fragments alone as above designed and synthesized in advance that is capable of guiding a site-specific editing of the first target gene to generate endogenous selection markers, optionally, at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes is further added, and incubation is carried out in vitro to form an RNP complex;

3) the above RNP complex is delivered into the recipient cell and contacts with the specific site of genomic DNA to achieve gene editing;

4) according to a phenotypic selectable trait generated by a site-specific editing of the first target gene by the RNP complex, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete organism that contains the editing event is isolated, and optionally, a cell, tissue, organ or complete organism that simultaneously contains the editing events of the first target gene and of the at least one of a second, third or more target genes is isolated.

[0070] In a specific embodiment, the "first target gene" is a gene locus encoding at least one phenotypic selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

[0071] In a specific embodiment, the "specific site of the first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs, which is capable of conferring the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.

[0072] In a specific embodiment, the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.

[0073] In a specific embodiment, the "certain selection pressure" may be an environmental pressure or a pressure resulted from an added compound; for example, the environmental pressure is high temperature, low temperature or hypoxia and the like; the pressure resulted from an added compound may be a pressure resulted from a salt ion concentration, antibiotic, cytotoxin or herbicide, and the like.

[0074] In a specific embodiment, the CRISPR/Cas protein is any CRISPR/Cas nuclease capable of performing genome editing.

[0075] In a specific embodiment, the feature "to sequentially generate two or more DNA breaks at the specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.

[0076] In a specific embodiment, the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target se-

quences.

[0077] In another specific embodiment, the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences. In this way, the editable window of a DNA molecule is larger.

[0078] In a specific embodiment, the "second, third or more target genes" refer to other genes that are different in coding from the first target gene.

[0079] In a specific embodiment, the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" shares the same Cas protein with the crRNA or sgRNA targeting the first target gene.

[0080] In another specific embodiment, the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" and the crRNA or sgRNA targeting the first target gene use Cas proteins that recognize different PAM sequences. In this way, there are more selectable editing sites on the second target gene.

[0081] In a specific embodiment, the method for delivering the RNP complex into cells is selected from, but not limited to: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) a microinjection; 5) a gene gun bombardment; and so on.

[0082] The present invention further provides a method for non-transgenic transient editing of a plant genome, comprising the following steps:

20 1) a combination of at least two crRNA fragments or a combination of at least two sgRNA fragments is designed and synthesized for a specific site of a first target gene of a recipient plant cell or tissue, the crRNA combination combined with tracrRNA or the sgRNA combination alone is capable of guiding a corresponding Cas protein to sequentially generate two or more DNA breaks at the specific site in the first target gene of the recipient cell and to spontaneously repair them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair;

30 2) an appropriate amount of CRISPR/Cas protein or corresponding mRNA thereof is mixed with the combination of crRNA fragments and tracrRNA fragment or with the combination of sgRNA fragments alone as above designed and synthesized in advance that is capable of guiding the site-specific editing of the first target gene to generate endogenous selection markers, optionally, at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes is further added, and incubation is carried out in vitro to form an RNP complex;

35 3) the above RNP complex is delivered into the recipient plant cell or tissue and contacts with the specific site of genomic DNA to achieve gene editing;

40 4) according to a phenotypic selectable trait generated by a site-specific editing of the first target gene by the RNP complex, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete plant that contains the editing event is isolated, and optionally, a cell, tissue, organ or complete plant that simultaneously contains the editing events of the first target gene and of the at least one of second, third or more target genes is isolated.

[0083] In a specific embodiment, the "first target gene" is a gene locus encoding at least one phenotypic selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

[0084] In a specific embodiment, the "specific site of the first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs at the site, which can confer the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.

[0085] In a specific embodiment, the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.

[0086] In a specific embodiment, the "certain selection pressure" may be an environmental pressure or a pressure resulted from an added compound; for example, the environmental pressure is preferably high temperature, low temperature or hypoxia and the like; the pressure resulted from an added compound may be a pressure resulted from a salt ion concentration, antibiotic, cytotoxin, herbicide, etc.

[0087] In a specific embodiment, the "recipient plant cell or tissue" is any cell or tissue that can serve as a recipient for transient expression and can be regenerated into a complete plant through tissue culture. Specifically, the cell is a protoplast cell or a suspension cell; the tissue is preferably a callus, immature embryo, mature embryo, leaf, shoot tip, young spike, hypocotyl, etc.

[0088] In a specific embodiment, the CRISPR/Cas protein is any CRISPR/Cas nuclease capable of performing genome editing.

[0089] In a specific embodiment, the feature "to sequentially generate two or more DNA breaks at the specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.

[0090] In a specific embodiment, the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.

[0091] In another specific embodiment, the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences. In this way, the editable window of a DNA molecule is larger.

[0092] In a specific embodiment, the "second, third or more target genes" refer to other genes that are different in coding from the first target gene.

[0093] In a specific embodiment, the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" shares the same Cas protein with the crRNA or sgRNA targeting the first target gene.

[0094] In another specific embodiment, the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" and the crRNA or sgRNA targeting the first target gene use Cas proteins that recognize different PAM sequences. In this way, there are more selectable editing sites on the second target gene.

[0095] In a specific embodiment, the method for delivering the RNP complex into plant cells is selected from, but not limited to: 1) a PEG-mediated protoplast transformation method; 2) a microinjection; 3) a gene gun bombardment; 4) a silicon carbide fiber-mediated method; 5) a vacuum infiltration method, or any other transient introduction method. The gene gun bombardment is preferred.

[0096] In a specific embodiment, the "first target gene" is at least one endogenous gene that encodes at least one phenotypic selectable trait selected from herbicide resistance/tolerance, wherein the herbicide resistance/tolerance is selected from the group consisting of resistance/tolerance to EPSPS inhibitor (including glyphosate); resistance/tolerance to glutamine synthesis inhibitor (including glufosinate); resistance/tolerance to ALS or AHAS inhibitor (including imidazoline or sulfonylurea); resistance/tolerance to ACCase inhibitor (including aryloxyphenoxypropionic acid (FOP)); resistance/tolerance to carotenoid biosynthesis inhibitor, including carotenoid biosynthesis inhibitors of phytoene desaturase (PDS) step, 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors or other carotenoid biosynthesis target inhibitors; resistance/tolerance to cellulose inhibitor; resistance/tolerance to lipid synthesis inhibitor; resistance/tolerance to long-chain fatty acid inhibitor; resistance/tolerance to microtubule assembly inhibitor; resistance/tolerance to photosystem I electron shunting agent; resistance/tolerance to photosystem II inhibitor (including carbamates, triazines and triazones); resistance/tolerance to PPO inhibitor; and resistance/tolerance to synthetic growth hormone (including dicamba, 2,4-D (i.e., 2,4-dichlorophenoxyacetic acid)). Wherein, the first target gene is selected from PsbA, ALS, EPSPS, ACCase, PPO, HPPD, PDS, GS, DOXPS, TIR1, AFB5, and some types of mutations generated after sequential cuttings and repairs at specific sites of these herbicide target genes may confer the recipient plant cells with resistance/tolerance to the corresponding herbicides.

[0097] In a specific embodiment, the "first target gene" is ALS, and the "specific site of gene" refers to site A122, P197, R198, D204, A205, D376, R377, W574, S653 or G654 in an *Arabidopsis* AtALS protein amino acid sequence (e.g., as shown in SEQ ID NO:1), and amino acid sites in an ALS protein of another plant which correspond to the above-mentioned amino acid sites by using the AtALS amino acid sequence as reference standard. The crRNA or sgRNA targets a target sequence comprising a sequence encoding an AtALS protein amino acid sequence site selected from the group consisting of A122, P197, R198, D204, A205, D376, R377, W574, S653, G654 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an ALS protein of another plant which corresponds to the above-mentioned amino acid sites, and any combination thereof, by using the AtALS amino acid sequence as reference standard. The ALS W574 site is preferred. The selection pressure is preferably a treatment with pyroxasulfuron or nicosulfuron.

[0098] In a specific embodiment, the "first target gene" is ACCase, and the "specific site of gene" refers to site 11781, E1874, N1878, W1999, W2027, 12041, D2078, C2088 or G2096 in an *Alopecurus myosuroides* AmACCase protein amino acid sequence (e.g., as shown in SEQ ID NO: 3, and the gene sequence is as shown in SEQ ID NO: 4), and amino acid sites in an ACCase protein of another monocotyledonous plant which correspond to the above-mentioned amino acid sites by using the AmACCase amino acid sequence as reference standard. The crRNA or sgRNA targets a target sequence comprising a sequence encoding an AmACCase amino acid sequence site selected from the group consisting of 11781, E1874, N1878, W1999, W2027, 12041, D2078, C2088, G2096 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an ACCase protein of another monocotyledonous plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the AmACCase

amino acid sequence as reference standard. ACCase W2027 site is preferred. The selection pressure is preferably a treatment with quizalofop-p-ethyl.

**[0099]** In a specific embodiment, the "first target gene" is HPPD, and the "specific site of gene" refers to site H141, L276, P277, N338, G342, R346, D370, P386, K418 or G419 in an *Oryza sativa* OsHPPD protein amino acid sequence (as shown in SEQ ID NO: 5, and the genome sequence is as shown in SEQ ID NO: 6), and amino acid sites in an HPPD protein of another plant which correspond to the above-mentioned amino acid sites by using the OsHPPD amino acid sequence as reference standard. The crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsHPPD amino acid sequence site selected from the group consisting of H141, L276, P277, N338, G342, R346, D370, P386, K418, G419 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an HPPD protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the OsHPPD amino acid sequence as reference standard. The selection pressure is preferably a treatment with biscarfentrazone.

**[0100]** In a specific embodiment, the "first target gene" is PPO, and the "specific site of gene" refers to site S128, V217, S223, V364, K373, L423, Y425 or W470 in an *Oryza sativa* OsPPO1 protein amino acid sequence (as shown in SEQ ID NO: 7, and the genome sequence is as shown in SEQ ID NO: 8), and amino acid sites in a PPO protein of another plant which correspond to the above-mentioned amino acid sites by using the amino acid sequence of OsPPO1 as reference standard. The crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsPPO1 amino acid sequence site selected from the group consisting of S128, V217, S223, V364, K373, L423, Y425, W470 or any combination thereof, and a target sequence comprising a sequence of the above-mentioned amino acid sites corresponding to a PPO protein of another plant and any combination thereof using the OsPPO1 amino acid sequence as reference standard. The selection pressure is preferably a treatment with saflufenacil.

**[0101]** In a specific embodiment, the "first target gene" is TIR1, and the "specific site of gene" refers to site F93, F357, C413 or S448 in an *Oryza sativa* OsTIR1 protein amino acid sequence (as shown in SEQ ID NO: 9, and the genome sequence is as shown in SEQ ID NO: 10), and amino acid sites in a TIR1 protein of another plant which correspond to the above-mentioned amino acid sites by using the OsTIR1 amino acid sequence as reference standard. The crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsTIR1 amino acid sequence site selected from the group consisting of F93, F357, C413, S448 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in a TIR1 protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the OsTIR1 amino acid sequence as reference standard. The selection pressure is preferably 2,4-D treatment.

**[0102]** The present invention further provides a non-transgenic transient editing system using the aforementioned method.

**[0103]** The present invention additionally provides a use of the aforementioned non-transgenic transient editing system as a selection marker.

**[0104]** The present invention additionally provides a use of the aforementioned non-transgenic transient editing system in treatment of a disease.

**[0105]** The present invention additionally provides a use of the aforementioned non-transgenic transient editing system in biological breeding.

**[0106]** The present invention additionally provides a genetically modified plant obtained by the aforementioned method, the genome of which contains an editing event of a first target gene, and the genetically modified plant is obtained in a non-transgenic manner.

**[0107]** The present invention additionally provides a genetically modified plant obtained by the aforementioned method, the genome of which contains an editing event of a first target gene, and further contains at least one second target gene editing event, and the genetically modified plant is obtained in a non-transgenic manner.

**[0108]** The present invention additionally provides a genetically modified plant obtained by the aforementioned method, the genome of which contains at least one second target gene editing event, and the genetically modified plant is obtained in a non-transgenic manner, wherein the first target gene editing event has been removed by genetic separation.

**[0109]** The present invention further provides a genome of the genetically modified plant obtained by the aforementioned method, the genome comprising: 1) an editing event of a first target gene; 2) an editing event of the first target gene and an editing event of at least one second target gene; or 3) at least one second target gene editing event, wherein the editing event of the first target gene has been removed by genetic separation; wherein the genetically modified plant is obtained in a non-transgenic manner.

**[0110]** Another aspect of the present invention provides a new plant gene mutation obtained by the aforementioned method.

**[0111]** The present invention also provides a new mutation generated in a plant, which comprises one or a combination of two or more of the following types:

substitution of aspartic acid at a site corresponding to *Arabidopsis* ALS376 with any other amino acid, substitution of tryptophan at a site corresponding to *Arabidopsis* ALS574 with any other amino acid, substitution of serine at a site

corresponding to *Arabidopsis* ALS653 with any other amino acid, or substitution of serine at a site corresponding to *Arabidopsis* ALS654 with any other amino acid; or substitution of tryptophan at a site corresponding to *Alopecurus myosuroides* ACCase2027 with any other amino acid.

[0112] In a specific embodiment, wherein the aspartic acid at a site corresponding to *Arabidopsis* ALS376 is substituted by glutamic acid (D376E), the tryptophan at a site corresponding to *Arabidopsis* ALS574 is substituted by leucine or methionine (W574L or W574M), the serine at a site corresponding to *Arabidopsis* ALS653 is substituted by asparagine or arginine (S653N or S653R), or the glycine at a site corresponding to *Arabidopsis* ALS654 is substituted by aspartic acid (G654D), wherein the sites of amino acids are mentioned by using the sites of corresponding amino acids in *Arabidopsis thaliana* as reference; or, the tryptophan at a site corresponding to *Alopecurus myosuroides* ACCase2027 is substituted by leucine or cysteine (W2027L or W2027C), wherein the site of amino acid is mentioned by using the site of corresponding amino acid in *Alopecurus myosuroides* as reference.

[0113] In another specific embodiment, the mutation type is S653R/G654D, wherein the sites of amino acids are mentioned by using the sites of corresponding amino acids in *Arabidopsis thaliana* as reference.

[0114] In a specific embodiment, the aspartic acid at site 350 of *Oryza sativa* ALS is substituted by any other amino acid, the tryptophan at site 548 of *Oryza sativa* ALS is substituted by any other amino acid, or the tryptophan at site 561 of *Solanum tuberosum L.* ALS2 is substituted by any other amino acid; or, the tryptophan at site 2038 of *Oryza sativa* ACCase2 is substituted by any other amino acid.

[0115] In another specific embodiment, the aspartic acid at site 350 of *Oryza sativa* ALS is substituted by glutamic acid (D350E), the tryptophan at site 548 of *Oryza sativa* ALS is substituted by leucine or methionine (W548L or W548M), or the tryptophan at site 561 of *Solanum tuberosum L.* ALS2 is substituted by leucine or methionine (W561L or W561M); or, the tryptophan at site 2038 of *Oryza sativa* ACCase2 is substituted by leucine or cysteine (W2038L or W2038C), wherein the amino acid sequence of the *Oryza sativa* ALS protein is shown in SEQ ID NO: 11, the amino acid sequence of the *Solanum tuberosum L.* StALS2 protein is shown in SEQ ID NO: 19, and the amino acid sequence of the *Oryza sativa* ACCase2 protein is shown in SEQ ID NO: 13.

[0116] The present invention additionally provides a protein or biologically active fragment thereof that has the aforementioned new mutation.

[0117] The present invention also provides a nucleic acid, which comprises a nucleic acid sequence or complementary sequence thereof that encodes the protein or biologically active fragment thereof.

[0118] The present invention additionally provides a recombinant expression vector, which comprises the nucleic acid and a promoter operably linked thereto.

[0119] The present invention further provides an expression cassette, which comprises the nucleic acid.

[0120] The present invention further provides a plant cell, which comprises the expression cassette.

[0121] The present invention further provides a plant regenerated by using the plant cell.

[0122] Another aspect of the present invention provides a method for producing a plant with improved resistance or tolerance to herbicides, which comprises regenerating the plant cell into a plant.

[0123] Another aspect of the present invention provides a method for controlling weeds in a plant cultivation site, wherein the plant includes the aforementioned plant or a plant produced by the aforementioned method, wherein the method comprises applying to the cultivation site one or more herbicides in an effective amount to control the weeds.

[0124] Another aspect of the present invention also provides a use of the new mutation, the protein or biologically active fragment thereof, the nucleic acid, the recombinant expression vector or the expression cassette in improving resistance or tolerance of a plant cell, a plant tissue, a plant part or a plant to herbicides.

[0125] The present invention has the following excellent technical effects:

Based on sequences generated from new repair event generated by sequential editing, new targets can be designed, which can sequentially form mutations for many times at a specific site in the genome, thereby exponentially enriching the types of repair events after DNA breaks, and realizing new types of base substitution, deletion and insertion mutations that cannot be obtained by a single gene editing. That is, the programmed sequential cutting/editing scheme adopted by the present invention, which uses the sequence generated from a previous gene editing repair as the later gene editing target, can endow CRISPR/Cas with new functions of single-base editing and site-precise deletion and insertion through simple knockout.

[0126] The invention can realize the screening of gene editing events in the absence of exogenous markers, further realize the non-transgenic gene editing and can effectively screen editing events, and can greatly reduce the biological safety concerns of the method in cell therapy and biological breeding.

[0127] In particular, the plant non-transgenic transient editing method provided by the present invention only involves Cas protein and artificially synthesized small fragments of gRNA or sgRNA, without the participation of exogenous DNA in the whole process, and produces endogenous resistance selection markers by editing the first target gene via continuous targeting, so that the editing event can be effectively screened, actually without genetic modification operations involved, and thus the method is equivalent to chemical mutagenesis or radiation-induced breeding, also does not require continuous multiple generations of separation and detection of exogenous transgenic components, thereby shortening

the breeding cycle, ensuring the biological safety, saving supervision and approval cost, and providing great application prospects in precise breeding of plants.

#### Detailed description of invention

5

**[0128]** In the present invention, unless otherwise specified, the scientific and technical terms used herein have the meanings commonly understood by those skilled in the art. In addition, protein and nucleic acid chemistry, molecular biology, cell and tissue culture, microbiology, immunology related terms and laboratory procedures used herein are all terms and routine procedures widely used in the corresponding fields. At the same time, in order to better understand the present invention, definitions and explanations of related terms are provided below.

10

**[0129]** The term "genome" as used herein refers to all complements of genetic material (genes and non-coding sequences) present in each cell or virus or organelle of an organism, and/or complete genome inherited from a parent as a unit (haploid).

15

**[0130]** The term "gene editing" refers to strategies and techniques for targeted specific modification of any genetic information or genome of living organisms. Therefore, the term includes editing of gene coding regions, but also includes editing of regions other than gene coding regions of the genome. It also includes editing or modifying other genetic information of nuclei (if present) and cells.

20

**[0131]** The term "CRISPR/Cas nuclease" may be a CRISPR-based nuclease or a nucleic acid sequence encoding the same, including but not limited to: 1) Cas9, including SpCas9, ScCas9, SaCas9, xCas9, VRER-Cas9, EQR-Cas9, SpG-Cas9, SpRY-Cas9, SpCas9-NG, NG-Cas9, NGA-Cas9 (VQR), etc.; 2) Cas12, including LbCpf1, FnCpf1, AsCpf1, MAD7, etc., or any variant or derivative of the aforementioned CRISPR-based nuclease; preferably, wherein the at least one CRISPR-based nuclease comprises a mutation compared to the corresponding wild-type sequence, so that the obtained CRISPR-based nuclease recognizes a different PAM sequence. As used herein, "CRISPR-based nuclease" is any nuclease that has been identified in a naturally occurring CRISPR system, which is subsequently isolated from its natural background, and has preferably been modified or combined into a recombinant construct of interest, suitable as a tool for targeted genome engineering. As long as the original wild-type CRISPR-based nuclease provides DNA recognition, i.e., binding properties, any CRISPR-based nuclease can be used and optionally reprogrammed or otherwise mutated so as to be suitable for various embodiments of the invention.

25

**[0132]** The term "CRISPR" refers to a sequence-specific genetic manipulation technique that relies on clustered regularly interspaced short palindromic repeats, which is different from RNA interference that regulates gene expression at the transcriptional level.

30

**[0133]** "Cas9 nuclease" and "Cas9" are used interchangeably herein, and refer to RNA-guided nuclease comprising Cas9 protein or fragment thereof (for example, a protein containing the active DNA cleavage domain of Cas9 and/or the gRNA binding domain of Cas9). Cas9 is a component of the CRISPR/Cas (clustered regularly interspaced short palindromic repeats and associated systems) genome editing system. It can target and cut DNA target sequences under the guidance of guide RNA to form DNA double-strand breaks (DSB).

35

**[0134]** "Cas protein" or "Cas polypeptide" refers to a polypeptide encoded by Cas (CRISPR-associated) gene. Cas protein includes Cas endonuclease. Cas protein can be a bacterial or archaeal protein. For example, the types I to III CRISPR Cas proteins herein generally originate from prokaryotes; the type I and type III Cas proteins can be derived from bacteria or archaea species, and the type II Cas protein (i.e., Cas9) can be derived from bacterial species. "Cas proteins" include Cas9 protein, Cpf1 protein, C2c1 protein, C2c2 protein, C2c3 protein, Cas3, Cas3-HD, Cas5, Cas7, Cas8, Cas10, Cas12a, Cas12b, or a combination or complex thereof.

40

**[0135]** "Cas9 variant" or "Cas9 endonuclease variant" refers to a variant of the parent Cas9 endonuclease, wherein when associated with crRNA and tracrRNA or with sgRNA, the Cas9 endonuclease variant retains the abilities of recognizing, binding to all or part of a DNA target sequence and optionally unwinding all or part of a DNA target sequence, nicking all or part of a DNA target sequence, or cutting all or part of a DNA target sequence. The Cas9 endonuclease variants include the Cas9 endonuclease variants described herein, wherein the Cas9 endonuclease variants are different from the parent Cas9 endonuclease in the following manner: the Cas9 endonuclease variants (when complexed with gRNA to form a polynucleotide-directed endonuclease complex capable of modifying a target site) have at least one improved property, such as, but not limited to, increased transformation efficiency, increased DNA editing efficiency, decreased off-target cutting, or any combination thereof, as compared to the parent Cas9 endonuclease (complexed with the same gRNA to form a polynucleotide-guided endonuclease complex capable of modifying the same target site).

45

**[0136]** The Cas9 endonuclease variants described herein include variants that can bind to and nick double-stranded DNA target sites when associated with crRNA and tracrRNA or with sgRNA, while the parent Cas endonuclease can bind to the target site and result in double strand break (cleavage) when associated with crRNA and tracrRNA or with sgRNA.

50

**[0137]** "Guide RNA" and "gRNA" are used interchangeably herein, and refer to a guide RNA sequence used to target a specific gene for correction using CRISPR technology, which usually consists of crRNA and tracrRNA molecules that

are partially complementary to form a complex, wherein crRNA contains a sequence that has sufficient complementarity with the target sequence so as to hybridize with the target sequence and direct the CRISPR complex (Cas9+crRNA+tracrRNA) to specifically bind to the target sequence. However, it is known in the art that a single guide RNA (sgRNA) can be designed, which contains both the properties of crRNA and tracrRNA.

**[0138]** The terms "single guide RNA" and "sgRNA" are used interchangeably herein, and refer to the synthetic fusion of two RNA molecules, which comprises a fusion of a crRNA (CRISPR RNA) of a variable targeting domain (linked to a tracr pairing sequence hybridized to tracrRNA) and a tracrRNA (trans-activating CRISPR RNA). The sgRNA may comprise crRNA or crRNA fragments and tracrRNA or tracrRNA fragments of the type II CRISPR/Cas system that can form a complex with the type II Cas endonuclease, wherein the guide RNA/Cas endonuclease complex can guide the Cas endonuclease to a DNA target site so that the Cas endonuclease can recognize, optionally bind to the DNA target site, and optionally nick the DNA target site or cut (introduce a single-strand or double-strand break) the DNA target site.

**[0139]** In certain embodiments, the guide RNA(s) and Cas9 can be delivered to a cell as a ribonucleoprotein (RNP) complex. RNP is composed of purified Cas9 protein complexed with gRNA, and it is well known in the art that RNP can be effectively delivered to many types of cells, including but not limited to stem cells and immune cells (Addgene, Cambridge, MA, Mirus Bio LLC, Madison, WI).

**[0140]** The protospacer adjacent motif (PAM) herein refers to a short nucleotide sequence adjacent to a (targeted) target sequence (prespacer) recognized by the gRNA/Cas endonuclease system. If the target DNA sequence is not adjacent to an appropriate PAM sequence, the Cas endonuclease may not be able to successfully recognize the target DNA sequence. The sequence and length of PAM herein can be different depending on the Cas protein or Cas protein complex in use. The PAM sequence can be of any length, but is typically in length of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 nucleotides.

**[0141]** As used herein, the term "organism" or "living body" includes animals, plants, fungi, bacteria, and the like.

**[0142]** As used herein, the term "host cell" includes plant cells, animal cells, fungal cells, bacterial cells, and the like.

**[0143]** In the present invention, "animal" includes but is not limited to vertebrates, such as humans, non-human mammals, birds, fish, reptiles, amphibians, etc., as well as invertebrates, such as insects.

**[0144]** In the present invention, the "plant" should be understood to mean any differentiated multicellular organism capable of performing photosynthesis, in particular monocotyledonous or dicotyledonous plants, for example, (1) food crops: Oryza spp., like Oryza sativa, Oryza latifolia, Oryza sativa, Oryza glaberrima; Triticum spp., like Triticum aestivum, T. Turgidumssp. durum; Hordeum spp., like Hordeum vulgare, Hordeum arizonicum; Secale cereale; Avena spp., like Avena sativa, Avena fatua, Avena byzantine, Avena fatua var.sativa, Avena hybrida; Echinochloa spp., like Pennisetum glaucum, Sorghum, Sorghum bicolor, Sorghum vulgare, Tritcale, Zea mays or Maize, Millet, Rice, Foxtail millet, Proso millet, Sorghum bicolor, Panicum, Fagopyrum spp., Panicum miliaceum, Setaria italica, Zizania palustris, Eragrostis tef, Panicum miliaceum, Eleusine coracana; (2) legume crops: Glycine spp. like Glycine max, Soja hispida, Soja max, Vicia spp., Vigna spp., Pisum spp., field bean, Lupinus spp., Vicia, Tamarindus indica, Lens culinaris, Lathyrus spp., Lablab, broad bean, mung bean, red bean, chickpea; (3) oil crops: Arachis hypogaea, Arachis spp, Sesamum spp., Helianthus spp. like Helianthus annuus, Elaeis like Eiaeis guineensis, Elaeis oleifera, soybean, Brassicanapus, Brassica oleracea, Sesamum orientale, Brassica juncea, Oilseed rape, Camellia oleifera, oil palm, olive, castor-oil plant, Brassica napus L., canola; (4) fiber crops: Agave sisalana, Gossypium spp. like Gossypium, Gossypium barbadense, Gossypium hirsutum, Hibiscus cannabinus, Agave sisalana, Musa textilis Nee, Linum usitatissimum, Corchorus capsularis L, Boehmeria nivea (L.), Cannabis sativa, Cannabis sativa; (5) fruit crops: Ziziphus spp., Cucumis spp., Passiflora edulis, Vitis spp., Vaccinium spp., Pyrus communis, Prunus spp., Psidium spp., Punica granatum, Malus spp., Citrullus lanatus, Citrus spp., Ficus carica, Fortunella spp., Fragaria spp., Crataegus spp., Diospyros spp., Eugenia uniflora, Eriobotrya japonica, Dimocarpus longan, Carica papaya, Cocos spp., Averrhoa carambola, Actinidia spp., Prunus amygdalus, Musa spp. (musa acuminate), Persea spp. (Persea Americana), Psidium guajava, Mammea Americana, Mangifera indica, Canarium album (Oleaeuropaea), Caricapapaya, Cocos nucifera, Malpighia emarginata, Manilkara zapota, Ananas comosus, Annona spp., Citrus reticulate (Citrus spp.), Artocarpus spp., Litchi chinensis, Ribes spp., Rubus spp., pear, peach, apricot, plum, red bayberry, lemon, kumquat, durian, orange, strawberry, blueberry, ham melon, muskmelon, date palm, walnut tree, cherry tree; (6) rhizome crops: Manihot spp., Ipomoea batatas, Colocasia esculenta, tuber mustard, Allium cepa (onion), eleocharis tuberosa (water chestnut), Cyperus rotundus, Rhizoma dioscoreae; (7) vegetable crops: Spinacia spp., Phaseolus spp., Lactuca sativa, Momordica spp., Petroselinum crispum, Capsicum spp., Solanum spp. (such as Solanum tuberosum, Solanum integrifolium, Solanum lycopersicum), Lycopersicon spp. (such as Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma spp., Kale, Luffa acutangula, lentil, okra, onion, potato, artichoke, asparagus, broccoli, Brussels sprouts, cabbage, carrot, cauliflower, celery, collard greens, squash, Benincasa hispida, Asparagus officinalis, Apium graveolens, Amaranthus spp., Allium spp., Abelmoschus spp., Cichorium endivia, Cucurbita spp., Coriandrum sativum, B.carinata, Rapbanus sativus, Brassica spp. (such as Brassica napus, Brassica rapa spp., canola, oilseed rape, turnip rape, turnip rape, leaf mustard, cabbage, black mustard, canola (rapeseed), Brussels sprout, Solanaceae (eggplant), Capsicum annum (sweet pepper), cucumber, luffa, Chinese cabbage, rape, cabbage, calabash, Chinese chives, lotus, lotus root, lettuce; (8) flower crops: Tropaeolum minus, Tropaeolum

majus, Canna indica, Opuntia spp., Tagetes spp., Cymbidium (orchid), Crinum asiaticum L., Clivia, Hippeastrum rutilum, Rosa rugosa, Rosa Chinensis, Jasminum sambac, Tulipa gesneriana L., Cerasus sp., Pharbitis nil (L.) Choisy, Calendula officinalis L., Nelumbo sp., Bellis perennis L., Dianthus caryophyllus, Petunia hybrida, Tulipa gesneriana L., Lilium brownie, Prunus mume, Narcissus tazetta L., Jasminum nudiflorum Lindl., Primula malacoides, Daphne odora, Camellia japonica, Michelia alba, Magnolia liliiflora, Viburnum macrocephalum, Clivia miniata, Malus spectabilis, Paeonia suffruticosa, Paeonia lactiflora, Syzygium aromaticum, Rhododendron simsii, Rhododendron hybridum, Michelia figo (Lour.) Spreng., Cercis chinensis, Kerria japonica, Weigela florida, Fructus forsythiae, Jasminum mesnyi, Parochetus communis, Cyclamen persicum Mill., Phalaenopsis hybrid, Dendrobium nobile, Hyacinthus orientalis, Iris tectorum Maxim, Zantedeschia aethiopica, Calendula officinalis, Hippeastrum rutilum, Begonia semperflorens-hybr, Fuchsia hybrida, Begonia maculata Raddi, Geranium, Epipremnum aureum; (9) medicinal crops: Carthamus tinctorius, Mentha spp., Rheum rhabarbarum, Crocus sativus, Lycium chinense, Polygonatum odoratum, Polygonatum Kingianum, Anemarrhena asphodeloides Bunge, Radix ophiopogonis, Fritillaria cirrhosa, Curcuma aromatica, Amomum villosum Lour., Polygonum multiflorum, Rheum officinale, Glycyrrhiza uralensis Fisch, Astragalus membranaceus, Panax ginseng, Panax notoginseng, Acanthopanax gracilistylus, Angelica sinensis, Ligusticum wallichii, Bupleurum sinenses DC., Datura stramonium Linn., 15 Datura metel L., Mentha haplocalyx, Leonurus sibiricus L., Agastache rugosus, Scutellaria baicalensis, Prunella vulgaris L., Pyrethrum carneum, Ginkgo biloba L., Cinchona ledgeriana, Hevea brasiliensis (wild), Medicago sativa Linn, Piper Nigrum L., Radix Isatidis, Atractylodes macrocephala Koidz; (10) raw material crops: Hevea brasiliensis, Ricinus communis, Vernicia fordii, Morus alba L., Hops Humulus lupulus, Betula, Alnus cremastogyne Burk., Rhus verniciflua stokes; (11) pasture crops: Agropyron spp., Trifolium spp., Miscanthus sinensis, Pennisetum sp., Phalaris arundinacea, Panicum 20 virgatum, prairiegrasses, Indiangrass, Big bluestem grass, Phleum pratense, turf, cyperaceae (Kobresia pygmaea, Carex pediformis, Carex humilis), Medicago sativa Linn, Phleum pratense L., Medicago sativa, Melilotus suavcolen, Astragalus sinicus, Crotalaria juncea, Sesbania cannabina, Azolla imbircata, Eichhornia crassipes, Amorpha fruticosa, Lupinus micranthus, Trifolium, Astragalus adsurgens pall, Pistia stratiotes linn, Alternanthera philoxeroides, Lolium; (12) sugar crops: Saccharum spp., Beta vulgaris; (13) beverage crops: Camellia sinensis, Camellia Sinensis, tea, Coffee (Coffea spp.), Theobroma cacao, Humulus lupulus Linn.; (14) lawn plants: Ammophila arenaria, Poa spp. (Poa pratensis (blue-grass)), Agrostis spp. (Agrostis matsumurae, Agrostis palustris), Lolium spp. (Lolium), Festuca spp. (Festuca ovina L.), Zoysia spp. (Zoysia japonica), Cynodon spp. (Cynodon dactylon/bermudagrass), Stenotaphrum secundatum tum (Stenotaphrum secundatum), Paspalum spp., Eremochloa ophiuroides (centipedegrass), Axonopus spp. (carpetweed), Bouteloua dactyloides (buffalograss), Bouteloua var. spp. (Bouteloua gracilis), Digitaria sanguinalis, Cyperus rotundus, Kyllinga brevifolia, Cyperus amuricus, Erigeron canadensis, Hydrocotylesibthorpioides, Kummerowia striata, Euphorbia humifusa, Viola arvensis, Carex rigescens, Carex heterostachya, turf; (15) tree crops: Pinus spp., Salix spp., Acer spp., Hibiscus spp., Eucalyptus spp., Ginkgo biloba, Bambusa sp., Populus spp., Prosopis spp., Quercus spp., Phoenix spp., Fagus spp., Ceiba pentandra, Cinnamomum spp., Corchorus spp., Phragmites australis, Physalis spp., Desmodium spp., Populus, Hedera helix, Populus tomentosa Carr, Viburnum odoratissimum, Ginkgo biloba L., Quercus, Ailanthus altissima, Schima superba, Ilex purpurea, Platanus acerifolia, ligustrum lucidum, Buxus megistophylla Lev., Dahurian larch, Acacia mearnsii, Pinus massoniana, Pinus khasys, Pinus yunnanensis, Pinus finlaysoniana, Pinus tabuliformis, Pinus koraiensis, Juglans nigra, Citrus limon, Platanus acerifolia, Syzygium jambos, Davidia involucrate, Bombax malabarica L., Ceiba pentandra (L.), Bauhinia blakeana, Albizia saman, Albizzia julibrissin, Erythrina corallodendron, Erythrina indica, Magnolia gradiflora, Cycas revolute, Lagerstroemia indica, coniferous, macrophanerophytes, Frutex; (16) nut crops: Bertholletia excelsa, Castanea spp., Corylus spp., Carya spp., Juglans spp., Pistacia vera, Anacardium occidentale, Macadamia (Macadamia integrifolia), Carya illinoensis Koch, Macadamia, Pistachio, Badam, other plants that produce nuts; (17) others: arabidopsis thaliana, Brassica eruciformis, Cenchrus echinatus, Setaria faberi, eleusine indica, Cadaba farinose, algae, Carex elata, ornamental plants, Carissa macrocarpa, Cynara spp., Daucus carota, Dioscorea spp., Erianthus sp., Festuca arundinacea, Hemerocallis fulva, Lotus spp., Luzula sylvatica, Medicago sativa, 30 Melilotus spp., Morus nigra, Nicotiana spp., Olea spp., Ornithopus spp., Pastinaca sativa, Sambucus spp., Sinapis sp., Syzygium spp., Tripsacum dactyloides, Triticosecale rimpau, Viola odorata, and the like.

**[0145]** In a specific embodiment, the plant is selected from rice, corn, wheat, soybean, sunflower, sorghum, rape, alfalfa, cotton, barley, millet, sugarcane, tomato, tobacco, cassava, potato, sweet potato, Chinese cabbage, cabbage, cucumber, Chinese rose, *Scindapsus aureus*, watermelon, melon, strawberry, blueberry, grape, apple, citrus, peach, pear, banana, etc.

**[0146]** As used herein, the term "plant" includes a whole plant and any progeny, cell, tissue or part of plant. The term "plant part" includes any part of a plant, including, for example, but not limited to: seed (including mature seed, immature embryo without seed coat, and immature seed); plant cutting; plant cell; plant cell culture; plant organ (e.g., pollen, embryo, flower, fruit, bud, leaf, root, stem, and related explant). Plant tissue or plant organ can be a seed, callus tissue, or any other plant cell population organized into a structural or functional unit. The plant cell or tissue culture can regenerate a plant that has the physiological and morphological characteristics of the plant from which the cell or tissue is derived, and can regenerate a plant that has substantially the same genotype as the plant. In contrast, some plant cells cannot regenerate plants. The regenerable cells in plant cells or tissue cultures can be embryos, protoplasts,

meristematic cells, callus, pollen, leaves, anthers, roots, root tips, silk, flowers, kernels, spikes, cobs, husks, or stems.

**[0147]** The plant parts comprise harvestable parts and parts that can be used to propagate offspring plants. The plant parts that can be used for propagation include, for example, but not limited to: seeds; fruits; cuttings; seedlings; tubers; and rootstocks. The harvestable parts of plants can be any of useful parts of plants, including, for example, but not limited to: flowers; pollen; seedlings; tubers; leaves; stems; fruits; seeds; and roots.

**[0148]** The plant cells are the structural and physiological units of plants. As used herein, the plant cells include protoplasts and protoplasts with partial cell walls. The plant cells may be in a form of isolated single cells or cell aggregates (e.g., loose callus and cultured cells), and may be part of higher order tissue units (e.g., plant tissues, plant organs, and plants). Therefore, the plant cells can be protoplasts, gamete-producing cells, or cells or collection of cells capable of regenerating a whole plant. Therefore, in the embodiments herein, a seed containing a plurality of plant cells and capable of regenerating into a whole plant is considered as a "plant part".

**[0149]** As used herein, the term "protoplast" refers to a plant cell whose cell wall is completely or partially removed and whose lipid bilayer membrane is exposed. Typically, the protoplast is an isolated plant cell without cell wall, which has the potential to regenerate a cell culture or a whole plant.

**[0150]** The plant "offspring" includes any subsequent generations of the plant.

**[0151]** The term "bacteria" means all prokaryotes, including all organisms in the Kingdom Procyotae. The term "bacteria" includes all microorganisms considered to be bacteria, including Mycoplasma, Chlamydia, Actinomyces, Streptomyce, and Rickettsia. All forms of bacteria are included in this definition, including cocci, bacilli, spirilla, spheroplasts, protoplasts, etc. The term also includes prokaryotes that are Gram-negative or Gram-positive. "Gram-negative" and "Gram-positive" mean a staining pattern using Gram staining methods well known in the art (see, for example, Finegold and Martin, Diagnostic Microbiology, 6th Ed., CV Mosby St. Louis, pp. 13-15[1982]). "Gram-positive bacteria" are bacteria that can retain the original dye used for Gram staining, causing the stained cells to appear dark blue to purple under a microscope. "Gram-negative bacteria" do not retain the original dye used for Gram staining, but can be stained with a counter stain. Therefore, Gram-negative bacteria appear red after the Gram staining reaction.

**[0152]** As used herein, the term "fungi" refers to eukaryotic organisms such as molds and yeasts, including dimorphic fungi.

**[0153]** The terms "herbicide tolerance" and "herbicide resistance" can be used interchangeably, and both refer to herbicide tolerance and herbicide resistance. "Improvement in herbicide tolerance" and "improvement in herbicide resistance" mean that the tolerance or resistance to the herbicide is improved compared to a plant containing wild-type gene.

**[0154]** The term "wild-type" refers to a nucleic acid molecule or protein that can be found in nature.

**[0155]** In the present invention, the term "cultivation site" comprises a site where the plant of the present invention is cultivated, such as soil, and also comprises, for example, plant seeds, plant seedlings and grown plants. The term "weed-controlling effective amount" refers to an amount of herbicide that is sufficient to affect the growth or development of the target weed, for example, to prevent or inhibit the growth or development of the target weed, or to kill the weed. Advantageously, the weed-controlling effective amount does not significantly affect the growth and/or development of the plant seeds, plant seedlings or plants of the present invention. Those skilled in the art can determine such weed-controlling effective amount through routine experiments.

**[0156]** The term "Target DNA" as used herein refers to a DNA polynucleotide comprising a "target site" or "target sequence".

**[0157]** The term "lysis" means the cleavage of the covalent backbone of a DNA molecule. The lysis can be initiated by a variety of methods, including but not limited to enzymatic or chemical hydrolysis of phosphodiester bonds. Both single-strand and double-strand lysis is possible, and double-strand lysis may occur due to two distinct single-strand lysis events. The DNA lysis may result in blunt or staggered ends. In certain embodiments, a complex comprising DNA-targeting RNA and a site-specific modification polypeptide is used for a targeted double-strand DNA lysis.

**[0158]** The term "gene" comprises a nucleic acid fragment expressing a functional molecule (such as, but not limited to, specific protein), including regulatory sequences before (5' non-coding sequences) and after (3' non-coding sequences) a coding sequence.

**[0159]** The DNA sequence that "encodes" a specific RNA is a DNA nucleic acid sequence that can be transcribed into RNA. The DNA polynucleotides can encode a RNA (mRNA) that can be translated into a protein, or the DNA polynucleotides can encode a RNA that cannot be translated into a protein (for example, tRNA, rRNA, or DNA-targeting RNA; which are also known as "non-coding" RNA or "ncRNA").

**[0160]** The terms "polypeptide", "peptide" and "protein" are used interchangeably in the present invention, and refer to a polymer of amino acid residues. The terms are applied to amino acid polymers in which one or more amino acid residues are artificially chemical analogs of corresponding and naturally occurring amino acids, as well as to naturally occurring amino acid polymers. The terms "polypeptide", "peptide", "amino acid sequence" and "protein" may also include their modification forms, including but not limited to glycosylation, lipid linkage, sulfation,  $\gamma$ -carboxylation of glutamic acid residue, hydroxylation and ADP-ribosylation.

**[0161]** The term "biologically active fragment" refers to a fragment that has one or more amino acid residues deleted

from the N and/or C-terminus of a protein while still retaining its functional activity.

[0162] For the terms related to amino acid substitution used in the description, the first letter represents a naturally occurring amino acid at a certain position in a specific sequence, the following number represents the position in the corresponding sequence, and the second letter represents a different amino acid for substituting the naturally occurring amino acid. For example, W574L means that tryptophan at position 574 is substituted by leucine. For double or multiple mutations, each mutation is separated by "/".

[0163] The terms "polynucleotide" and "nucleic acid" are used interchangeably and comprise DNA, RNA or hybrids thereof, which may be double-stranded or single-stranded.

[0164] The terms "nucleotide sequence" and "nucleic acid sequence" both refer to the sequence of bases in DNA or RNA.

[0165] As used in the present invention, "expression cassette", "expression vector" and "expression construct" refer to a vector such as a recombinant vector suitable for expression of a nucleotide sequence of interest in a plant. The term "expression" refers to the production of a functional product. For example, the expression of a nucleotide sequence may refer to the transcription of the nucleotide sequence (such as transcription to generate mRNA or functional RNA) and/or the translation of RNA into a precursor or mature protein.

[0166] The "expression construct" of the present invention can be a linear nucleic acid fragment, a circular plasmid, a viral vector, or, in some embodiments, can be an RNA (such as mRNA) that can be translated.

[0167] The "expression construct" of the present invention may comprise regulatory sequences and nucleotide sequences of interest from different sources, or regulatory sequences and nucleotide sequences of interest from the same source but arranged in a way different from those normally occurring in nature.

[0168] The terms "recombinant expression vector" or "DNA construct" are used interchangeably herein and refer to a DNA molecule comprising a vector and at least one insert. Recombinant expression vectors are usually produced for the purpose of expression and/or propagation of the insert or for the construction of other recombinant nucleotide sequences. The insert may be operably or may be inoperably linked to a promoter sequence and may be operably or may be inoperably linked to a DNA regulatory sequence.

[0169] The terms "regulatory sequence" and "regulatory element" can be used interchangeably and refer to a nucleotide sequence that is located at the upstream (5' non-coding sequence), middle or downstream (3' non-coding sequence) of a coding sequence, and affects the transcription, RNA processing, stability or translation of a related coding sequence. Plant expression regulatory elements refer to nucleotide sequences that can control the transcription, RNA processing or stability or translation of a nucleotide sequence of interest in plants.

[0170] The regulatory sequences may include, but are not limited to, promoters, translation leader sequences, introns, and polyA recognition sequences.

[0171] The term "promoter" refers to a nucleic acid fragment capable of controlling the transcription of another nucleic acid fragment. In some embodiments of the present invention, the promoter is a promoter capable of controlling gene transcription in plant cells, regardless of whether it is derived from plant cells. The promoter can be a constitutive promoter or a tissue-specific promoter or a developmentally regulated promoter or an inducible promoter.

[0172] The term "constitutive promoter" refers to a promoter that will generally cause gene expression in most cell types in most cases. "Tissue-specific promoter" and "tissue-preferred promoter" are used interchangeably, and refer to a promoter that is mainly but not necessarily exclusively expressed in a tissue or organ, and also expressed in a specific cell or cell type. "Developmentally regulated promoter" refers to a promoter whose activity is determined by a developmental event. "Inducible promoter" responds to an endogenous or exogenous stimulus (environment, hormone, chemical signal, etc.) to selectively express an operably linked DNA sequence.

[0173] As used herein, the term "operably linked" refers to a connection of a regulatory element (for example, but not limited to, promoter sequence, transcription termination sequence, etc.) to a nucleic acid sequence (for example, a coding sequence or open reading frame) such that the transcription of the nucleotide sequence is controlled and regulated by the transcription regulatory element. The techniques for operably linking regulatory element region to nucleic acid molecule are known in the art.

[0174] The "introducing" a nucleic acid molecule (such as a plasmid, linear nucleic acid fragment, RNA, etc.) or protein into a plant refers to transforming a cell of the plant with the nucleic acid or protein so that the nucleic acid or protein can function in the plant cell. The term "transformation" used in the present invention comprises stable transformation and transient transformation.

[0175] The term "stable transformation" refers to that the introduction of an exogenous nucleotide sequence into a plant genome results in a stable inheritance of the exogenous gene. Once stably transformed, the exogenous nucleic acid sequence is stably integrated into the genome of the plant and any successive generations thereof.

[0176] The term "transient transformation" refers to that the introduction of a nucleic acid molecule or protein into a plant cell to perform function does not result in a stable inheritance of the foreign gene. In transient transformation, the exogenous nucleic acid sequence is not integrated into the genome of the plant.

[0177] Unless otherwise defined, all technical and scientific terms used herein have the same meanings as commonly

understood by those of ordinary skill in the art to which the present invention pertains. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, the preferred methods and materials are now described.

**[0178]** All publications and patents cited in this description are incorporated herein by reference as if each individual publication or patent is exactly and individually indicated to be incorporated by reference, and is incorporated herein by reference to disclose and describe methods and/or materials related to the publications cited. The citation of any publication which it was published before the filing date should not be interpreted as an admission that the present invention is not eligible to precede the publications of the existing invention. In addition, the publication date provided may be different from the actual publication date, which may require independent verification.

**[0179]** Unless specifically stated or implied, as used herein, the terms "a", "a/an" and "the" mean "at least one." All patents, patent applications, and publications mentioned or cited herein are incorporated herein by reference in their entirety, with the same degree of citation as if they were individually cited.

#### Brief Description of Drawings

**[0180]**

Figure 1 shows a schematic diagram of the method for generating new mutations in organisms according to the present invention. In the figure, only Cas9 with NGG as PAM is taken as an example; in the same way, other Cas9 variants with different PAM (such as NG) can also be used.

Figure 2 shows a schematic diagram of gRNA design at the ALS gene sites W574 and S653 of *Arabidopsis thaliana*.

Figure 3 shows the *Arabidopsis thaliana* T2 generation herbicide-resistant strains transformed with two different programmed sequential cutting/editing vectors. pQY743 and pQY745 are vector numbers. Resistant strains were capable of rooting normally, but the wild-type Col-0 and non-resistant strains were not.

Figure 4 shows a diagram of sequencing peaks of the ALS gene of the T2 generation strain of *Arabidopsis thaliana* resistant to imazapic, in which the T indicated by the arrow was mutated from G, resulting in a W574L mutation.

Figure 5 shows a design of the programmed sequential cutting/editing scheme at W574 site of ALS gene of *Arabidopsis thaliana*. The T-DNA sequence expressed four genes, sgRNA1, sgRNA2, Cas9 and HygR. Among them, sgRNA1 and Cas9 formed a complex, which cut the W574 codon of ALS in the genome, and was expected to form -G genotype through cellular spontaneous repair. This new sequence could be recognized and cut by the complex formed by sgRNA2 and Cas9, and formed +T genotype through cellular spontaneous repair, leading to W574L.

Figure 6 shows the resistant seedlings of *Arabidopsis thaliana* screened by imazapic and the sequencing results of ALS W574 site.

Figure 7 shows a design of the programmed sequential cutting/editing scheme at S653 site of ALS gene of *Arabidopsis thaliana*. The T-DNA sequence expressed four genes, sgRNA1, sgRNA2, Cas9 and HygR. Among them, sgRNA1 and Cas9 formed a complex, which cut the S653 codon of ALS in the genome. After cellular spontaneous repair, a -G genotype was formed. This sequence could be recognized and cut by the complex formed by sgRNA2 and Cas9. After cellular spontaneous repair, a +A genotype was formed, resulting in S653N.

Figure 8 shows the resistant seedlings of *Arabidopsis thaliana* screened by imazapic and the sequencing results of ALS S653 site. The left panel shows the screening results of resistant seedlings and the ratio of resistant seedlings, and the right panel shows a diagram of sequencing peaks and mutation type of S653 site.

Figure 9 shows a design of the programmed sequential cutting/editing scheme at W574 site of ALS gene of *Arabidopsis thaliana*. The T-DNA sequence expressed four genes, sgRNA1, sgRNA2, Cas9 and HygR. Among them, sgRNA1 and Cas9 formed a complex, which cut the W574 codon of ALS in the genome. After cellular spontaneous repair, the +A genotype was formed. This sequence could be recognized and cut by the complex formed by sgRNA2 and Cas9. After cellular spontaneous repair, the -G genotype was formed, resulting in W574M. The two cuttings used different PAM sites.

Figure 10 shows a design of the programmed sequential cutting/editing scheme at W2038 site of ACCase2 gene of *Oryza sativa*. This site corresponds to site W2027 of ACCase2 gene of *Alopecurus myosuroides*. The T-DNA

sequence expressed four genes, sgRNA1, sgRNA2, Cas9 and HygR. Among them, sgRNA1 and Cas9 formed a complex, which cut the W2038 codon of ACCase in the genome. A-G genotype was formed after cellular spontaneous repair. This sequence could be recognized and cut by the complex formed by sgRNA2 and Cas9. After cellular spontaneous repair, a +T genotype was formed, resulting in W2038L.

5 Figure 11 shows the resistant callus of *Oryza sativa* co-screened with hygromycin (50 ug/L) and quizalofop-p (50 ug/L) and the sequencing results of W2038 site.

10 Figure 12 shows a diagram of polyacrylamide gel electrophoresis of SpCas9 and NGA-Cas9 proteins which were produced by prokaryotic expression and purified. The band indicated by arrow is the Cas9 protein band.

15 Figure 13 shows the in vitro cutting activity of purified Cas9 protein on DNA fragments containing OsALS W548 target site and OsACCase2 W2038 target site, which was detected by agarose gel electrophoresis, and it can be seen that only when Cas9 protein and sgRNA fragment were added at the same time, the DNA fragments could be cut to an expected size.

20 Figure 14 shows a diagram of the sequencing peaks of the OsALS W548 target site of RNP transformed *Oryza sativa* protoplasts, and the site corresponds to ALS W574 site of *Arabidopsis thaliana*. At the place indicated by the arrow, in addition to the original G base signal peak, there was a T base signal peak obtained by mutation, which led to a W548L mutation.

25 Figure 15 shows the resistant callus of *Oryza sativa* edited by RNP at OsACCase2 W2038 site and screened with 50ug/L quizalofop-p. The arrow indicates the resistant callus.

30 Figure 16 shows a diagram of sequencing peaks of the OsACCase2 W2038 target site of *Oryza sativa* seedlings differentiated from resistant callus. The T indicated by the arrow was mutated from G, resulting in a W2038L mutation.

35 Figure 17 shows the resistance test of T1 generation OsACCase2 W2038L edited seedlings to Haloxylfop-p. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No.5 (a rice variety), and seedlings Nos. 2 to 4 show the results of wild-type Huaidao No.5 and two RNP gene gun-transformed T1 generation W2038L edited strains QY367-7-12 and QY367-7-18, all of which were treated with 5 g/mu active ingredient of Haloxylfop-p (mu, a unit of area, 1 mu = 1/15 hectares). It is clearly that the edited strains developed resistance to Haloxylfop-p.

40 Figure 18 shows the resistance test of T1 generation OsACCase2 W2038L edited seedlings to quizalofop-p. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No.5, and seedlings Nos. 2 to 4 show the results of wild-type Huaidao No.5 and two RNP gene gun-transformed T1 generation W2038L edited strains QY367-5-10 and QY367-5-21, all of which were treated with 5 g/mu active ingredient of quizalofop-p. It is clearly that the edited strains developed resistance to quizalofop-p.

45 Figure 19 shows the resistant callus of *Oryza sativa* edited by RNP simultaneously at OsACCase2 W2038 site and OsBADH2 gene and screened with 50ug/L quizalofop-p. The arrow indicates the resistant callus.

50 Figure 20 shows a diagram of sequencing peaks of OsACCase2 W2038 target site of T0 generation double-site edited seedlings. The T indicated by the arrow was mutated from G, resulting in a W2038L mutation.

55 Figure 21 shows a diagram of sequencing peaks of OsBADH2 target site of T0 generation double-site edited seedling. +A homozygous mutation occurred at the point indicated by an arrow.

Figure 22 shows the resistant callus of *Oryza sativa* edited by RNP simultaneously at OsALS W548 site and OsSWEET14 gene and screened with 5mg/L pyroxslam. The arrow indicates the resistant callus.

Figure 23 shows a diagram of sequencing peaks of OsALS W548 target site of T0 generation double-site edited seedlings. Both G base and T base signal peaks are present at the place indicated by an arrow, resulting in W548L mutation.

Figure 24 shows a diagram of sequencing peaks of OsSWEET14 target site of T0 generation double-site edited seedlings. -C homozygous mutation occurred at the site indicated by an arrow.

Figure 25 shows the resistance test of T1 generation OsALS W548 site and OsSWEET14 double-site edited seedlings to nicosulfuron. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No.5, and seedlings Nos. 2 to 3 show the results of the wild-type Huaidao No.5 and the RNP gene gun-transformed T1 generation W548L edited strain QY360-7-11, both of which were treated with 4 g/mu active ingredient of nicosulfuron. It is clearly that the edited strain developed resistance to nicosulfuron.

Figure 26 shows the resistance test of T1 generation OsALS W548 site and OsSWEET14 double-site edited seedlings to flucarbazone-Na. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No.5, and seedlings Nos. 2 to 3 show the results of the wild-type Huaidao No.5 and the RNP gene gun-transformed T1 generation W548L edited strain QY360-7-9, both of which were treated with 2 g/mu active ingredient of flucarbazone-Na. It is clearly that the edited strain developed resistance to flucarbazone-Na.

Figure 27 shows the resistance test of T1 generation OsALS W548 site and OsSWEET14 double-site edited seedlings to imazapic. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No. 5, and seedlings Nos. 2 to 3 show the results of the wild-type Huaidao No.5 and the RNP gene gun-transformed T1 generation W548L edited strain QY360-7-11, both of which were treated with 7 g/mu active ingredient of imazapic. It is clearly that the edited strain developed resistance to imazapic.

Figure 28 shows the resistance test of T1 generation OsALS W548 site and OsSWEET14 double-site edited seedlings to pyroxslam. From left to right, seedling No. 1 shows the results of the water treatment control of wild-type Huaidao No. 5, and seedlings Nos. 2 to 4 show the results of the wild-type Huaidao No.5 and the RNP gene gun-transformed T1 generation W548L edited strains QY360-7-2 and QY360-7-11, all of which were treated with 2 g/mu active ingredient of pyroxslam. It is clearly that the edited strains developed resistance to pyroxslam.

Figure 29 shows a scheme of programmed sequential cutting/editing of HBB gene in 293T cells. A. The design of HBB gene sites for programmed sequential cutting/editing in 293T cells is shown; B. The conversion efficiency of programmed sequential cutting/editing of HBB gene in 293T cells after 48 hours of transformation of each editing vector; C. The ratio of gene editing types generated from programmed sequential cutting/editing of HBB gene and single site cutting/editing of HBB gene in 293T cells; WT: wild type, indel: a genotype for deletion or insertion, C->T SNP: a genotype with C to T base substitution at cutting.

#### Description of Sequence Listing

**[0181]** The main sequences involved in the present invention are summarized as follows, and related sequences are provided in the sequence listing.

SEQ ID NO:	Sequence description
1	Amino acid sequence of ALS protein from <i>Arabidopsis thaliana</i>
2	DNA sequence of ALS gene from <i>Arabidopsis thaliana</i>
3	Amino acid sequence of AmACCase protein from <i>Alopecurus myosuroides</i>
4	DNA sequence of AmACCase gene from <i>Alopecurus myosuroides</i>
5	Amino acid sequence of OsHPPD protein from <i>Oryza sativa</i>
6	Genomic DNA sequence of OsHPPD protein from <i>Oryza sativa</i>
7	Amino acid sequence of OsPPO 1 protein from <i>Oryza sativa</i>
8	Genomic DNA sequence of OsPPO1 protein from <i>Oryza sativa</i>
9	Amino acid sequence of OsTIR1 protein from <i>Oryza sativa</i>
10	Genomic DNA sequence of OsTIR1 protein from <i>Oryza sativa</i>
11	Amino acid sequence of ALS protein from <i>Oryza sativa</i>
12	DNA sequence of ALS gene from <i>Oryza sativa</i>
13	Amino acid sequence of ACCase2 protein from <i>Oryza sativa</i>
14	Genomic DNA sequence of ACCase2 from <i>Oryza sativa</i>

(continued)

SEQ ID NO:	Sequence description
15	DNA sequence of SpCas9 protein optimized by plant codon
16	DNA sequence of NGA-Cas9 protein optimized by plant codon
17	Genomic DNA sequence of OsBADH2 from <i>Oryza sativa</i>
18	Genomic DNA sequence of OsSWEET14 from <i>Oryza sativa</i>
19	Amino acid sequence of StALS2 protein from potato ( <i>Solanum tuberosum L.</i> )
20	DNA sequence of StALS2 gene from potato ( <i>Solanum tuberosum L.</i> )
21	DNA sequence of HBB (hemoglobin subunit beta) gene in human embryonic kidney cell 293 T
22	CDS sequence of HBB gene in human embryonic kidney cell 293T
23	Amino acid sequence of HBB gene in human embryonic kidney cell 293T

**Embodiments for Carrying Out the Invention**

[0182] The present invention will be further explained in conjunction with the following examples. The following examples are illustrated by way of examples, but the protection scope of the present invention should not be limited to these examples. The experimental methods in the following examples, unless otherwise specified, were the methods described in commonly used molecular biology, tissue culture technology and agronomy manuals. For example, specific steps could be found in: "Molecular Cloning: A Laboratory Manual (3rd edition)" (Sambrook, J., Russell, David W., 2001, Cold Spring Harbor), "Plant Propagation by Tissue Culture" (Edwin F. George , Michael A. Hall, Geert-Jan De Klerk, 2008, Springer). The materials, reagents, instruments, etc. used in the following examples could be obtained from commercial sources unless otherwise specified.

**Example 1: Designing predictable base substitutions introduced by programmed sequential cutting/editing for W574 and S653 sites of ALS gene of *Arabidopsis thaliana***

## A. Experimental materials

1. *Arabidopsis thaliana* material

[0183] The wild-type *Arabidopsis thaliana* Col-0 was a model variety of dicotyledonous plant, its original seeds were provided by the Department of Weeds, College of Plant Protection, China Agricultural University, and the propagation and preservation thereof were performed by our laboratory according to standard methods in this field.

## 2. Vectors

[0184] The vector plasmids pCBC-dT1T2 (Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ 2014. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol. Nov 29;14(1):327, see <https://www.addgene.org/50590/>for details), pHEE401E (see Wang ZP, Xing HL, Dong L, Zhang HY, Han CY, Wang XC, Chen QJ Genome Biol. 2015 Jul 21; 16:144. doi: 10.1186/s13059-015-0715-0. <https://www.addgene.org/71287/>for specific information), and pHEE401E-NG (the mutation as reported in Nishimasu et al. 2018 Engineered CRISPR-Cas9 nuclease with expanded Targeting space. Science 361(6408):1259-1262. doi: 10.1126/science.aas9129 was introduced into pHEE401E to construct vector pHEE401E-NG capable of recognizing NG PAM) were purchased from Addgene website or constructed by our laboratory in accordance with conventional molecular biology methods, and kept by our laboratory.

## 3. Main equipment

[0185] Pipette gun, water bath, PCR instrument (Bio-rad T100), electrophoresis instrument (WIX-EP600), gel imager, electric blast dryer, centrifuge (Eppendorf 5424R), high-throughput tissue lyser, shaker, electronic balance, pH meter, etc.

## 4. Main reagents

[0186] High-fidelity DNA polymerase (purchased from Tsingke Bio), agarose gel recovery kit and plasmid extraction kit (purchased from Sparkjade), Bsal and T4 DNA ligases (purchased from NEB), Trans5 $\alpha$  competent cells and EHA105 competent cells (purchased from TransGen Biotech, Beijing, China), GV3101 *Agrobacterium* competent cells (purchased from Shanghai AngyuBio), Tris, EDTA, kanamycin, cephalosporin, hygromycin, agarose, yeast powder, tryptone, NaCl (purchased from Sangon Biotech), MS powder, sucrose, Silwet-77, hygromycin (purchased from Solarbio), nucleic acid dye (Dured), absolute ethanol (purchased from Sinopharm), etc.

## 10 5. Preparation of main solutions

## [0187]

1) Seed disinfectant: 4mL of 10% SDS, 20mL NaClO, added with water to make up 200mL.

15 2) SDS extraction buffer: 40mL of 1M Tris - HCl (pH 8.0), 50mL of 0.1M EDTA, 10mL of 5M NaCl, 10mL of 10% SDS, added with water to make up 200mL.

20 3) Infection solution: 1.5g of sucrose, 9 $\mu$ L of Silwet-77, added with 30mL of ultrapure water.

25 4) LB solid medium: 5g of yeast powder, 10g of tryptone, 10g of NaCl, 15g of agar, added with water to make up 1L, sterilize at 121°C for 15 minutes, poured on a plate for later use.

5) 50× TAE stock: 242g of Tris, 37.2g of Na<sub>2</sub>EDTA·2H<sub>2</sub>O, added with 800mL of ultrapure water, stirred well to dissolve, added with 57.1mL of acetic acid, stirred well, finally diluted to 1L with deionized water, and stored at room temperature.

30 6) MS solid medium: 4.42g of MS powder and 10g of sucrose were weighed, added with 800mL of ultrapure water, adjusted to pH 5.8, added with water to make up 1L, added 10g of phytigel, sterilized at 121°C for 15 minutes, and poured on a plate for later use.

## B. Experimental methods

## 1. Design and construction of CRISPR/Cas9 dual target vector

## 35 1.1 Target design

[0188] The *Arabidopsis thaliana* ALS gene sequence was shown in SEQ ID NO: 2. The target sequence gRNA1 (5'-GCATGGTTATGCAATGGGA-3') of 19 bases was designed using the AGA near *Arabidopsis thaliana* ALS574 site as PAM, it was predicted that one G base would be deleted between the first 3-4 sites of PAM after editing, then a second target sequence gRNA2 (5'-GGCATGGTTATGCAATGGA-3') was designed based on the sequence generated from the deletion, and it was predicted that one T base would be inserted via a second editing, thereby realizing conversion of TGG-TTG, as shown in Figure 2.

[0189] Similarly, the target sequence gRNA3 (5'- TGCCGATGATCCCGAGTGG-3') of 19 bases was designed using the TGG near *Arabidopsis thaliana* ALS653 site as PAM, it was predicted that one G base would be deleted between the first 3-4 positions of PAM after editing, then a second target sequence gRNA4 (5'-TTGCCGATGATCCCGATGG-3') was designed based on the sequence generated from the deletion, and it was predicted that one A base would be inserted after the second editing, thereby realizing conversion of AGT-AAT, as shown in Figure 2.

## 50 1.2 Vector construction

[0190] The method described by Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ 2014. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol. Nov 29;14(1): 327 was followed. Specifically, dT1T2 plasmid was used as a template to amplify the ALS574 and 653 sites dual-target fragment respectively to construct an sgRNA expression cassette. The vector backbones of pHEE401E and pHEE401E-NG were digested with Bsal, and the bands were cut from the gel and recovered, and the target fragment was directly used for the ligation reaction after digestion. T4 DNA ligase was used to ligate the vector backbones and the target fragment, the ligation products were transformed into Trans5 $\alpha$  competent cells, different monoclonal cultures were picked out for sequencing. After

confirmation via sequencing, the Sparkjade High Purity Plasmid Mini Extraction Kit was used to extract the plasmids to obtain the recombinant plasmids, which were respectively named as pQY743 and pQY745.

2. Design of primers for target detection

[0191] Primers for target detection took ALS574 and ALS653 target sites as centers, wherein the primer for upstream detection was about 100bp from the ALS574 target site, and the primer for downstream detection was about 280bp from the ALS653 target site. The primer sequences were as follows: 574/653checking-F: 5'ATTGACGGAGATGGAAGCTT3', and 574/653checking-R: 5'CCAAACTGAGCCAGTCACAA3'.

3. Establishment of *Arabidopsis thaliana* genetic transformation system

3.1 *Agrobacterium* transformation

[0192] The constructed recombinant plasmids were transformed into *Agrobacterium* GV3101 competent cells to obtain recombinant *Agrobacterium* cells.

3.2 Preparation of *Agrobacterium* infection solution

[0193]

- 1) The activated *Agrobacterium* was picked and inoculated in 30ml of YEP liquid medium (containing 25mg/L Rif and 50mg/L Kan), and cultured under shaking at 200 rpm and 28°C overnight until the OD600 value was about 1.0-1.5.
- 2) After centrifugation at 6000 rpm for 10 minutes, the *Agrobacterium* was collected, and the supernatant was discarded.
- 3) The *Agrobacterium* was resuspended in the infection solution (no need to adjust pH) to OD600=0.8 for later use.

3.3 Transformation of *Arabidopsis thaliana*

[0194]

- 1) Before plant transformation, the attention was paid to whether the plant grew well, the inflorescence was luxuriant, and there was no stress response. The first transformation was carried out when plant height was about 20 cm. Watering could be properly carried out when the soil was dry. The grown siliques were cut off with scissors on the day before transformation.
- 2) The inflorescence of the plant to be transformed was immersed in the above solution for 30 seconds to 1 minute under gently stirring, and there should be a layer of liquid membrane on the infiltrated plant.
- 3) After the transformation was completed, the plant was cultured in the dark for 24 hours, and then taken out and placed in a normal light environment for growth.
- 4) After one week, the second transformation could be carried out in the same way.

3.4 Seed harvest

[0195] After the seeds were mature, they were harvested. After harvesting, the seeds were dried in an oven at 37°C for about one week.

4. Selection of transgenic plant

[0196] The seeds were treated with disinfectant for 5 minutes, washed with deionized water for 5 times, and evenly spread on the MS selection medium (containing 30µg/mL hygromycin, 100µg/mL cephalosporin), the medium was placed in a light incubator (temperature 22°C, 16 hours light, 8 hours dark, light intensity 100-150µmol/m<sup>2</sup>/s, humidity 75%), and one week later, the positive seedlings were selected and transplanted to soil.

## 5. Detection of T1 mutant plants

## 5.1 Extraction of genomic DNA

5 [0197]

- 1) The leaves of *Arabidopsis thaliana* were cut and placed in a 2mL centrifuge tube, added with steel balls, and the leaves were ground with the high-throughput tissue lyser.
- 10 2) After the grinding was complete, 400 $\mu$ L of SDS extraction buffer was added, mixed upside down, and incubated in a 65°C water bath for 15 minutes and mixed upside down every 5 minutes.
- 15 3) Centrifugation was carried out at 13,000 rpm for 5 minutes.
- 20 4) 300 $\mu$ L of the supernatant was pipetted and transferred to a new 1.5mL centrifuge tube, an equal volume of isopropanol pre-cooled at -20°C was added to the centrifuge tube, and the centrifuge tube was placed at -20°C for 1 hour or overnight.
- 25 5) Centrifugation was carried out at 13,000 rpm for 10 minutes and the supernatant was discarded.
- 30 6) 500 $\mu$ L of 70% ethanol was added to the centrifuge tube to wash the pellet, the washing solution was discarded after centrifugation (be careful not to discard the pellet), dried at room temperature, 30 $\mu$ L of ultrapure water was added to dissolve the DNA, and the DNA was stored at -20°C.

## 25 5.2 PCR amplification

[0198] The extracted T1 plant genome was used as a template, the detection primers were used to amplify target fragment, 5 $\mu$ L of amplified product was pipetted and detected by 1% agarose gel electrophoresis, and imaged with the gel imager. The remaining product was delivered to the sequencing company to directly perform sequencing.

## 30 6. Detection of T2 mutant plant

[0199] After the seeds of the T1 strain were harvested from single plant, the seeds of different strains of two vectors were selected and spread on the imazapic selection medium (MS medium + 0.24 $\mu$ g/mL imazapic) to perform selection, and the positive seedlings were transplanted to the soil one week later and subjected to molecular detection, in which the method was the same as step 5.

[0200] The used primers and sequences thereof:

Primer name	Sequence (5'-3')
ALS574-F	ATATATGGTCTCGATTGGCATGGTTATGCAATGGGAGTT TTAGAGCTAGAAATA
ALS574-R	ATTATTGGTCTCGAAACTCCATTGCATAACCATGCCAA TCTCTTAGTCGACTC
ALS653-F	ATATATGGTCTCGATTGTGCCGATGATCCCGAGTGGGTT TTAGAGCTAGAAATA
ALS653-R	ATTATTGGTCTCGAAACCCATCGGGATCATCGGCAACAA TCTCTTAGTCGACTC
ALS574/653checking-F	ATTGACGGAGATGGAAGCTT
ALS574/653checking-R	CCAAACTGAGCCAGTCACAA

## C. Experimental results

## 1. Genotype detection of T1 plant

[0201] T1 seeds were selected by MS hygromycin resistance medium, a total of 32 positive seedlings were obtained for the pQY743 vector, and a total of 18 positive seedlings were obtained for the pQY745 vector. For each vector, 10 seedlings were selected and leaf genomic DNA thereof was extracted to detect the target site. It was found that there was no editing occurred in the T1 generation at ALS574 site, and there were editing events that met the design expectations at ALS653 site. The detection results are shown in Table 1:

Table 1: Mutation types of T1 plants

ALS site	Vector number	T1 plant number	Edit type
574	pQY743	1-10	WT
653	pQY745	3,4-8,10	Chimera
		1,2,9	Heterozygote

## 2. Selection results of T2 generation seeds

[0202] After the T2 generation seeds of single plant were harvested, they are spread on the imazapic resistant medium to perform selection, it could be seen that the wild-type Col-0 could not grow on the resistant medium, while the positive plants of the mutant strains could grow normally on the resistant medium, as shown in Figure 3.

[0203] For each vector, 10 seedlings were selected and genomic DNA thereof was extracted from leaves for molecular detection. It was found that there were 6 strains for the pQY743 vector that had homozygous mutation in line with the expectations, i.e., mutation from TGG to TTG, as shown in Figure 4. In addition, there was 1 strain with heterozygous mutation, and there were 3 strains with chimeric mutation. The detection results are shown in Table 2.

Table 2: Mutation types of T2 plants

ALS site	Vector No.	T2 plant No.	Edit type
574	pQY743	2,3,5,6,7,9	Homozygote
		4	Heterozygote
		1,8,10	Chimera

[0204] The above results showed that by using the technical solution of programmed sequential cutting/editing of the present invention, the design and realization of expected mutations at the target site could be achieved, and the base substitution mutations could be realized by designing sequential sgRNA combinations only with Cas9 protein.

Example 2: Achieving multiple mutation types by programmed sequential cutting/editing at W574 and S653 sites of *Arabidopsis thaliana* ALS gene

[0205] The operation steps for vector design, construction, and *Arabidopsis thaliana* transformation and selection were performed by referring to Example 1. For AtALS W574 site, the vector design was the same as that of Example 1. The schematic diagram of the vector was shown in Figure 5, and it was expected that the W574L mutation was to be realized by first -G then +T at a specific site. No editing event was detected in the T1 generation of the vector transformed *Arabidopsis thaliana*. The T2 generation of the transgenic *Arabidopsis thaliana* was selected with 0.24 mg/L imazapic, and a large number of herbicide-resistant plants were obtained, as shown in the left panel of Figure 6, the molecular detection of these resistant plants was performed, the PCR product sequencing results showed that not only the expected W574L mutation appeared at the W574 site, but also another resistant mutation W574M appeared, as shown in the right panel Figure 6.

[0206] For the AtALS S653 site, the vector design was the same as that of Example 1. The vector diagram was shown in Figure 7, and it was expected that the S653N mutation was to be realized by first -G then +A at a specific site. An expected editing event of S653N was detected in the T1 generation of the vector transformed *Arabidopsis thaliana*. The T2 generation of the transgenic *Arabidopsis thaliana* was continuously selected with 0.24 mg/L imazapic, and a large number of herbicide-resistant plants were obtained, as shown in the left panel of Figure 8, molecular detection was

performed to these resistant plants, the PCR product sequencing results showed that not only the expected S653N mutation occurred at S653 site, but also two other resistant mutations S653R and S653R/G654D occurred, as shown in the right panel of Figure 8.

**[0207]** The above results showed that by using the technical solution of programmed sequential cutting/editing of the present invention, in addition to the expected mutations designed for the target site, multiple functional mutation types could also be generated by designing sequential sgRNA combinations with only corresponding Cas9 protein, so that it is a suitable tool for creating new functional mutations.

Example 3: Generating W574M resistance mutation by using different PAMs to perform programmed sequential cutting/editing near the W574 site of *Arabidopsis thaliana* ALS gene

**[0208]** The operation steps for vector design, construction, and *Arabidopsis thaliana* transformation and selection were performed by referring to Example 1, except that, for the sequence 5'CTTGGCATGGTTATGCAATGgg3' near the AtALS W574 site, the GG closer to the W574 site was used first as PAM to design sgRNA1: 5'CTTGGCATGGGTATGCAATG3', in which the W574 site was underlined and the NG PAM was shown in italics. It was predicted that a new sequence 5'CTTGGCATGGTTATGCAAATGGGAag3' was to be formed by +A after cutting and repair. The AG was used as a new PAM site to design sgRNA2: 5'CTTGGCATGGTTATGCAAATGGGA3', and -G genotype was formed after spontaneous repair of cells, resulting in W574M. That was, different PAM sites were used for the two cuttings in this scheme. The vector diagram was shown in Figure 9. The vector was used to transform *Arabidopsis thaliana*, the T1 generation transgenic strain was subjected to genotype detection from which the expected editing event of W574M was detected, and the plant showed resistance to imazapic treatment.

**[0209]** The results showed that by using the technical solution of the present invention and using different PAMs to perform programmed sequential cutting/editing, the editing could be carried out in a wider sequence range to obtain amino acid substitution.

Example 4: Designing predictable base substitutions for D350 and W548 sites of *Oryza sativa* ALS gene

**[0210]** The mutation as reported in Nishimasu et al. 2018 Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science 361(6408): 1259-1262. doi: 10.1126/science.aas9129 was introduced into pHUE411 (A CRISPR/Cas9 toolkit for multiplex genome editing in plants. Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ. BMC Plant Biol. 2014 Nov 29;14(1):327. 10.1186/s12870-014-0327-y, see details in <https://www.addgene.org/71287/>) to construct the vector pHUE411-NG capable of recognizing NG PAM

**[0211]** The *Oryza sativa* ALS gene sequence was shown in SEQ ID NO: 12. By using 5'GGCGTGCGGTTGATGATCG3' (the underlined part was the OsALS-D350 site corresponding to *Arabidopsis thaliana* ALS-D376) and a new sequence 5'GGCGTGCGGTTGATGACG3' that was predicted to be generated from editing as targets, a dual-target vector was constructed according to the method described in Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ. BMC Plant Biol. 2014, and it was expected to obtain the conversion of GAT-GAA, resulting in OsALS D350E mutation.

**[0212]** By using 5'GGTATGGTTGTGCAATGGGA3' (the underlined part was the OsALS-W548 site corresponding to *Arabidopsis thaliana* ALS-W574) and a new sequence 5'GGTATGGTTGTGCAATGGA 3' that was predicted to be generated from editing as targets, a dual-target vector was constructed, and it was expected to obtain the conversion of TGG-TTG, resulting in OsALS W548L mutation.

**[0213]** Then the two vectors were transferred into *Oryza sativa* to obtain transgenic plants, and the identification indicated that the plants with expected substitutions D350E and W548L were obtained. The results of herbicide resistance biotest in field showed that the D350E and W548L mutants acquired the resistance to ALS inhibitor herbicides.

Example 5: Designing predictable base substitutions and selecting multiple mutation types for W2038 site of *Oryza sativa* ACCase2 gene

**[0214]** The *Oryza sativa* ACCase2 gene sequence was shown in SEQ ID NO: 14, in which OsACCase2 W2038 site corresponded to the ACCase W2027 site of *Alopecurus myosuroides*. The AGG close to this site was used as PAM to design sgRNA1: 5'TTCATCCTCGCTAAC-TGAG3', and it was predicted that a new sequence was formed by -G after cutting and repair. The AGG was continuously used as PAM to design sgRNA2: 5'CTTC-ATCCTCGCTAACTGAG3', and +T genotype was formed after cutting and repair again, resulting in the W2038L mutation. The sgRNA1 and sgRNA2 were constructed on the pHUE411 vector to form an editing vector, and the vector diagram was shown in Figure 10.

**[0215]** The editing vector was used to transform the callus of Huaidao No 5 (a rice variety), and after 3 weeks of co-selection with 50 µg/L hygromycin and 50 µg/L quizalop-p, a large number of resistant calli were obtained, as shown on the left panel of Figure 11. The resistant callus was taken for genotype identification, and it was found that not only

the expected W2038L mutation but also the W2038C mutation occurred, as shown on the right panel of Figure 11.

**[0216]** The above results of programmed sequential cutting/editing of *Oryza sativa* gene showed that the technical solution of the present invention was applicable to both monocotyledonous and dicotyledonous plants.

5 Example 6: Expression and purification of SpCas9 and NGA-Cas9 proteins

10 1. Experimental instrument and reagents

**[0217]**

Table 3: Experimental reagents

Reagent	Preparation and usage
Q5 DNA polymerase	Purchased from NEB
I5 DNA polymerase and I5 reaction solution	Purchased from Tsingke Bio
dNTPs	Shanghai Sangon
Ni-NTA resin	Shanghai Sangon
Tris/NaCl/imidazole	Shanghai Sangong
DTT	Shanghai Sangong
β-mercaptoethanol	Sigma
SDS-PAGE precast gel	Nanjing GenScript
Superdex200 column	GE Healthcare

25 Table 4: Experimental Instruments

Instrument	Model
PCR instrument	Bio-rad T100
Constant temperature shaking bed	Ounuo HNY-200B
Cell disrupter	Ningbo Xinzhi JY92-IIN
Electrophoresis apparatus	WIX-EP600
Protein purifier	GE AKTA pure
High speed large capacity refrigerated centrifuge	Xiangzhi ZX21K
High-speed centrifuge	Eppendorf 5424R
White light film viewer	Beijing Liuyi WD-9406
Nucleic acid gel imager	Beijing Sage ChampGel™ 5000
Micro spectrophotometer	DS-II

40 2. Experimental method

2.1 Construction of pET15b-Cas9 expression vector

**[0218]** The DNA sequences of SpCas9 and NGA-Cas9 proteins obtained after plant codon optimization were shown in SEQ ID NO: 15 and SEQ ID NO: 16, respectively, and the sequences were synthesized by GenScript as template DNAs.

**[0219]** After the NG-Cas9 and NGA-Cas9 sequences were amplified separately, the two fragments were ligated to a pET15b expression vector by infusion method, transformed into DH5a, and sequenced after verification.

50 Table 5: PCR verification amplification system and primers

Reaction system, 50 ul	
2 × I5 reaction solution	25 ul
Template DNA	1 ul
Forward and reverse primers	2 ul each
Ultrapure water	Added to 50 ul

(continued)

	Primer sequence:	5'-3'
5	pET15b-NG Cas9 -F	gtggccgcggcagccatatggattacaaggaccacgacg
	pET15b-NG Cas9-R	tttgttagcagccggatcctactcttcgttgcctgc
	pET15b-NGA Cas9 -F	tgcggccggcagccatatgtggattacaaggaccacgacg
	pET15b-NG Cas9 -R	tttgttagcagccggatcctactcttcgttgcctgtccc

## 10 2.2 Expression and purification of proteins

**[0220]** The constructed expression vectors were transformed into *Escherichia coli* Rosetta (DE3), the expression thereof was induced by IPTG, and the bacteria were harvested, lysed and purified by Ni-NTA column. The specific method was as follows:

- 15 a) The recombinant expression vectors were transformed into Rosetta (DE3) strain, a single colony was picked into 10 ml of LB medium, CmR+Amp (pET15b) or CmR+Kana (pET28a) resistance, cultured at 200 rpm at 37°C overnight, transferred to a 2 L shake flask containing 1 L of LB medium, cultured at 37°C and 200 rpm until the OD600 reached 0.6-0.8, cooled to 18°C, the expression thereof was induced with 0.5 mM IPTG overnight, and the bacteria were harvested by centrifugation at 4000g.
- 20 b) The harvested bacteria were resuspended in Ni-buffer A: 50 mM HEPES, pH7.4, 500 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, added with PMSF with a final concentration of 1 mM and 250 ul of Cocktail inhibitor, and then mixed well.
- 25 c) After the resuspended cells were disrupted by an ultrasonic disruptor, they are centrifuged at 40000g for 30 minutes at 4°C, and the supernatant was collected and passed through a Ni-NTA column.
- 30 d) Purification via Ni-NTA column: the lysate supernatant was combined with resin for 20 minutes, eluted with buffer A containing 50 mM imidazole to remove impurities, and finally eluted with an elution buffer containing 400 mM imidazole.
- e) SDS-PAGE gel electrophoresis system was used to detect the purification effect of protein.
- 35 f) Dialysis was performed by changing the buffer to 50 mM HEPES pH 7.5, 150 mM KCl, 1 mM DTT, 3% glycerol.
- h) The final sample was subjected to SDS-PAGE gel electrophoresis to detect the purification effect of the protein. After concentration by ultrafiltration, the protein was frozen at -80°C for later use.

## 40 2.3 Expression and purification results of Cas9 fusion proteins

**[0221]** The purification results of SpCas9 and NGA-Cas9 proteins were shown in Figure 12. The arrows indicated the Cas9 protein bands, indicating that a higher purity was achieved.

45 Example 7: In vitro enzyme cleavage activity of SpCas9 and NGA Cas9 proteins separately detected to OsACCase2 W2038 and OsALS W548 target sites

**[0222]**

- 50 1. The DNA sequences of OsALS and OsACCase2 genes were input into the CRISPOR online tool (<http://crispor.te-for.net/>), the following sgRNAs were designed for the W548 of OsALS (the 548<sup>th</sup> amino acid of OsALS, the amino acid sequence of *Oryza sativa* ALS protein was shown in SEQ ID NO: 11) corresponding to the amino acid site 574 of *Arabidopsis thaliana* ALS protein and the W2038 of OsACCase2 (the 2038<sup>th</sup> amino acid of OsACCase2, the amino acid sequence of *Oryza sativa* ACCase2 protein was shown in SEQ ID NO: 13) corresponding to the amino acid site 2027 of *Alopecurus myosuroides*, and these sgRNAs were synthesized by GenScript:

55

&gt;sgRNA1-548-G:

5'GGGUUAUGGUUUGUGCAAUGGGAguuuuagagcuagaaaaugcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'

5 > sgRNA1-2038-G:  
  
10 5'GUUCAUCCUCGCUAACUGGAGguuuuagagcuagaaaaugcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'

15 2. The specific detection primers, OsALS265AA-F: 5'ggcttgcgtctgggtggc3' and OsALS-end-R:  
5'ccatgccaaggcacatcaaacaag3', were used to amplify the fragment containing the OsALS W548 target site, and the PCR product was 1200 bp in length.  
The specific detection primers, OsACC1750AA-F: 5'gcgaagaagactatgctgtattgg3' and OsACC2196AA-R:  
5'cttaatcacacaccttcgcagcc3', were used to amplify the fragment containing the OsACCase W2038 target site, and the PCR product was 1500 bp in length.

20 The PCR system was shown in the following table:

Component	Volume
2× I5 reaction solution	25 μL
Forward primer (10 μM)	2 μL
Reverse primer (10 μM)	2 μL
Genomic DNA template	2 μL
Ultrapure water	Added to 50 μL

30 3. The PCR reaction was established, and the reaction conditions were shown in the following table:

Step	Temperature	Time
Pre-denaturation	98 °C	30 s
30-35 amplification cycles	98 °C	15 s
	58 °C	15 s
	72 °C	30 s
Final extension	72 °C	3 min

45 4. The PCR products were detected by agarose gel electrophoresis and subjected to sequencing for further verification. After being correctly verified, the gel was cut to recover DNA fragments. The recovered DNA fragments were dissolved by adding 30ul of RNase-free ultrapure water, and the concentration thereof was measured.  
5. The following detection system was used to detect Cas9 protein activity, and after the system was prepared, it was incubated at 37°C for 1 hour.

Components	Volume
Cas9 protein	0.5μL (lug)
sgRNA	0.5μL (lug)
10× Cas9 reaction buffer	2 μL
Recovered DNA fragment	XμL (100ng)
RNase-free ultrapure water	Added to 20 μL

55 6. After the reaction was completed, the system was treated at 65°C for 10 minutes, added with 4 ul of 6× DNA loading

buffer, and run on 2% agarose gel to detect band size.

**[0223]** The results were shown in Figure 13, and it can be seen that the purified Cas9 proteins could cut DNA double-strand at the designed target sites only in the presence of sgRNA.

5 Example 8: Achieving W548L mutation at OsALS548 site by transforming *Oryza sativa* protoplast with two different targeting RNP complexes

**[0224]**

10 1. The *Oryza sativa* protoplast preparation and PEG-mediated transformation were performed based on the partially modified method that had been published (Bart et al., 2006), and the specific preparation steps were as follows:

15 (1) The *Oryza sativa* seedlings for protoplasts were firstly prepared. The rice variety was Nipponbare. The rice seeds were hulled first, and the hulled seeds were rinsed with 75% ethanol for 1 minute, treated with 5% (v/v) sodium hypochlorite for 20 minutes, then washed with sterile water for more than 5 times, placed on an ultra-clean table and blow-dried, and then placed in a tissue culture flask containing 1/2 MS medium, with 20 seeds in each flask. Protoplasts were prepared by incubating the seeds at 26°C and 12 hours of light for about 10 days.

20 (2) The leaf sheath of the seedlings was selected, cut with a sharp Gillette razor blade into about 1 mm pieces, and placed in 0.6M mannitol and MES culture medium (formulation: 0.6M mannitol, 0.4M MES, pH 5.7) for later use. After all the materials were cut, they were transferred to 20 mL of enzymolysis solution (formulation: 1.5% cellulase R10/RS (YaKult Honsha), 0.5% Mecerozyme R10 (YaKult Honsha), 0.5M mannitol, 20mM KCl, 20mM MES, pH5.7, 10mM CaCl<sub>2</sub>, 0.1% BSA, 5 mM β-mercaptoethanol), wrapped in aluminum foil and placed in a shaker at 28°C. The enzymolysis was carried out at 50 rpm in the dark for about 4 hours, and the speed was increased to 100 rpm for the last 2 minutes;

25 (3) After enzymolysis, an equal volume of W5 solution (formulation: 154mM NaCl, 125mM CaCl<sub>2</sub>, 5mM KCl, 15mM MES) was added, shaken horizontally for 10 seconds to release the protoplast. The cells resulted from enzymolysis were filtered through a 300-mesh sieve and centrifuged at 150 g for 5 minutes to collect protoplast;

30 (4) The cells were rinsed twice with W5 solution, and the protoplast was collected by centrifugation at 150 g for 5 minutes;

35 (5) The protoplast was resuspended with an appropriate amount of MMG solution (formulation: 3.05g/L MgCl<sub>2</sub>, 1g/L MES, 91.2g/L mannitol), and the protoplast concentration was about 2×10<sup>6</sup> cells/mL.

2. Preparation of RNP complex:

40 **[0225]** According to the OsALS548 target site sequence GGGTATGGTTGTGCAATGGGAgga (the OsALS W548 site was underlined, corresponding to *Arabidopsis thaliana* ALS W574, and the PAM site recognized by Cas9 protein was shown in italic lowercase), the purified NGA-Cas9 protein was selected to prepare the RNP complex. The GGA was used as PAM to design >CrRNA1-548-G: 5'-GGGUAUGGUUGUGCAAUGGGAguuuuagagcuaugcu-3', it was predicted that one G base would be deleted between the first 3-4 positions of PAM after editing, and then the sequence resulted from the above deletion was used to design a second >CrRNA2-548+T: 5'-UGGUAUGGUUGUGCAAUGGGAguuuuagcuaugcu-3', it was predicted that one T base would be inserted after the second editing to obtain the conversion of TGG-TTG.

45 **[0226]** >CrRNA1-548-G and >CrRNA1-548+T were synthesized by GenScript Biotechnology Company, and sgRNA was also synthesized:

50 >sgRNA1-548-G:

55 5'GGGUAUGGUUGUGCAAUGGGAguuuuagagcuaugaaauagcaaguuaaaaauaaggcuauguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'

>sgRNA2-548+T:

5'UGGUUAUGGUUGCAAUAGGAguuuuagagcuagaaaaugcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'.

- 5 [0227] The synthesized crRNA and GenCRISPR tracrRNA (GenScript SC1933) were mixed equimolarly, added with crRNA&tracrRNA annealing buffer (GenScript SC1957-B) and annealed to prepare gRNA according to the instructions. The tracrRNA sequence was 5'-agcauagcaaguuaaaaaggcuaguccguuaacuugaaaaaguggcaccgagucggugcuuu-3'.  
10 [0228] The RNP reaction system was prepared according to the following table, and after the reaction system was prepared, it was incubated at 25°C for 10 minutes.

Components	Volume
NGA Cas9 protein	10µL (20 µg)
gRNA or sgRNA	10µL (20 µg)
10x× Cas9 reaction buffer	10 µL
RNase-free ultrapure water	70 µL

20 3. Protoplast transformation

[0229]

- 25 (1) 200 µl of MMG resuspended protoplast as prepared above was taken and added with the RNP complex (20µg of Cas9 protein, 20µg of sgRNA) generated after the incubation, and gently flicked to mixed well.  
  
 (2) An equal volume of 40% (w/v) PEG solution (formulation: 40%(w/v) PEG, 0.5M mannitol, 100mM CaCl<sub>2</sub>) was added, gently flicked to mix well, and stood at 28°C in dark for 15 minutes;  
  
 30 (3) After induction and transformation, 1.5 mL of W5 solution was added slowly, gently flicked to mix the cells well, and the cells were collected by centrifugation at 150 g for 3 minutes. This step was repeated once;  
  
 (4) the cells were resuspended by adding 1.5 mL of W5 solution, placed in a 28°C incubator and cultured in dark for 12-16 hours. The cells should be incubated for 48 to 60 hours if they were used to extract protoplast genomic DNA.

35 4. Detection of genome target editing event

[0230]

- 40 (1) The protoplast DNA was extracted using the partially modified CTAB method, and the method was specifically described as follows: the protoplast was centrifuged and the supernatant was discarded, 500 µL of DNA extraction solution was added, shaken and mixed well, and incubated in a 65°C water bath for 1 hour; the sample was cooled after the incubation in water bath, added with an equal volume of chloroform, inverted for several times to mix well and centrifuged at 10,000 rpm for 10 minutes; 400 µl of supernatant was taken and transferred to a new 1.5 mL centrifuge tube, added with 1 mL of 70% (v/v) ethanol, and stood to precipitate at -20°C for 20 minutes; the DNA was precipitated by centrifugation at 12,000 rpm for 15 minutes, the precipitate was dried in the air and then dissolved by adding 50 µl of ultrapure water, and stored at -20°C for later use.  
  
 45 (2) The gene-specific primers were used to amplify the fragment containing the W548 target site, and the following detection primers were designed for the target site:

50 OsALS500AA-F: 5'GGCTAACCCAGGTGTCACAG3',

OsALS-3'UTR-R: 5'CCATGCCAAGCACATCAAACAAG3'

55 The PCR reaction system was shown in the following table:

Components	Volume
2× I5 reaction solution	25 µL
Forward primer (10 µM)	2 µL
Reverse primer (10 µM)	2 µL
Genomic DNA template	2 µL
Ultrapure water	Added to 50 µL

5           (3) The PCR reaction was established, and the general reaction conditions were shown in the following table:

Step	Temperature	Time
Pre-denaturation	98 °C	30 s
30-35 amplification cycles	98 °C 58 °C 72 °C	15 s 15 s 30 s
Final extension	72 °C	3 min

10           (4) The detection was carried out by agarose gel electrophoresis, and the PCR fragments were recovered and sequenced.

## 5. Experimental results:

15           [0231] By either using the gRNA prepared with the synthetic crRNA and tracrRNA by annealing or directly using the synthetic sgRNA, an active RNP complex could be formed with the purified NGA-Cas9 protein. By sequencing the OsALS548 targeted site, the mutation from TGG to TTG could be detected, which demonstrated that the site-specific mutation of the target site in the cell could be achieved by the programmed sequential cutting/editing generated from the RNP complex in combination with crRNA or sgRNA in sequential order. As shown in Figure 14, at the arrow point in the diagram of protoplast OsALS548 target sequencing peaks, in addition to the original G base signal peak, there was also the T base signal peak generated from the mutation, which resulted in the W548L mutation at OsALS548 site.

20           Example 9: Achieving W2038L mutation at OsACCase2 W2038 site by bombarding *Oryza sativa* callus with two different targeted RNP complexes

### 35           1. Preparation of RNP complexes:

25           [0232] According to the OsACCase2 W2038 target site sequence GTTCATCCTCGCTAATGGAGgg, in which the OsACCase2 W2038 site, corresponding to the *Alopecurus myosuroides* ACCase W2027, was underlined, and the PAM site recognized by the Cas9 protein was shown in italic lowercase, the purified SpCas9 protein was selected to prepare the RNP complexes. The GGA was used as PAM to design >CrRNA1-2038-G: 5'-GUUCAUCCUCGCUAACUGGAGgg-uuuuagagcuauugcu-3', it was predicted that one G base would be deleted between the first 3-4 positions of PAM after editing, then the sequence generated from the deletion was used to design a second >CrRNA2-2038+T: 5'-UGUUCAUC-CUCGCUAACUGAGguuuuagagcuauugcu-3', it was predicted that one T base would be inserted after the second editing, thus the conversion of TGG-TTG was obtained.

30           [0233] >CrRNA1-2038-G and >CrRNA2-2038+T were synthesized by GenScript Biotechnology Company, and sgRNA was also synthesized:

45           >sgRNA1-2038-G:

50           5'GUUCAUCCUCGCUAACUGGAGgguuuuagagcuagaaaauagcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc 3'

55           >sgRNA2-2038+T:

5'UGUUCAUCCUCGCUAACUGAGguuuuagagcuagaaauagcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc 3'.

- 5 [0234] The synthesized crRNA and GenCRISPR tracrRNA (GenScript SC1933) were mixed equimolarly, added with crRNA&tracrRNA annealing buffer (GenScript SC1957-B) and annealed to prepare gRNA according to the instructions.  
 [0235] The RNP complexes were prepared by incubating in the same reaction system as in Example 8. Taking the amount of 10 gene gun bombardments for transformation as example: 20 µg of Cas9 protein, 20 µg of gRNA or sgRNA, 10 µl of 10× Cas9 reaction buffer, made up to 100 µl in total with RNase-free ultrapure water, incubated at 25°C for 10 minutes, and mixed gently.

10 2. Induction of *Oryza sativa* callus

- 15 [0236] The mature and plump *Oryza sativa* seeds were selected, hulled, and disinfected according to the following steps:

- 20 (1) The rice (*Oryza sativa*) seeds were hulled, wherein the rice variety was Huaidao No.5, purchased from the seed market.  
 (2) The seeds were washed with sterile water for unlimited times until the washed water became clear.  
 25 (3) After being sterilized with 70% alcohol for 1 minute, the seeds were then placed in 10% sodium hypochlorite on a horizontal shaker and incubated for 25 minutes under shaking.  
 (4) After sodium hypochlorite disinfection, the seeds were washed with sterile water for 5 times. The seeds were seeded on callus induction medium (formulation: MS powder (4.42g/L) + 2,4-D (2mg/L) + sucrose (30g/L) + phytagel (4g/L)), and cultivated in the dark at 28°C to induce callus.

30 3. Gene gun RNP bombardment:

[0237]

- 35 (1) Hypertonic cultivation  
 The callus with good embryogenicity was transferred to hypertonic medium (formulation: MS powder (4.42g/L) + 2,4-D (2mg/L) + sucrose (30g/L) + D-mannitol (0.4M) ) + phytagel (4g/L)), sterile operation was carried out on an ultra-clean bench, and cultivation was carried out in the dark at 25°C for 4-6 hours.
- 40 (2) Preparation of gold powder suspension: by using an imported 1.5 mL EP tube, 30 mg of gold powder (diameter 0.6 µm) was weighed; added with 1 mL of 70% ethanol, subjected to vortex well, and stood on ice for 10 minutes. After centrifugation for 1 minute, the supernatant was discarded; 1 mL of sterile water was added, subjected to vortex thoroughly, centrifuged for 1 minute, the supernatant was discarded, and the above operation was repeated for 3 times. 500 µL of sterilized glycerol (50%) was added, subjected to vortex thoroughly to prepare a gold powder suspension with a concentration of 60 µg/µl, which was stored at -20°C.
- 45 (3) 50µl of gold powder suspension (60µg/ml) was taken, added with 100µL of the prepared RNP complex, and mixed gently.
- 50 (4) 15 µl of RNP/gold powder mixture was taken and placed in the center of a cleavable membrane of PDS-1000 benchtop gene gun (Bio-Rad), blow-dried, and then bombarded according to the instructions of the instrument. Bombardment parameters: the vacuum degree was 26-28, the distance was 6 cm, and the air pressure was 1100 psi or 1350 psi.
- 55 (5) After the bombardment was completed, the callus was cultured on a hypertonic medium at 25°C overnight (for 16 hours) in the dark.

## 4. Selection, differentiation and rooting:

**[0238]**

- 5 (1) After the bombarded callus was cultured overnight, it was transferred to the induction medium, and cultured to recover at 28°C for one week.
- 10 (2) After one week of recovery, the callus was transferred to a selection medium (formulation: 4.1g/L N6 powder + 0.3 g/L hydrolyzed casein + 2.8 g/L proline + 2 mg/L 2,4-D + 3% sucrose + 50 ug/L quizalofop-p + 500 mg/L Cef (cephalosporin) + 0.1 g/L inositol + 0.35% phytigel, pH5.8) to perform selection for 3-4 weeks; for the prospective editing of OsACCase2 W2038, 50 ug/L quizalofop-p was used for selection, as shown in Figure 15.
- 15 (3) The selected resistant callus with good growth status was transferred to a differentiation medium (formulation: MS powder (4.42g/L) + KT (1mg/L) + sucrose (30g/L) + phytigel (4.5g/L) + 50 ug/L quizalofop-p, pH 5.8), and cultured at 28°C under light to induce differentiation for 2-4 weeks.
- 20 (4) The differentiated seedlings were transferred to a rooting medium (formulation: 1/2 MS powder (2.3g/L) + sucrose (30g/L) + phytigel (4.5g/L)) for rooting, the seedlings were tempered after the rooting was completed, and then transferred to pots filled with soil and placed in a greenhouse for cultivation.

## 5. Detection of target editing events in resistant callus and T0 tissue culture seedlings:

**[0239]** A total of 11 resistant calli were obtained in the selection, the DNA thereof was extracted by the CTAB method. The detection primers for the target site were designed as follows: OsACCase2038test-F: 5'CTGTAGGCATTGAAACTGCAGTG3', OsACCase2038test-R: 5'GCAATCCTGGAGTTCCCT-CTGACC3', and the PCR fragment containing OsACCase2 W2038 site was amplified, recovered and sequenced. The sequencing detection indicated that 10 out of them had the mutation from TGG to TTG at the OsACCase2 W2038 site, in which 3 samples were homozygous mutants.

**[0240]** For the T0 generation tissue culture seedlings obtained by resistant callus differentiation, the DNA thereof was extracted for detecting the editing target site sequence, and there was also a homozygous mutation from TGG to TTG at the OsACCase2 W2038 site, as shown in Figure 16.

## 6. Resistance test of T1 generation seedlings to ACCase inhibitor herbicides

**[0241]** After the propagation of the T0 strain containing the W2038L mutation, the T1 generation mutant seedlings were tested for herbicide resistance with quizalofop-p and haloxyfop-p in field concentrations. It could be seen that the OsACCase2 W2038L mutant strain was significantly resistant to these two ACCase inhibitor herbicides, as shown in Figures 17-18.

**[0242]** In summary, by either using the gRNA prepared with the synthetic crRNA and tracrRNA by annealing, or directly using the synthetic sgRNA, an active RNP complex could be formed with the purified SpCas9 protein; the gene gun bombarded calli could be selected by using quizalofop-p in the tissue culture stage; the homozygous mutation from TGG to TTG could be detected by sequencing the OsACCase2 W2038 target site of the T0 generation tissue culture seedlings; the mutation could be inherited to the T1 generation and showed resistance to ACCase inhibitor herbicides, which further demonstrated that the programmed sequential cutting/editing generated from the RNP complex in combination with the sequential targeting crRNA or sgRNA could achieve site-specific mutation of a target site in the cell, and could guide the production of cell-endogenous selection markers for tissue culture selection, thereby creating herbicide-resistant crops.

Example 10: Different targeting RNP complexes simultaneously edit OsACCase2 W2038 site and OsBADH2 gene of *Oryza sativa* callus

**[0243]** The RNP complex was prepared by the method according to Example 8, the gene gun bombardment and tissue culture procedures were the same as those of Example 9. Besides the crRNA or sgRNA targeting the OsACCase2 W2038 site, the crRNA or sgRNA targeting the OsBADH2 gene was added at the same time, and they were incubated with SpCas9 protein to form a targeting RNP complex for a second target gene OsBADH2. The gene gun bombardment was performed, the T0 generation tissue culture seedlings were obtained after recovery culture, screening, differentiation and rooting, and the T1 generation was obtained by propagation.

**[0244]** The *Oryza sativa* OsBADH2 genome sequence was shown in SEQ ID NO: 17. According to the CRISPOR

online tool (<http://crispor.tefor.net/>), the target site sequence CCAAGTACCTCCGCGCAATCG<sup>cg</sup> was selected, in which the PAM site recognized by the Cas9 protein was shown in italic lowercase, and the purified SpCas9 protein was selected to prepare an RNP complex. The CGG was used as PAM to design >CrRNA1-OsBADH2: 5'-CCAAGUACCUCCGCG-CAAUCGguuuuagagcuugcu-3', it was predicted that the resistant mutation of OsACCase2 W2038L and the knockout mutation event of OsBADH2 could be simultaneously detected in the resistant callus obtained by the selection of quizalofop-p.

**[0245]** >CrRNA1-OsBADH2 was synthesized by GenScript Biotechnology Company, and sgRNA was also synthesized:

>OsBADH2-sgRNA: 5'CCAAGUACCUCCGCGAAUCGguuuuagagcuagaaauagcaaguuaa  
aaauaaggcuaguccguuaucuacuugaaaaaguggcaccgagucggugc3'.

**[0246]** The resistant callus was selected according to the transformation steps in Example 9, as shown in Figure 19, and the resistant callus was selected and subjected to differentiation and seedling emergence, and the sequences at the OsACCase2 and OsBADH2 target sites of the callus and T0 generation tissue culture seedlings were sequenced. The OsBADH2 target detection primers were:

OsBADH2-check F: 5'CATCGGTACCCTCCTCTTC3'  
OsBADH2-check R: 5'ATCGATCGATTGGGGCTCA3'

**[0247]** As a result, a total of 13 resistant calli were obtained by selection, the OsACCase2 W2038L mutation event was detected in 11 out of them, and the detection of OsBADH2 target sequence for these 11 callus samples showed that 8 out of them simultaneously contained the editing event at OsBADH2 target site. In the T0 generation tissue culture seedlings obtained by differentiation, the existence of OsACCase2 W2038L mutation and OsBADH2 +A homozygous mutation was detected, as shown in Figures 20-21.

**[0248]** In sum, after the rice callus was subjected to gene gun bombardment with the sequential targeting crRNA- or sgRNA-targeting RNP complex to perform site-specific editing of OsACCase2 W2038 and simultaneously adding the targeting RNP complex targeting the second target gene OsBADH2, the selection of the callus could be performed in the tissue culture stage with quizalofop-p; the OsACCase2 W2038 mutation and the targeted knockout of OsBADH2 simultaneously occurred in 61% of the resistant callus, and the strains containing homozygous mutation of OsBADH2 were detected in the T0 generation tissue culture seedlings. These indicated that the sequential targeting in combination with the RNP transformation for site-specific editing of resistance genes could generate endogenous selection markers, and the addition of the corresponding selection pressure could simultaneously screen the editing events of the second target gene, thereby achieving the site-specific editing of genome by non-transgenic means.

Example 11: Different targeted RNP complexes simultaneously edit OsALS548 site and OsSWEET14 gene of *Oryza sativa* callus

**[0249]** The targeting RNP complexes for the OsALS548 site were prepared by the method according to Example 8, in which the crRNA and sgRNA sequences were the same as in Example 8, respectively:

>CrRNA1-548-G: 5'-GGGUUAUGGUUGUGCAAUGGGAguuuuagagcuugcu-3',  
>CrRNA2-548+T: 5'-UGGGUAUGGUUGUGCAAUGGGAguuuuagagcuugcu-3',  
>sgRNA1-548-G:

5'GGGUUAUGGUUGUGCAAUGGGAguuuuagagcuagaaauagcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3',

>sgRNA2-548+T:

5'UGGUUAUGGUUGCAAUAGGAguuuuagagcuagaaaaugcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'.

5 [0250] Referring to Example 8, the gRNA or sgRNA was incubated with NGA-Cas9 to prepare RNP complexes targeting OsALS548 site.

10 [0251] In addition to the targeted RNP complex for the OsALS548 site, the crRNA or sgRNA for the OsSWEET14 gene was added at the same time, and the targeting RNP complex for the second target gene OsSWEET14 was formed by incubation with the SpCas9 protein. The T0 generation tissue culture seedlings were obtained by performing gene gun bombardment, recovery culture, selection, differentiation, rooting, and the T1 generation was obtained by propagation. The selection pressure was 5 mg/L pyroxsulam.

15 [0252] The *Oryza sativa* OsSWEET14 genome sequence was shown in SEQ ID NO: 18. According to the CRISPOR online tool (<http://crispor.tefor.net/>), the target site sequence GAGCTTAGCACCTGGTGGAGggg was selected, in which the PAM sites recognized by the SpCas9 protein was shown in italic lowercase, and the purified SpCas9 protein was selected to prepare the RNP complex. The GGG was used as PAM to design:

>CrRNA1-OsSWEET14:

20 5'-GAGCUUAGCACCUGGUUGGAGguuuuagagcuaugcu-3', and it was predicted that the resistant mutation of OsALS W548L and the knockout mutation of OsSWEET14 could be simultaneously detected in the resistant callus obtained by the selection of pyroxsulam.

25 [0253] >CrRNA1-Os SWEET14 was synthesized by GenScript Biotechnology Company, and the sgRNA was also synthesized:

>Os SWEET14-sgRNA:

25 5'GAGCUUAGCACCUGGUUGGAGguuuuagagcuagaaaaugcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'.

30 [0254] The resistant callus was selected according to the gene gun bombardment and tissue culture procedures as described in Example 9, as shown in Figure 22, and the resistant callus was selected to perform differentiation and seedling emergence, in which the selection medium formulation was: 4.1g/L N6 powder + 0.3 g/L hydrolyzed casein + 2.8 g/L proline + 2 mg/L 2,4-D + 3% sucrose + 5 mg/L pyroxsulam + 500 mg/L Cef (cephalosporin) + 0.1 g/L inositol + 0.35% phytigel, pH 5.8.

35 [0255] The sequences at the OsALS548 site and OsSWEET14 target site of the callus and T0 generation tissue culture seedlings were sequenced. The OsSWEET14 target site detection primers were:

OsSWEET14-check F: 5' ATGGGTGCTGATGATTATCTTGAT3'

40 OsSWEET14-check R: 5' TGAAGAGACATGCCAGCCATTG3'

45 [0256] As a result, a total of 9 resistant calli were obtained by the selection, the OsALS W548L mutation event was detected in 8 out of them, and the detection of OsSWEET14 target sequence in these 8 callus samples showed that 5 out of them also contained the editing event at OsSWEET14 target site. In the T0 generation tissue culture seedlings obtained by differentiation, both the OsALS W548L mutation and the OsSWEET14-C homozygous mutation could be detected, as shown in Figures 23-24.

50 [0257] After the propagation of T0 strains in which the occurrence of OsALS W548L mutation was detected, the T1 generation mutant seedling strains were tested for herbicide resistance with pyroxsulam, imazapic, nicosulfuron and flucarbazone-Na at field concentrations. It could be seen that the OsALS W548L mutant strain showed significant resistance to all of these 4 ALS inhibitor herbicides, as shown in Figures 25-28.

55 [0258] In summary, by either using the gRNA prepared with the synthetic crRNA and tracrRNA by annealing or directly using the synthetic sgRNA, an active RNP complex could be formed with the purified NGA Cas9 protein, the site-specific mutation of a target site in the cells could be achieved by the programmed sequential cutting/editing generated from the RNP complex in combination with the sequential targeting crRNA or sgRNA, the selection of the callus bombarded by gene gun could be performed with pyroxsulam at the tissue culture stage, the mutation from TGG to TTG was detected by sequencing the OsALS548 target site of the T0 generation tissue culture seedlings, and this mutation could be inherited to the T1 generation and showed resistance to ALS inhibitor herbicides.

[0259] After the rice callus was subjected to gene gun bombardment by simultaneously adding the targeting RNP

complex targeting the second target gene OsSWEET14 and using the SpCas9 protein that recognized NGG PAM, the selection of the callus was performed with pyroxslam at the tissue culture stage. The OsALS W548L mutation and the targeted knockout of OsSWEET14 occurred simultaneously in 55% of the resistant callus, the occurrence of homozygous mutation of OsSWEET14 could be detected in the T0 generation tissue culture seedlings, which further indicated that the endogenous selection markers could be generated by the programmed sequential cutting/editing in combination with the site-specific editing of resistant genes generated from PNP transformation, and the editing events of the second target gene could be selected at the same time by simultaneously using the Cas9 proteins that recognized different PAM sites and adding a corresponding selection pressure, thereby achieving the site-specific editing of genome by non-transgenic means.

Example 12: Generation of W561L mutation at StALS561 site by bombarding potato (*Solanum tuberosum L.*) explants with two different targeted RNP complexes

[0260] The amino acid sequence of *Solanum tuberosum L.* StALS2 protein was shown in SEQ ID NO: 19, and the sequence of *Solanum tuberosum L.* StALS2 gene was shown in SEQ ID NO: 20. The methods for the preparation of RNP complexes and the gene gun bombardment referred to Examples 8-9, and the sgRNAs for the original sequence and the edited sequence designed for StALS2W561 site of *Solanum tuberosum L.* StALS2 corresponding to the *Arabidopsis thaliana* ALS574 site were as follows:

>StALS561- G:

5'GGGAAUGGUGGUUCAGUGGGAguuuuagagcuagaaauagcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3'

>StALS561 +T:

5'UGGGAAUGGUGGUUCAGUGGAguuuuagagcuagaaauagcaaguuaaaaauaaggcuaguccguua  
ucaacuugaaaaaguggcaccgagucggugc3',

[0261] And the RNP complexes were prepared according to the method described in Example 8.

[0262] The recipient potato variety was Atlantic or Favorita, and the leaves, stems and axillary buds thereof were used as explants, respectively. The methods for gene gun bombardment and selection and differentiation were as follows:

(1) Leaves: The whole leaf was cut and spread flat on the hypertonic M6 medium (formulation: 4.42g/L MS powder + 1 ml/L B5 vitamins (Phytotechlab, G219) + 30g/L sucrose + 8g/L agar + 2mg/L 2,4-D + 0.8mg/L zeatin riboside + 0.2M mannitol + 250mg/L Cef), about 5-6 leaves were used for each shot of bombardment. After pre-incubation for 24 hours, the bombardment with gold powder was performed. After bombardment, they were continuously cultured in hypertonic M6 medium in dark for 2 days, then transferred to M6 medium (4.42g/L MS powder+1 ml/L B5 vitamins (Phytotechlab, G219) + 30g/L sucrose + 8g/L agar + 2mg/L 2,4-D + 0.8mg/L zeatin riboside + 250mg/L Cef) and continuously cultured for 1 week, then transferred to M6 medium with selection pressure after 1 week, the selection of resistant callus was performed with the selection pressure of 20 $\mu$ g/L chlorsulfuron (formulation: 4.42g/L MS powder + 1 ml/L B5 vitamins (Phytotechlab, G219) + 30g/L sucrose + 8g/L agar + 2mg/L 2,4-D + 0.8mg/L zeatin riboside + 250mg/L Cef + 20 $\mu$ g/L chlorsulfuron), 4 weeks later, they were transferred to an bud induction medium R4 containing selection pressure (formulation: 4.42g/L MS powder + 1 ml/L B5 vitamins (Phytotechlab, G219) + 30g/L sucrose + 8g/L agar + 2mg/L GA3 + 0.8mg/L zeatin riboside + 250mg/L Cef + 20 $\mu$ g/L chlorsulfuron) until seedling emergence.

(2) Stems: The potato stems (without axillary buds) were taken and cut longitudinally with the incisions facing upwards, and placed on CIMI medium (4.42g/L MS powder + 20g/L sucrose + 8g/L agar + 0.5 mg/L zeatin nucleoside + 2mg/L 2,4-D + 250mg/L Cef + 20 $\mu$ g/L chlorsulfuron), followed by gene gun bombardment, cultured in the dark for 1 day after bombardment, the stems were transferred to fresh CIMI medium and continuously cultured for 1 week, then the stems were transferred to SIMI medium containing selection pressure (formulation: 4.42g/L MS powder + 20g/L sucrose + 8g/L agar + 1mg/L zeatin riboside + 0.1mg/L GA3 + 250mg/L Cef + 20 $\mu$ g/L chlorsulfuron) to perform the selection for resistant buds until seedling emergence.

(3) Axillary buds: The axillary buds were taken from potato stems, cut longitudinally, and transferred to CIMI medium, with the incisions facing upwards, placed on CIMI medium, followed by gene gun bombardment, cultured in the dark for 1 day after bombardment, they were transferred to fresh CIMI medium and continuously cultured for 1 week, and then they were transferred to SIMI medium containing selection pressure to perform the selection for resistant buds until seedling emergence.

5 [0263] The detection primers for StALS2 W561 site were:

10 StALS561-Check F: 5'GTGGATTAGGAGCAATGGGATT3'

15 StALS561-Check R: 5' TTATTTAGATAATAACATGCCTCG3'

[0264] After detection, the editing event of W561L occurred at the StALS2 W561 site of the resistance-screened T0 generation potato tissue culture seedlings, which demonstrated that the non-transgenic transient gene editing method provided by the present invention was suitable for crops such as potatoes from which an exogenous transgenic element can hardly be separated and removed by selfing or hybridization.

Example 13: Successful use of programmed sequential cutting/editing scheme for base substitution in human 293T cells

20 [0265] The HBB (hemoglobin subunit beta) gene in human embryonic kidney cell 293T (the DNA sequence thereof was shown in SEQ ID NO: 21, the CDS sequence thereof was shown in SEQ ID NO: 22, and the amino acid sequence thereof was shown in SEQ ID NO: 23) was selected, the target site for sequential targeting of sgRNA was designed in the region of the first exon, in which the first target was catggcactgtcgAGG. The sgRNA that recognized this target was named sgHBB, and it was predicted that the deletion of one C base could be generated at the sgRNA cut of this site. The second target was ccatggcactgtcgAGG, which recognized the sequence with the deletion of one C base generated from the cutting/editing of the first target, and the sgRNA of this target was named sgHBB-c. The sgRNA with no target site in 293T cells was designed and named sgNOTAR, which was used as a complementing plasmid for transfection in the experiment.

25 [0266] According to the above design, complementary single-stranded DNA fragments were synthesized respectively. After annealing, they were ligated into px458 (addgene: 48138) plasmids digested with BbsI enzyme, and transformed to *E. coli* DH5a competent. After the resultant *E. coli* single colonies were verified by sequencing, the plasmids were extracted and purified with an endotoxin-free plasmid extraction kit (Tiangen Bio).

30 [0267] The vigorously growing 293T cells were digested and isolated with 0.05% trypsin (Gibco), diluted with DMEM medium (10% fetal bovine serum; penicillin + streptomycin double resistant) and inoculated into 24-well culture plates, and placed in a carbon dioxide incubator overnight. On the next day, they were mixed separately with, according to sequential cutting/editing: sgHBB and sgHBB-c plasmids each 0.5ug; single target cutting/editing: sgHBB and sgNOTAR plasmids each 0.5ug; no target control: pEGFP-c1 plasmid 1μg. The transformation was performed with lipofectamine3000 (Invitrogen). There were 3 duplications for each group.

35 [0268] 48 hours After transformation, pictures were taken with a fluorescence microscope to record the transformation efficiency, and the total DNA of each well was extracted with a nucleic acid extraction kit (Omega).

40 [0269] The designed Hi-tom sequencing primers were as follows:

Hi-HBB-F: gaggtgagttacgggtgcGCTTACATTGCTTCTGACACAAC;

45 Hi-HBB-R: gagttggatgtggatggTCTATTGGTCTCCTAAACCTGTCTG.

[0270] These primers were used to perform PCR for each DNA sample. The high-throughput sequencing of PCR products was performed by using Hi-tom method (Sci China Life Sci. 2019 Jan;62(1):1-7. doi: 10.1007/s11427-018-9402-9).

50 [0271] The statistical data of the sequencing results were shown in Table 6 and Table 7. These data indicated that the sequential cutting/editing method at HBB site produced an editing outcome from C to T (resulting in P6S mutation) at a ratio of about 1.67%, while the single target cutting/editing method could not produce such base substitution.

Table 6: Editing types produced by different sgRNA combinations in the sequential cutting/editing experiment of HBB gene in 293T cells, and the readings of different genotypes detected by Hi-Tom sequencing

Genotype	HBB programmed sequential cutting/editing repeat 1	HBB programmed sequential cutting/editing repeat 2	HBB programmed sequential cutting/editing repeat 3	HBB single target cutting/editing repeat 1	HBB single target cutting/editing repeat 2	HBB single target cutting/editing repeat 3	GFP transformation repeat 1	GFP transformation repeat 2
WT	6025	6037	6003	5776	5636	5614	7564	7430
1Deletion	555	593	613	1170	1134	1120	252	288
2Deletion	594	625	410	465	509	405		
3Deletion	304	263	258		260	243		
4Deletion		222						
5Deletion			220		217			
6Deletion								
7Deletion	312	223	232	274	229	278		
Insertion				273		308		
C->T SNP	168		230					
total	7958	7963	7966	7958	7985	7968	7816	7718

Table 7: Combined statistics of ratios of editing types generated by different sgRNA combinations in the sequential cutting/editing experiment of HBB genes in 293T cells

Genotype	HBB programmed sequential cutting/editing	HBB single target cutting/editing	GFP transformation background	HBB programmed sequential cutting/editing ratio	HBB single target cutting/editing ratio	GFP transformation background ratio
WT	18065	17026	14994	75.63%	71.21%	96.52%
1Deletion	1761	3424	540	7.37%	14.32%	3.48%
2Deletion	1629	1379	0	6.82%	5.77%	0.00%
3Deletion	825	503	0	3.45%	2.10%	0.00%
4Deletion	222	0	0	0.93%	0.00%	0.00%
5Deletion	220	217	0	0.92%	0.91%	0.00%

6Deletion	0	0	0	0.00%	0.00%	0.00%
7Deletion	767	781	0	3.21%	3.27%	0.00%
Insertion	0	581	0	0.00%	2.43%	0.00%
C->T SNP	398	0	0	1.67%	0.00%	0.00%
total	23887	23911	15534	100.00%	100.00%	100.00%

Note: WT represents wild type; Deletion represents genotype of deletion; Insertion represents genotype of insertion; C->T SNP represents genotype with base substitution from C to T at the cut; total represents grand total.

[0272] In addition, sickle cell anemia and  $\beta$ -thalassemia are both inherited anemias caused by mutations in the HBB gene encoding hemoglobin  $\beta$  subunit of adult. Patients with these diseases may need blood transfusion or other therapies throughout their lives. The results of the above experiment demonstrated that through the technical solution of programmed sequential cutting/editing as provided by the present invention, a combination of crRNA or sgRNA could be designed for the mutation site of the HBB gene to induce the expected repair at the mutation site, so that the cell produced active hemoglobin and restored function, thereby achieving therapeutic effect. That is, the composition provided by the present invention has a use in treatment of diseases.

[0273] After a variety of tests at the same time, because this novel method is completely based on the existing functions of Cas9, the method of the present invention will be fully applicable to achieve the new functions of base substitution, deletion and insertion of specific fragments in other organisms (plants, animals, fungi or bacteria, etc.) where Cas9 can work well.

[0274] All publications and patent applications mentioned in the description are incorporated herein by reference, as if each publication or patent application is individually and specifically incorporated herein by reference.

[0275] Although the aforementioned invention has been described in more details by way of examples and embodiments for clear understanding, it is obvious that certain changes and modifications can be implemented within the scope of the appended claims, and such changes and modifications are all within the scope of the present invention.

35

40

45

50

55

## Sequence Listing

<110> Qingdao Kingagroot Chemical Compound Co., Ltd.

5 <120> A method for generating new mutations in organisms and use thereof

<130> IEC200334

10 <150> 201911081617X  
<151> 2019-11-07

15 <150> 2020108218772  
<151> 2020-08-15

<150> 2020109741512  
<151> 2020-09-16

20 <160> 23

<170> PatentIn version 3.5

25 <210> 1  
<211> 670  
<212> PRT  
<213> Arabidopsis thaliana ALS protein amino acid sequence (amino acid sequence)

30 <400> 1  
Met Ala Ala Ala Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe  
1 5 10 15  
Ser Thr Lys Pro Ser Pro Ser Ser Lys Ser Pro Leu Pro Ile Ser  
20 25 30  
Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser  
35 40 45  
Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala  
50 55 60  
Val Leu Asn Thr Thr Asn Val Thr Thr Pro Ser Pro Thr Lys  
65 70 75 80  
35 Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro  
85 90 95  
Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val  
100 105 110  
Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln  
115 120 125  
40 Ala Leu Thr Arg Ser Ser Ser Ile Arg Asn Val Leu Pro Arg His Glu  
130 135 140  
Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ser Ser Gly Lys  
145 150 155 160  
Pro Gly Ile Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val  
165 170 175  
45 Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Leu Val Ala Ile  
180 185 190  
Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln Glu  
195 200 205  
50 Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His Asn Tyr Leu  
210 215 220  
Val Met Asp Val Glu Asp Ile Pro Arg Ile Ile Glu Glu Ala Phe Phe  
225 230 235 240  
Leu Ala Thr Ser Gly Arg Pro Gly Pro Val Leu Val Asp Val Pro Lys  
245 250 255  
55 Asp Ile Gln Gln Leu Ala Ile Pro Asn Trp Glu Gln Ala Met Arg  
260 265 270  
Leu Pro Gly Tyr Met Ser Arg Met Pro Lys Pro Pro Glu Asp Ser His

	275	280	285
5	Leu Glu Gln Ile Val Arg	Leu Ile Ser Glu Ser Lys	Lys Pro Val Leu
	290	295	300
	Tyr Val Gly Gly Gly	Cys Leu Asn Ser Ser Asp	Glu Leu Gly Arg Phe
	305	310	315
	Val Glu Leu Thr Gly Ile Pro Val Ala	Ser Thr Leu Met Gly	Leu Gly
	325	330	335
	Ser Tyr Pro Cys Asp Asp Glu	Leu Ser Leu His Met	Leu Gly Met His
	340	345	350
10	Gly Thr Val Tyr Ala Asn	Tyr Ala Val Glu His	Ser Asp Leu Leu Leu
	355	360	365
	Ala Phe Gly Val Arg Phe	Asp Asp Arg Val Thr	Gly Lys Leu Glu Ala
	370	375	380
	Phe Ala Ser Arg Ala Lys	Ile Val His Ile Asp	Ile Asp Ser Ala Glu
	385	390	395
15	Ile Gly Lys Asn Lys	Thr Pro His Val Ser	Val Cys Gly Asp Val Lys
	405	410	415
	Leu Ala Leu Gln Gly Met Asn Lys	Val Leu Glu Asn Arg	Ala Glu Glu
	420	425	430
	Leu Lys Leu Asp Phe Gly Val	Trp Arg Asn Glu	Leu Asn Val Gln Lys
	435	440	445
20	Gln Lys Phe Pro Leu Ser Phe	Lys Thr Phe Gly	Glu Ala Ile Pro Pro
	450	455	460
	Gln Tyr Ala Ile Lys Val	Leu Asp Glu Leu	Thr Asp Gly Lys Ala Ile
	465	470	475
	Ile Ser Thr Gly Val Gly Gln His	Gln Met Trp Ala Ala	Gln Phe Tyr
	485	490	495
25	Asn Tyr Lys Lys Pro Arg Gln	Trp Leu Ser Ser Gly	Gly Leu Gly Ala
	500	505	510
	Met Gly Phe Gly Leu Pro Ala	Ala Ile Gly Ala Ser	Val Ala Asn Pro
	515	520	525
	Asp Ala Ile Val Val Asp	Ile Asp Gly Asp Gly	Ser Phe Ile Met Asn
	530	535	540
30	Val Gln Glu Leu Ala Thr	Ile Arg Val Glu Asn	Leu Pro Val Lys Val
	545	550	555
	Leu Leu Leu Asn Asn Gln His	Leu Gly Met Val	Met Gln Trp Glu Asp
	565	570	575
	Arg Phe Tyr Lys Ala Asn Arg	Ala His Thr Phe	Leu Gly Asp Pro Ala
35	580	585	590
	Gln Glu Asp Glu Ile Phe Pro	Asn Met Leu Leu Phe	Ala Ala Ala Cys
	595	600	605
	Gly Ile Pro Ala Ala Arg Val	Thr Lys Lys Ala Asp	Leu Arg Glu Ala
	610	615	620
40	Ile Gln Thr Met Leu Asp	Thr Pro Gly Pro	Tyr Leu Leu Asp Val Ile
	625	630	635
	Cys Pro His Gln Glu His	Val Leu Pro Met	Ile Pro Ser Gly Gly Thr
	645	650	655
	Phe Asn Asp Val Ile Thr	Glu Gly Asp Gly	Arg Ile Lys Tyr
	660	665	670
45	<210> 2		
	<211> 2013		
	<212> DNA		
	<213> Arabidopsis thaliana ALS gene DNA sequence (DNA sequence)		
50	<400> 2		
	atggcggcgg caacaacaac aacaacaaca ttttcttcga ttccttcgtc caccaaacca		60
	tctccttcgt cctccaaatc accattacca atctccagat ttcctccccc attctcccta		120
	aaccccaaca aatcatcctc ctccctccgc cggcgccgtt tcaaattccag ctctccctcc		180
	tccatctccg ccgtgctcaa cacaaccacc aatgtcacaa ccactccctc tccaacccaa		240
	cctacccaaac ccgaaacatt catctcccgta ttccgtccag atcaaccccg caaaggcgct		300
55	gatatctcg tggaaagcttt agaacgtcaa ggcttagaaa ccgttatccg ttaccctgg		360
	ggtgcatcaa tggagattca ccaaggctta acccgcttt cctcaatccg taacgtcctt		420

	cctcgtaacg aacaaggagg tgtattcgca gcagaaggat acgctcgatc ctcaggtaaa	480
	ccaggtatct gtatagccac ttccaggccc ggagctacaa atctcgtagt cgattagcc	540
	gatgcgttgt tagatagtgt tcctcttgc gcaatcacag gacaagtccc tcgtcgatg	600
5	attggtagac agagactccg attgttgagg taacgcgttc gattacgaag	660
	cataactatc ttgtgatgga ttttgaagat atcccttagga ttatttgagga agctttctt	720
	ttagctactt ctggtagacc tggacctgtt ttgttgatg ttccctaaaga tattcaacaa	780
	cagcttgcga ttccataattt ggaacaggct atgagattac ctggtttatgt gtctaggatg	840
10	cctaaacctc cggaaagattc tcatttggag cagattgtt ggttggatc tgagtctaag	900
	aaggctgtgt tttatgttgg tgggtgtt ggttggatc gcatgaaatt ggttaggtt	960
	gtttagctt cggggatccc ttttgcgagt acgttgcgtt ggctgggatc ttatccttgc	1020
	gatgtatgt tttatgttaca tatgttttggaa atgcatgggatc ctgtgtatgc aaattacgc	1080
	gtggagcata gtgattttgtt gttggcggtt gggtaaggat ttgtatgtatgc tgacgttgc	1140
15	aaggcttgggg cttttgttagt tagggctaaatg attgttgcata ttgtatatttgc ttcggctgag	1200
	attgggaaga ataagactcc tcatgtgtct gtgtgtggatc atgttaagat ggctttgc	1260
	gggatgaata aggttcttgc gaaaccggacg gaggagctt agcttgcattt tggagtttgg	1320
	aggaatgagt tttatgttaca gaaacagaag tttccgttgc gctttaagac gtttggggaa	1380
20	gctattccctc cacagtatgc gattaaggatc ctttgcgttgc tttatgttgc gaaacggata	1440
	ataagtactg gtgtcgccca acatcaaatttggccggcgc agtttgcata ttacaagaaaa	1500
	ccaaggccatc ggctatcatc aggaggccctt ggagctatgg gatttggact tcctgctgc	1560
	attggagcgt ctgttgcata ccctgtatgc atagttgtgg atattgcacgg agatggaa	1620
25	tttataatgc atgtgcataa gcttgcacttgc tttatgttgc agaatcttcc agtgcaggta	1680
	cttttataa acaaccagca ttttgcacttgc gttatgttgc ggttgcataaaa	1740
	gcttaccggc ttttgcacttgc ttttgcacttgc cccgttgcagg aggacgcgat attccgc	1800
	atgttgcgtt ttgcacttgc ttttgcacttgc ccagccggcga ggttgcataaaa gaaagcagat	1860
	ctccggaaatc ttttgcacttgc aatgttgcgtt acaccaggac ctttgcacttgc ggttgcgtt	1920
	tttttgcacttgc aagaacatgt gtttgcacttgc atcccgacttgc gtttgcacttgc caacgcgtt	1980
	ataacggaaatc ggttgcacttgc gtttgcacttgc ttttgcacttgc ttttgcacttgc ttttgcacttgc	2013

25                   <210> 3  
 <211> 2320  
 <212> PRT  
 <213> Alopecurus myosuroides ACCase protein amino acid sequence (amino acid  
 30                   sequence)

	<400> 3
	Met Gly Ser Thr His Leu Pro Ile Val Gly Phe Asn Ala Ser Thr Thr
	1                         5                         10                         15
	Pro Ser Leu Ser Thr Leu Arg Gln Ile Asn Ser Ala Ala Ala Phe
35	20                         25                         30
	Gln Ser Ser Ser Pro Ser Arg Ser Ser Lys Lys Lys Ser Arg Arg Val
	35                         40                         45
	Lys Ser Ile Arg Asp Asp Gly Asp Gly Ser Val Pro Asp Pro Ala Gly
	50                         55                         60
40	His Gly Gln Ser Ile Arg Gln Gly Leu Ala Gly Ile Ile Asp Leu Pro
	65                         70                         75                         80
	Lys Glu Gly Ala Ser Ala Pro Asp Val Asp Ile Ser His Gly Ser Glu
	85                         90                         95
	Asp His Lys Ala Ser Tyr Gln Met Asn Gly Ile Leu Asn Glu Ser His
	100                         105                         110
45	Asn Gly Arg His Ala Ser Leu Ser Lys Val Tyr Glu Phe Cys Thr Glu
	115                         120                         125
	Leu Gly Gly Lys Thr Pro Ile His Ser Val Leu Val Ala Asn Asn Gly
	130                         135                         140
	Met Ala Ala Ala Lys Phe Met Arg Ser Val Arg Thr Trp Ala Asn Asp
50	145                         150                         155                         160
	Thr Phe Gly Ser Glu Lys Ala Ile Gln Leu Ile Ala Met Ala Thr Pro
	165                         170                         175
	Glu Asp Met Arg Ile Asn Ala Glu His Ile Arg Ile Ala Asp Gln Phe
	180                         185                         190
55	Val Glu Val Pro Gly Gly Thr Asn Asn Asn Asn Tyr Ala Asn Val Gln
	195                         200                         205
	Leu Ile Val Glu Ile Ala Glu Arg Thr Gly Val Ser Ala Val Trp Pro
	210                         215                         220

## EP 3 889 266 A1

	Gly Trp Gly His Ala Ser Glu Asn Pro Glu Leu Pro Asp Ala Leu Thr			
225	230	235	240	
	Ala Lys Gly Ile Val Phe Leu Gly Pro Pro Ala Ser Ser Met Asn Ala			
	245	250	255	
5	Leu Gly Asp Lys Val Gly Ser Ala Leu Ile Ala Gln Ala Ala Gly Val			
	260	265	270	
	Pro Thr Leu Ala Trp Ser Gly Ser His Val Glu Ile Pro Leu Glu Leu			
	275	280	285	
10	Cys Leu Asp Ser Ile Pro Glu Glu Met Tyr Arg Lys Ala Cys Val Thr			
	290	295	300	
	Thr Ala Asp Glu Ala Val Ala Ser Cys Gln Met Ile Gly Tyr Pro Ala			
	305	310	315	320
	Met Ile Lys Ala Ser Trp Gly Gly Gly Lys Gly Ile Arg Lys Val			
	325	330	335	
15	Asn Asn Asp Asp Glu Val Lys Ala Leu Phe Lys Gln Val Gln Gly Glu			
	340	345	350	
	Val Pro Gly Ser Pro Ile Phe Ile Met Arg Leu Ala Ser Gln Ser Arg			
	355	360	365	
	His Leu Glu Val Gln Leu Leu Cys Asp Glu Tyr Gly Asn Val Ala Ala			
	370	375	380	
20	Leu His Ser Arg Asp Cys Ser Val Gln Arg Arg His Gln Lys Ile Ile			
	385	390	395	400
	Glu Glu Gly Pro Val Thr Val Ala Pro Arg Glu Thr Val Lys Glu Leu			
	405	410	415	
	Glu Gln Ala Ala Arg Arg Leu Ala Lys Ala Val Gly Tyr Val Gly Ala			
	420	425	430	
25	Ala Thr Val Glu Tyr Leu Tyr Ser Met Glu Thr Gly Glu Tyr Tyr Phe			
	435	440	445	
	Leu Glu Leu Asn Pro Arg Leu Gln Val Glu His Pro Val Thr Glu Ser			
	450	455	460	
	Ile Ala Glu Val Asn Leu Pro Ala Ala Gln Val Ala Val Gly Met Gly			
	465	470	475	480
30	Ile Pro Leu Trp Gln Ile Pro Glu Ile Arg Arg Phe Tyr Gly Met Asp			
	485	490	495	
	Asn Gly Gly Tyr Asp Ile Trp Arg Lys Thr Ala Ala Leu Ala Thr			
	500	505	510	
	Pro Phe Asn Phe Asp Glu Val Asp Ser Gln Trp Pro Lys Gly His Cys			
	515	520	525	
35	Val Ala Val Arg Ile Thr Ser Glu Asn Pro Asp Asp Gly Phe Lys Pro			
	530	535	540	
	Thr Gly Gly Lys Val Lys Glu Ile Ser Phe Lys Ser Lys Pro Asn Val			
	545	550	555	560
	Trp Gly Tyr Phe Ser Val Lys Ser Gly Gly Ile His Glu Phe Ala			
	565	570	575	
40	Asp Ser Gln Phe Gly His Val Phe Ala Tyr Gly Glu Thr Arg Ser Ala			
	580	585	590	
	Ala Ile Thr Ser Met Ser Leu Ala Leu Lys Glu Ile Gln Ile Arg Gly			
	595	600	605	
	Glu Ile His Thr Asn Val Asp Tyr Thr Val Asp Leu Leu Asn Ala Pro			
	610	615	620	
45	Asp Phe Arg Glu Asn Thr Ile His Thr Gly Trp Leu Asp Thr Arg Ile			
	625	630	635	640
	Ala Met Arg Val Gln Ala Glu Arg Pro Pro Trp Tyr Ile Ser Val Val			
	645	650	655	
50	Gly Gly Ala Leu Tyr Lys Thr Ile Thr Thr Asn Ala Glu Thr Val Ser			
	660	665	670	
	Glu Tyr Val Ser Tyr Leu Ile Lys Gly Gln Ile Pro Pro Lys His Ile			
	675	680	685	
	Ser Leu Val His Ser Thr Ile Ser Leu Asn Ile Glu Glu Ser Lys Tyr			
	690	695	700	
55	Thr Ile Glu Ile Val Arg Ser Gly Gln Gly Ser Tyr Arg Leu Arg Leu			
	705	710	715	720
	Asn Gly Ser Leu Ile Glu Ala Asn Val Gln Thr Leu Cys Asp Gly Gly			

	725	730	735
	Leu Leu Met Gln Leu Asp Gly Asn Ser His Val Ile Tyr Ala Glu Glu		
5	740	745	750
	Glu Ala Gly Gly Thr Arg Leu Leu Ile Asp Gly Lys Thr Cys Leu Leu		
	755	760	765
	Gln Asn Asp His Asp Pro Ser Arg Leu Leu Ala Glu Thr Pro Cys Lys		
	770	775	780
	Leu Leu Arg Phe Leu Ile Ala Asp Gly Ala His Val Asp Ala Asp Val		
10	785	790	800
	Pro Tyr Ala Glu Val Glu Val Met Lys Met Cys Met Pro Leu Leu Ser		
	805	810	815
	Pro Ala Ala Gly Val Ile Asn Val Leu Leu Ser Glu Gly Gln Ala Met		
	820	825	830
	Gln Ala Gly Asp Leu Ile Ala Arg Leu Asp Leu Asp Asp Pro Ser Ala		
15	835	840	845
	Val Lys Arg Ala Glu Pro Phe Glu Gly Ser Phe Pro Glu Met Ser Leu		
	850	855	860
	Pro Ile Ala Ala Ser Gly Gln Val His Lys Arg Cys Ala Ala Ser Leu		
	865	870	880
	Asn Ala Ala Arg Met Val Leu Ala Gly Tyr Asp His Ala Ala Asn Lys		
	885	890	895
20	Val Val Gln Asp Leu Val Trp Cys Leu Asp Thr Pro Ala Leu Pro Phe		
	900	905	910
	Leu Gln Trp Glu Glu Leu Met Ser Val Leu Ala Thr Arg Leu Pro Arg		
	915	920	925
	Arg Leu Lys Ser Glu Leu Glu Gly Lys Tyr Asn Glu Tyr Lys Leu Asn		
25	930	935	940
	Val Asp His Val Lys Ile Lys Asp Phe Pro Thr Glu Met Leu Arg Glu		
	945	950	960
	Thr Ile Glu Glu Asn Leu Ala Cys Val Ser Glu Lys Glu Met Val Thr		
	965	970	975
	Ile Glu Arg Leu Val Asp Pro Leu Met Ser Leu Leu Lys Ser Tyr Glu		
30	980	985	990
	Gly Gly Arg Glu Ser His Ala His Phe Ile Val Lys Ser Leu Phe Glu		
	995	1000	1005
	Glu Tyr Leu Ser Val Glu Glu Leu Phe Ser Asp Gly Ile Gln Ser Asp		
	1010	1015	1020
35	Val Ile Glu Arg Leu Arg Leu Gln Tyr Ser Lys Asp Leu Gln Lys Val		
	1025	1030	1035
	Val Asp Ile Val Leu Ser His Gln Gly Val Arg Asn Lys Thr Lys Leu		
	1045	1050	1055
	Ile Leu Ala Leu Met Glu Lys Leu Val Tyr Pro Asn Pro Ala Ala Tyr		
	1060	1065	1070
40	Arg Asp Gln Leu Ile Arg Phe Ser Ser Leu Asn His Lys Arg Tyr Tyr		
	1075	1080	1085
	Lys Leu Ala Leu Lys Ala Ser Glu Leu Leu Glu Gln Thr Lys Leu Ser		
	1090	1095	1100
	Glu Leu Arg Thr Ser Ile Ala Arg Asn Leu Ser Ala Leu Asp Met Phe		
45	1105	1110	1115
	Thr Glu Glu Lys Ala Asp Phe Ser Leu Gln Asp Arg Lys Leu Ala Ile		
	1125	1130	1135
	Asn Glu Ser Met Gly Asp Leu Val Thr Ala Pro Leu Pro Val Glu Asp		
	1140	1145	1150
	Ala Leu Val Ser Leu Phe Asp Cys Thr Asp Gln Thr Leu Gln Gln Arg		
50	1155	1160	1165
	Val Ile Gln Thr Tyr Ile Ser Arg Leu Tyr Gln Pro Gln Leu Val Lys		
	1170	1175	1180
	Asp Ser Ile Gln Leu Lys Tyr Gln Asp Ser Gly Val Ile Ala Leu Trp		
	1185	1190	1200
	Glu Phe Thr Glu Gly Asn His Glu Lys Arg Leu Gly Ala Met Val Ile		
55	1205	1210	1215
	Leu Lys Ser Leu Glu Ser Val Ser Thr Ala Ile Gly Ala Ala Leu Lys		
	1220	1225	1230

## EP 3 889 266 A1

	Asp Ala Ser His Tyr Ala Ser Ser Ala Gly Asn Thr Val His Ile Ala	
	1235 1240 1245	
	Leu Leu Asp Ala Asp Thr Gln Leu Asn Thr Thr Glu Asp Ser Gly Asp	
5	1250 1255 1260	
	Asn Asp Gln Ala Gln Asp Lys Met Asp Lys Leu Ser Phe Val Leu Lys	
	1265 1270 1275 1280	
	Gln Asp Val Val Met Ala Asp Leu Arg Ala Ala Asp Val Lys Val Val	
	1285 1290 1295	
10	Ser Cys Ile Val Gln Arg Asp Gly Ala Ile Met Pro Met Arg Arg Thr	
	1300 1305 1310	
	Phe Leu Leu Ser Glu Glu Lys Leu Cys Tyr Glu Glu Pro Ile Leu	
	1315 1320 1325	
	Arg His Val Glu Pro Pro Leu Ser Ala Leu Leu Glu Leu Asp Lys Leu	
	1330 1335 1340	
15	Lys Val Lys Gly Tyr Asn Glu Met Lys Tyr Thr Pro Ser Arg Asp Arg	
	1345 1350 1355 1360	
	Gln Trp His Ile Tyr Thr Leu Arg Asn Thr Glu Asn Pro Lys Met Leu	
	1365 1370 1375	
	His Arg Val Phe Phe Arg Thr Leu Val Arg Gln Pro Ser Ala Gly Asn	
	1380 1385 1390	
20	Arg Phe Thr Ser Asp His Ile Thr Asp Val Glu Val Gly His Ala Glu	
	1395 1400 1405	
	Glu Pro Leu Ser Phe Thr Ser Ser Ile Leu Lys Ser Leu Lys Ile	
	1410 1415 1420	
	Ala Lys Glu Glu Leu Glu Leu His Ala Ile Arg Thr Gly His Ser His	
	1425 1430 1435 1440	
25	Met Tyr Leu Cys Ile Leu Lys Glu Gln Lys Leu Leu Asp Leu Val Pro	
	1445 1450 1455	
	Val Ser Gly Asn Thr Val Val Asp Val Gly Gln Asp Glu Ala Thr Ala	
	1460 1465 1470	
	Cys Ser Leu Leu Lys Glu Met Ala Leu Lys Ile His Glu Leu Val Gly	
	1475 1480 1485	
30	Ala Arg Met His His Leu Ser Val Cys Gln Trp Glu Val Lys Leu Lys	
	1490 1495 1500	
	Leu Val Ser Asp Gly Pro Ala Ser Gly Ser Trp Arg Val Val Thr Thr	
	1505 1510 1515 1520	
	Asn Val Thr Gly His Thr Cys Thr Val Asp Ile Tyr Arg Glu Val Glu	
	1525 1530 1535	
35	Asp Thr Glu Ser Gln Lys Leu Val Tyr His Ser Thr Ala Leu Ser Ser	
	1540 1545 1550	
	Gly Pro Leu His Gly Val Ala Leu Asn Thr Ser Tyr Gln Pro Leu Ser	
	1555 1560 1565	
	Val Ile Asp Leu Lys Arg Cys Ser Ala Arg Asn Asn Lys Thr Thr Tyr	
	1570 1575 1580	
40	Cys Tyr Asp Phe Pro Leu Thr Phe Glu Ala Ala Val Gln Lys Ser Trp	
	1585 1590 1595 1600	
	Ser Asn Ile Ser Ser Glu Asn Asn Gln Cys Tyr Val Lys Ala Thr Glu	
	1605 1610 1615	
	Leu Val Phe Ala Glu Lys Asn Gly Ser Trp Gly Thr Pro Ile Ile Pro	
	1620 1625 1630	
45	Met Gln Arg Ala Ala Gly Leu Asn Asp Ile Gly Met Val Ala Trp Ile	
	1635 1640 1645	
	Leu Asp Met Ser Thr Pro Glu Phe Pro Ser Gly Arg Gln Ile Ile Val	
	1650 1655 1660	
	Ile Ala Asn Asp Ile Thr Phe Arg Ala Gly Ser Phe Gly Pro Arg Glu	
50	1665 1670 1675 1680	
	Asp Ala Phe Phe Glu Ala Val Thr Asn Leu Ala Cys Glu Lys Lys Leu	
	1685 1690 1695	
	Pro Leu Ile Tyr Leu Ala Ala Asn Ser Gly Ala Arg Ile Gly Ile Ala	
	1700 1705 1710	
55	Asp Glu Val Lys Ser Cys Phe Arg Val Gly Trp Thr Asp Asp Ser Ser	
	1715 1720 1725	
	Pro Glu Arg Gly Phe Arg Tyr Ile Tyr Met Thr Asp Glu Asp His Asp	

	1730	1735	1740
	Arg Ile Gly Ser Ser Val	Ile Ala His Lys Met	Gln Leu Asp Ser Gly
5	1745	1750	1755
	Glu Ile Arg Trp Val	Ile Asp Ser Val Val	Gly Lys Glu Asp Gly Leu
	1765	1770	1775
	Gly Val Glu Asn Ile His Gly Ser Ala	Ala Ile Ala Ser Ala	Tyr Ser
	1780	1785	1790
10	Arg Ala Tyr Glu Glu Thr Phe	Thr Leu Thr Phe Val	Thr Gly Arg Thr
	1795	1800	1805
	Val Gly Ile Gly Ala Tyr Leu	Ala Arg Leu Gly	Ile Arg Cys Ile Gln
	1810	1815	1820
	Arg Ile Asp Gln Pro Ile	Ile Leu Thr Gly Phe	Ser Ala Leu Asn Lys
	1825	1830	1835
	Leu Leu Gly Arg Glu Val	Tyr Ser Ser His	Met Gln Leu Gly Gly Pro
	1845	1850	1855
15	Lys Ile Met Ala Thr Asn Gly Val	Val His Leu Thr Val	Pro Asp Asp
	1860	1865	1870
	Leu Glu Gly Val Ser Asn Ile	Leu Arg Trp Leu Ser	Tyr Val Pro Ala
	1875	1880	1885
20	Asn Ile Gly Gly Pro Leu Pro	Ile Thr Lys Ser Leu	Asp Pro Ile Asp
	1890	1895	1900
	Arg Pro Val Ala Tyr Ile	Pro Glu Asn Thr Cys	Asp Pro Arg Ala Ala
	1905	1910	1915
	Ile Ser Gly Ile Asp Asp Ser Gln	Gly Lys Trp Leu Gly	Gly Met Phe
	1925	1930	1935
25	Asp Lys Asp Ser Phe Val Glu	Thr Phe Glu Gly Trp Ala	Lys Thr Val
	1940	1945	1950
	Val Thr Gly Arg Ala Lys	Leu Gly Ile Pro Val	Gly Val Ile Ala
	1955	1960	1965
	Val Glu Thr Gln Thr Met	Met Gln Leu Val	Pro Ala Asp Pro Gly Gln
	1970	1975	1980
30	Pro Asp Ser His Glu Arg	Ser Val Pro Arg Ala	Gly Gln Val Trp Phe
	1985	1990	1995
	Pro Asp Ser Ala Thr Lys	Thr Ala Gln Ala	Met Leu Asp Phe Asn Arg
	2005	2010	2015
	Glu Gly Leu Pro Leu Phe	Ile Leu Ala Asn Trp Arg	Gly Phe Ser Gly
	2020	2025	2030
35	Gly Gln Arg Asp Leu Phe	Glu Gly Ile Leu Gln Ala	Gly Ser Thr Ile
	2035	2040	2045
	Val Glu Asn Leu Arg Thr	Tyr Asn Gln Pro Ala	Phe Val Tyr Ile Pro
	2050	2055	2060
	Lys Ala Ala Glu Leu Arg	Gly Gly Ala Trp Val	Val Ile Asp Ser Lys
	2065	2070	2075
40	Ile Asn Pro Asp Arg Ile	Glu Cys Tyr Ala	Glu Arg Thr Ala Lys Gly
	2085	2090	2095
	Asn Val Leu Glu Pro Gln	Gly Leu Ile Glu	Ile Lys Phe Arg Ser Glu
	2100	2105	2110
	Glu Leu Lys Glu Cys Met	Gly Arg Leu Asp Pro	Glu Leu Ile Asp Leu
	2115	2120	2125
45	Lys Ala Arg Leu Gln	Gly Ala Asn Gly	Ser Leu Ser Asp Gly Glu Ser
	2130	2135	2140
	Leu Gln Lys Ser Ile	Glu Ala Arg Lys Lys	Gln Leu Leu Pro Leu Tyr
	2145	2150	2155
	Thr Gln Ile Ala Val	Arg Phe Ala Glu	Leu His Asp Thr Ser Leu Arg
	2165	2170	2175
50	Met Ala Ala Lys Gly Val	Ile Arg Lys Val Val	Asp Trp Glu Asp Ser
	2180	2185	2190
	Arg Ser Phe Phe Tyr	Lys Arg Leu Arg	Arg Arg Leu Ser Glu Asp Val
	2195	2200	2205
55	Leu Ala Lys Glu Ile Arg	Gly Val Ile Gly	Glu Lys Phe Pro His Lys
	2210	2215	2220
	Ser Ala Ile Glu Leu Ile	Lys Lys Trp Tyr	Leu Ala Ser Glu Ala Ala
	2225	2230	2235
			2240

	Ala Ala Gly Ser Thr Asp Trp Asp Asp Asp Ala Phe Val Ala Trp	
	2245 2250 2255	
	Arg Glu Asn Pro Glu Asn Tyr Lys Glu Tyr Ile Lys Glu Leu Arg Ala	
	2260 2265 2270	
5	Gln Arg Val Ser Arg Leu Leu Ser Asp Val Ala Gly Ser Ser Ser Asp	
	2275 2280 2285	
	Leu Gln Ala Leu Pro Gln Gly Leu Ser Met Leu Leu Asp Lys Met Asp	
	2290 2295 2300	
10	Pro Ser Lys Arg Ala Gln Phe Ile Glu Glu Val Met Lys Val Leu Lys	
	2305 2310 2315 2320	
	<210> 4	
	<211> 6963	
	<212> DNA	
15	<213> Alopecurus myosuroides ACCase gene DNA sequence (DNA sequence)	
	<400> 4	
	atgggatcca cacatctgcc cattgtcggt ttaatgcac ccacaacacc atcgctatcc 60	
	actttcgcc agataaactc agctgcgtct gcattccaaat cttcgcccccc ttcaagggtca 120	
	tccaaagaaga aaagccgacg tgtaagtca ataagggtatc atggcgatgg aagcgtgcca 180	
20	gaccctgcag gccatggcca gtctattcgc caagggtctcg ctggcatcat cgacccccc 240	
	aaggaggcg catcagctcc agatgtggac atttcacatc ggtctgaaga ccacaaggcc 300	
	tcctaccaaa tgaatggat actgaatgaa tcacataacg ggaggcacgc ctctctgtct 360	
	aaagtttatg aattttgcac ggaatttgggt gaaaaaacac caattcacag tgtagttatc 420	
25	gccaacaatg gaatggcagc agctaagttc atgcggatgt tccggacatg ggctaatgtat 480	
	acattttgggt cagagaaggc gattcagttt atagctatgg caactccgga agacatgaga 540	
	ataaatgcag agcacatttt aattgtgtat cagtttgggt aagtacctgg tgaacaaac 600	
30	aataacaact atgcaaatgt ccaactcata gtggagatag cagagagaac tgggtgtctcc 660	
	gccgttggc ctgggtgggg ccatgcattt gagaatcctg aacttccaga tgcactaact 720	
	gcaaaaggaa ttgtttttct tggccacca gcatcatcaa tgaacgcact aggcgacaag 780	
	gttgggtcag ctctcattgc tcaagcagca ggggttccccctt ctcttgcgtt gagggtatca 840	
35	catgtggaaa ttccatttgc acatttttttgc gactcgatac ctgaggagat gtataggaaa 900	
	gcctgtgtta caaccgttgc tgaagcagtt gcaagttgtc agatgattgg ttaccctgcc 960	
	atgatcaagg catcctgggg tgggtgggt aaagggtttaa taatgtatc 1020	
	gagggtggaaag cactgtttaa gcaagtgacag ggtgaagttc ctggcccccc gatattttatc 1080	
	atgagacttgc catctcagag tcgtcatctt gaagtccagc tgctttgtca tgaatatggc 1140	
40	aatgttagcag cacttcacag tcgtgattgc agtgtgcaac gacgacacca aaagattatc 1200	
	gaggaaggac cagttactgt tgctccctgt gaaacagtga aagagctaga gcaagcagca 1260	
	aggaggcttgc ttaaggccgtt gggtaacgtc ggtgctgtca ctgttgcata tctctacagc 1320	
45	atggagactgt gtgaactata ttttctggag ctaatccac ggttgcagggt tgaggcaccc 1380	
	gtcaccgggtt ctagactgtca agtaaattttt cctgcagcccc aagttgcagggt tggatgggt 1440	
	ataccctttt ggcagattcc agagatcaga cttttctacg gaatggacaa tggaggaggc 1500	
	tatgatattt ggaggaaaaac agcagcttc tctactccat tcaactttgtca tgaagtagat 1560	
50	tctcaatggc cgaagggtca ttgtgtggca gtaggtatcaa ccagtgcataa tccagatgtat 1620	
	ggattcaagc ctactgggtt aaaagttaaag gagataagtt taaaagttaa gccaatgtc 1680	
	tggggatatt tctcatttgc tctgtgttgc ggcattcatg aatttgcggg ttctcattttt 1740	
	ggacacgttt ttgcctatgg agagactaga tcagcagcaa taaccagcat gtcttgcata 1800	
	ctaaaagaga ttcaaaatttcg tggagaaattt cataccaaacg ttgattacac ggttgcatttc 1860	
	ttgaatggccc cagacttcag agaaaaacacg atccataccg gttggcttgc taccagaata 1920	
55	gctatgcgtt tcaagcttgc gaggccccc tggatattt cagttgttgg aggagcttca 1980	
	tataaaacaa taaccaccaa tgcggagacc gtttctgaat atgttagtca tctcatcaag 2040	
	ggtcagatttcc caccacaaatgc catatccctt gtcatttcaat tattttttttaa gaatatacg 2100	
	gaaagccaaat atacaatttgc gattgtggagg agtggacagg gtagctacag atggagactg 2160	
	aatggatcac ttatgttgc caatgtacaa acatattgtt atggaggccct ttaatgttgc 2220	
	ctggatggaa atagccatgt tattttgttgc gagaagaag cgggtgttgc acggcttctt 2280	
	attgtatggaa aaacatgtttt gctacagaat gccatgttgc cgtcaagggtt attagcttgc 2340	
	acacccttgc aacttcttcg tttcttgcattt gccatgttgc cttatgttgc tggatgttgc 2400	
	ccatacgcgg aagttgggtt tatgtatgc tgcatttttttgc tcttgcgttgc tggatgttgc 2460	
	gtcattatgc ttttgcatttgc tgagggccag gggatgttgc ctttgcatttgc tttatgttgc 2520	
	cttgatctcg atgacccttc tgctgttgc agagccgagc catttttttgc atctttttttttt 2580	
	gaaatgttgc ttcttatttgc tgcttgcatttgc caatgttgc tttatgttgc tggatgttgc 2640	
	aacgctgttgc gaatgttgc tgcaggat gccatgttgc ccaacaaatgttgc tggatgttgc 2700	
	ttggatgttgc gcttgcatttgc accttgcatttgc ctttgcatttgc aatggggaaatgttgc tggatgttgc 2760	
55	tttttagtgc ttagacttgc aagacgttgc tggatgttgc aatacaatgttgc tggatgttgc 2820	

	tacaaggtaa atgttgacca tgtgaagatc aaggattcc ctaccgagat gcttagagag	2880
5	acaatcgagg aaaatcttgc atgtgttcc gagaaggaaa tggtgacaat tgagaggctt	2940
	gttgaccctc tcatgagcct gctgaagtca tacgagggtg ggagagaaag ccatgcccac	3000
	tttattgtca agtcccttt tgaggagtt ctctcggtt aggaactatt cagtatggc	3060
	attcagtctg acgtgattga acgcctgcgct cacaataataa gtaaagacct ccagaagggtt	3120
	gtagacattt ttttgtctca ccagggtgtg agaaacaaaa caaagctgat actcgccgctc	3180
10	atggagaaac tggcttatcc aaaccctgct gcctacagag atcagttgat tcgctttct	3240
	tccctcaacc ataaaagata ttataagttt gctcttaag ctatgtact tcttgaacaa	3300
	accaagctca gCGAactCCG cacaaggattt gcaaggaacc tttcagcgct ggatatgttc	3360
	accgaggaaa aggcaattt ctccttgcgaa gacagaaaat tggccattaa tgagagcatg	3420
	ggagatttag tcaactgcccc actgcccattt gaagatgcac ttgtttctt gtttattgt	3480
15	actgatcaa ctcttcagca gagagtgtt cagacatatac tatctcgatt ataccagcct	3540
	caacttgta aggtatgtt ccagctgaaa tatcaggatt ctgggtttat tgctttatgg	3600
	gaattactg aagggaaatca tgagaagaga ttgggtgcta tggttatccat gaagtcaacta	3660
	gaatctgtgt caacagccat tggagctgtt ctaaaggatg catcacatta tgcagctct	3720
	gcgggcaaca cggtgcatat tgctttgtt gatgctgata cccaactgaa tacaactgaa	3780
	gatagtgggtg ataatgacca agtcaagac aagatggata aactttctt tttactgaaa	3840
	caagatgttgc tcatggctga tctacgtgt gctgatgtca aggttggtag ttgcattgtt	3900
	caaagagatg gagaatcat gctatgcgc cgtagcttcc tcttgcaga gaaaaaactt	3960
20	tgttacgagg aagagccat tttcggcat gtggagccctc cactttctgc acttcttgag	4020
	ttggataaat tggaaatgtt gggatataat gagatgttggt atacaccgtc acgtgatcg	4080
	cagtggcata tatacacact tagaaataact gaaaatccaa aaatgtcga cagggtattt	4140
	ttccgaacac ttgtcagaca acccagtgc gcaacagggtt ttcacatcaga ccatatcact	4200
	gatgttgaag taggacacgc agaggaacctt cttcattttt cttcaagcgat catattaaaa	4260
25	tcgttgaaga ttgtctaaaga agaatttggag cttcacgcga tcaggactgg ccattctcat	4320
	atgtacttgtt gcatattgaa agagcaaaag ctcttgcattt tttttctgtt ttcagggaac	4380
	actgttgtgg atgttgttca agatgttgc actgtcatgtt ctcttttgcgaa agaaatggct	4440
	ttaaagatac atgaacttgtt tggtcagaatg atgcattatc tttctgtatg ccagtggaa	4500
	gtgaaacttta agtttgttggat cgatggcctt gcaatgttgc gctggagatg tggaaacacc	4560
30	aatgttactg gtcacacactg cactgttggat atctaccggg aggtcgaaga tacagaatca	4620
	cagaaacttag tataccactc caccgcattt tcatctggc ctttgcatttgc ttttgcactg	4680
	aataactctgtt atcagccattt gatgttattt gattttttttt gttgctctgc caggaacaac	4740
	aaaactacat actgtatgtt tttccatttgc acatttgcgat ctgcagtgca gaatgtcg	4800
	tctaaccattt ccagtgaaaaa caaccaatgtt ttttttttttgcgat ttttttttttgcgat	4860
	gaaaagaatg ggtcggtgggg cactccata attccatgc agcgtgtgc tgggctgaaat	4920
	gacatttggta ttgttgcctt gatcttggac atgtccactc ctgaattttcc cagcggcaga	4980
35	cagatcattt ttatcgcaaa ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5040
	gatgcatttt tggaaatgttgc aaccaacactg gtttgcgat ttttttttttgcgat ttttttttttgcgat	5100
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5160
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5220
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5280
40	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5340
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5400
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5460
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5520
45	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5580
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5640
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5700
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5760
50	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5820
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5880
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	5940
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6000
55	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6060
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6120
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6180
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6240
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6300
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6360
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6420
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6480
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6540
	tttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat ttttttttttgcgat	6600

5	cgaggaggc tatccgagga cgttctggca aaggagatta gaggtgtaat tggtgagaag	6660
	tttcctcaca aatcagcgat cgagctgatc aaaaaatggt acttggcttc tgagggcagct	6720
	gcagcaggaa gcaccgactg ggatgacac gatgctttg tcgcctggag ggagaaccct	6780
	aaaaactata aggagtata caaagagctt aggctcaaa ggatctcg ttgctctca	6840
	gatgttgcag gctccagttc ggatttacaa gcctgccgc agggctttc catgctacta	6900
	gataagatgg atccctctaa gagagcacag tttatcgagg aggtcatgaa gtcctgaaa	6960
	tga	6963

<210> 5  
<211> 446

&lt;212&gt; PRT

&lt;213&gt; Oryza sativa L. HPPD protein amino acid sequence (amino acid sequence)

&lt;400&gt; 5

Met Pro Pro Thr Pro Thr Ala Thr Thr Gly Ala Val Ser Ala

1 5 10 15

Ala Ala Ala Ala Gly Glu Asn Ala Gly Phe Arg Leu Val Gly His Arg  
20 25 30Arg Phe Val Arg Ala Asn Pro Arg Ser Asp Arg Phe Gln Ala Leu Ala  
35 40 45Phe His His Val Glu Leu Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly  
50 55 60Arg Phe Ala Phe Ala Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu  
65 70 75 80Ser Thr Gly Asn Ser Ala His Ala Ser Leu Leu Leu Arg Ser Ala Ser  
85 90 95Val Ala Phe Leu Phe Thr Ala Pro Tyr Gly Gly Asp His Gly Val Gly  
100 105 110Ala Asp Ala Ala Thr Thr Ala Ser Ile Pro Ser Phe Ser Pro Gly Ala  
115 120 125Ala Arg Arg Phe Ala Ala Asp His Gly Leu Ala Val His Ala Val Ala  
130 135 140Leu Arg Val Ala Asp Ala Ala Asp Ala Phe Arg Ala Ser Val Ala Ala  
145 150 155 160Gly Ala Arg Pro Ala Phe Gln Pro Ala Asp Leu Gly Gly Phe Gly  
165 170 175Leu Ala Glu Val Glu Leu Tyr Gly Asp Val Val Leu Arg Phe Val Ser  
180 185 190His Pro Asp Gly Ala Asp Ala Pro Phe Leu Pro Gly Phe Glu Gly Val  
195 200 205Ser Asn Pro Gly Ala Val Asp Tyr Gly Leu Arg Arg Phe Asp His Val  
210 215 220Val Gly Asn Val Pro Glu Leu Ala Pro Val Ala Ala Tyr Ile Ser Gly  
225 230 235 240Phe Thr Gly Phe His Glu Phe Ala Glu Phe Thr Ala Glu Asp Val Gly  
245 250 255Thr Ala Glu Ser Gly Leu Asn Ser Val Val Leu Ala Asn Asn Ala Glu  
260 265 270Thr Val Leu Leu Pro Leu Asn Glu Pro Val His Gly Thr Lys Arg Arg  
275 280 285Ser Gln Ile Gln Thr Tyr Leu Asp His His Gly Pro Gly Val Gln  
290 295 300His Ile Ala Leu Ala Ser Asp Asp Val Leu Gly Thr Leu Arg Glu Met  
305 310 315 320Arg Ala Arg Ser Ala Met Gly Gly Phe Glu Phe Leu Ala Pro Pro Pro  
325 330 335Pro Asn Tyr Tyr Asp Gly Val Arg Arg Ala Gly Asp Val Leu Ser  
340 345 350Glu Glu Gln Ile Asn Glu Cys Gln Glu Leu Gly Val Leu Val Asp Arg  
355 360 365Asp Asp Gln Gly Val Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp  
370 375 380

Arg Pro Thr Phe Phe Leu Glu Met Ile Gln Arg Ile Gly Cys Met Glu

385	390	395	400
Lys Asp Glu Ser Gly Gln Glu Tyr Gln Lys	Gly Gly Cys Gly Gly	Phe	
405	410	415	
Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys	Ser Ile Glu Glu Tyr	Glu	
420	425	430	
5 Lys Ser Leu Glu Ala Lys Gln Ala Pro Thr	Val Gln Gly Ser		
435	440	445	

<210> 6  
<211> 2099

<212> DNA

<213> Oryza sativa L. HPPD genome DNA sequence (DNA sequence)

<400> 6

atgcctccca ctcccacccc	caccggcacc accggcgccg	tctcgccgc tgccggggcg	60
ggggagaacg cgggttccg	cctcgctcggg caccggcgct	tctcgccgc caacccgggg	120
15 agcgaccgg tccaggcgct	cgcgttccac cacgtcgagc	tctcgccgc cgacggcgcc	180
tccggccgg gccgggttcgc	cttcggccctg ggccgcggc	tccggccag gtccgaccc	240
tccacggga actccgcga	cgcgttccctc ctccgttgcgt	ccgcgttccgt cggtttccctc	300
20 ttcaccggcc cctacggcg	cgaccacggc gtccggcgccg	acggccgcac caccgcctcc	360
atcccttctc tctcccccagg	cgccgcggcg aggttcgcgc	cgaccacgg cctcgccgtg	420
cacggccgtgg cgctgcgcgt	cgccgcgcgc gccgacgcct	tccgcgcag cgtcgccggcc	480
25 ggtgcgcgc cggcggtcca	gcccggccac ctccggggtg	gttcggccct cgccggagggt	540
gagctctaeg ggcacgtcg	gttcggccctc gtcaagccacc	cgacggcgcc cgacggcgcc	600
ttccctccgg gtttcgaggg	cgtcagcaac cccggccgg	tggactacgg cttccggccgg	660
ttcgaccacg tcgtcgccaa	cgtgccccgg ctgcgtccgg	tagccgcgtta catctccggg	720
ttcacccggg tccacgagtt	cgcccgaggtc accggccgagg	acgtgggcac cgccgagagc	780
25 ggcctcaact cgggtgtgt	cgccaaacaac gcgagaccc	tgctgtgcc gctcaacgag	840
ccgggtgcacg gcaccaagcg	gcccggccacg atacagacgt	acgtggacca ccacggccgc	900
ccgggggtgc agcacatgc	gctggccacg gacgacgtgc	tcgggacgc gagggagatg	960
cgccgcgcgt ccggcatggg	cggcttcgag ttcttggcgc	cgccgcgcgc caactactac	1020
30 gacggcgtgc ggcggcgcgc	cggggacgtg ctctcgagg	agcagatcaa cgagtgcac	1080
gagctcgcccc tgctcggtga	cagggtatgac caggggtgt	tgctccagat cttcaccaag	1140
ccagtaggag acaggtaaaa	ttctcacctc ttcatgtat	aaaatggctt atgaatttcag	1200
atttgcagtt atttggc	acatagcatac gattaggcgc	agaaagggtt caagcattat	1260
35 gaaattaatc cagaatgtct	gaataataca gtataatata	tgatagttag ctctgtata	1320
ctccatggat actctttatg	tgtctccatg aatccatgt	gccccttct gaagattgt	1380
acactagaaa ggaataaaag	ctgaatgtgc ataggaaaaa	aatgaaaagc caatgtgt	1440
35 ctgtttatgc cttcttgcaa	gcatatccca gttcttttt	gcccgcgt tgtaatgcag	1500
atagccagcc acatatacg	acttaattag ttagtactcc	ctctcacaat gtaagtcat	1560
ctagtatccc ccacattcat	attgatgcta atctatctag	attcatttgc atcaatatga	1620
atatggaaaa tactagaatg	acttacattg taaaacggag	gaagtattac ttactacatc	1680
40 taaggtccat ggattccctt	ttttacaaaa gaaagaaaaga	atcttatggc aactccatca	1740
gcataaaacca gcaatgtgc	tggaaacaac taaaactta	ggttcaggag gtttaattt	1800
tctttaagct taatagtctg	attcagtcg tattctaatt	tctgctgcatt cttgtctatt	1860
gttattttct ctctgtgact	ccaaatctaa ctgatcgc	tatttcactc agggcaaccc	1920
40 ttttcttggaa gatgatacaa	aggattgggt gcatggagaa	ggatgagagt gggcaggag	1980
accagaaggg cggctgcggc	gggtttggga aggcaactt	ctcgagctg ttcagttcca	2040
ttgaggagta tgagaaatcc	tttgcggccca agcaagcccc	tacagttcaa ggatccctag	2099

45 <210> 7  
<211> 536  
<212> PRT

<213> Oryza sativa L. PPO1 protein amino acid sequence (amino acid sequence)

50 <400> 7

Met Ala Ala Ala Ala Ala Met	Ala Thr Ala Thr Ser Ala Thr Ala	
1 5	10	15
Ala Pro Pro Leu Arg Ile Arg Asp	Ala Ala Arg Arg Thr Arg Arg Arg	
20 25	30	
55 Gly His Val Arg Cys Ala Val Ala Ser Gly Ala Ala Glu Ala Pro Ala		
35 40	45	
Ala Pro Gly Ala Arg Val Ser Ala Asp Cys Val Val Val Gly Gly Gly		

	50	55	60
	Ile Ser Gly Leu Cys Thr Ala Gln Ala Leu Ala Thr Lys His Gly Val		
5	65 70 75 80		
	Gly Asp Val Leu Val Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile		
	85 90 95		
	Thr Thr Ala Glu Arg Ala Gly Glu Gly Tyr Leu Trp Glu Glu Gly Pro		
	100 105 110		
	Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser		
	115 120 125		
10	Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe		
	130 135 140		
	Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Gly Asp		
	145 150 155 160		
	Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly		
	165 170 175		
15	Leu Gly Ala Leu Gly Val Arg Ala Pro Pro Pro Gly Arg Glu Glu Ser		
	180 185 190		
	Val Glu Asp Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg		
	195 200 205		
	Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys		
20	210 215 220		
	Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Asp Thr		
	225 230 235 240		
	Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly		
	245 250 255		
25	Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly		
	260 265 270		
	Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala		
	275 280 285		
	Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr		
	290 295 300		
30	Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr		
	305 310 315 320		
	Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile		
	325 330 335		
	Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala		
	340 345 350		
35	Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr		
	355 360 365		
	Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly		
	370 375 380		
	Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu		
	385 390 395 400		
40	Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro		
	405 410 415		
	Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn Thr		
	420 425 430		
	Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg		
	435 440 445		
45	Asp Leu Arg Lys Met Leu Ile Asn Pro Lys Ala Val Asp Pro Leu Val		
	450 455 460		
	Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly		
	465 470 475 480		
	His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly Gly		
	485 490 495		
50	Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu		
	500 505 510		
	Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp		
	515 520 525		
55	Tyr Leu Thr Lys Tyr Ala Tyr Lys		
	530 535		

<211> 3246  
 <212> DNA  
 <213> Oryza sativa L. PPO1 genome DNA sequence (DNA sequence)

5	<400>	8					
	atggccgcgg	ccgcccgcagg	catggccacc	gccacccctcg	ccacggcagc	gcccggcgtc	60
	cgcattcgcg	acgcccgcgg	gaggaccgcg	cgacgcggcc	acgttcgcgt	cgccgtcgcc	120
	agcggccgg	ccgaggcgcc	cgccggccccc	ggggcgcggg	tgtcggcgga	ctgcgtcggt	180
10	gtgggcggcg	gcatcagcgg	gtctcgcacc	gcmcaggcgc	tggccacaaa	gcacggcgtc	240
	ggcgcacgtgc	tcgtcacgg	ggcccgcgcc	cgccccggcg	gcaacatcac	caccgcggag	300
	cgccggcg	agggtcacct	ctgggaggag	gggccaaca	gttccagcc	ttccgacccc	360
15	gtcctcacca	tggccgtacg	tcttctgtct	cccttctct	tctgatttctc	tcgcggcgga	420
	gacgacatga	atgggaatgg	tggtgcattgg	atggggcgc	gcagggtggac	agcgggctca	480
	aggacgatct	cgtttcggg	gacccaaacg	cgccgcgggt	cgtgctgtgg	gaggggaagc	540
20	taaggccgt	gcccgtccaag	cccggcggcc	tggcgttctt	cgacctcatg	agcatccccg	600
	gcaagctcag	ggccggccctt	ggcgcgtcg	gcgttcgagc	gccaccccca	gtttgtgtgc	660
	tctcccgct	gtgcatttctt	gattcacttg	tggaaattcga	tttgtgcgt	cgtttccggc	720
	gaagggttca	ggggcgttag	gagtgcgtgg	aggacttcgt	gcggcgtcaac	ctcggcgcgg	780
25	aggtcttga	gcccgttattt	gaggcatttct	gttcagggtgt	ttattgttagt	gtgcatttgc	840
	tgtttgtttt	tttatgtattt	agataagaat	acggtgattt	cggtgcattag	gtgtgtatgc	900
	tggtgatcct	tcaaagctca	gtatgaaggc	tgcatattggg	aagggtgttga	ggctggagga	960
30	tactggaggt	agcattatttgc	gttgcattttt	caaaacaatc	caggagagg	ggaaaaaaccc	1020
	caaaccgcgg	aggatccgt	gagtggaaaa	ttgccttctt	tgttggatta	attgtccatt	1080
	gtgttacact	gatatgcctt	caccatttt	agccgccttc	caacgcctaa	ggggcagacaa	1140
35	gttgcatttt	tcaggaagg	tctgactatg	ctcccgatg	ctattacatc	taggtttgtt	1200
	atcattgtct	ttgttatttta	cctagtttctt	caactatgg	tattaggtgc	tgttagattgt	1260
	tcagatagat	gacatttgc	caacaatcta	gttagattga	ttgttatggc	ttgttgaatt	1320
40	aactgtttca	tttgcattcat	tgcctcagca	catatgaagc	atatggatag	atttttcaat	1380
	catttatccc	tcaataaacac	aattttgaca	ccagtgcatttca	tcctttttt	cattctgttt	1440
	catctgttca	tccacaatat	aattaagcat	acaaaagagg	cactcttga	tgcacaatttc	1500
45	atagtgttgc	gggttaat	tcatgttgc	tctttgaggg	acaattcggt	acacccttaac	1560
	atgaactagt	aatgttttggg	gtgttttacc	attgttcttgc	catttcttcc	attttcagg	1620
	tgggttagcaa	agtcacaaactt	tcatgttgc	tgacaaggat	tacaaagtca	gacaacaaag	1680
50	gatatgcatt	agtgttatgaa	acaccagaag	gggtgttctc	gttgcaagct	aaaactgttg	1740
	tcatgaccat	cccatcatat	gttgcatttgc	atatcttgc	gccacttca	gttgcatttata	1800
	tatattttat	taactttcttgc	ttcccaaat	acactgcagc	acttcatttgc	ttcctgagg	1860
	cctcgattca	tttttgcgtt	gacagggat	agtatttatt	tgcactttt	aagggttattaa	1920
55	ttcaacat	ccacttggaaa	tatacatatc	ctacacatcc	tgtcaacata	ctgtctaaac	1980
	agcattttgt	tttgcatttgc	tggcatctca	gcagccatt	actatcttgc	ggggacaaagc	2040
	cacatttctt	ataaaatcttgc	tcggaaatca	cttttgcattt	ttttagatt	agtgttgc	2100
	tagaatttgg	cttgcatttgc	tacttgcatt	gttgcatttgc	tttgcatttgc	ggacaatttc	2160
	ttcacaatc	tagcattat	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	2220
	ggttacat	catatgcata	aatttgcatt	tttgcatttgc	tttgcatttgc	tttgcatttgc	2280
	gtcttttacac	tgttattttt	agacatattt	tttgcatttgc	tttgcatttgc	tttgcatttgc	2340
	cccttccaca	gagtgcatt	gcagatgc	tgtcaatatt	ctattatcca	ccagggtgc	2400
	ctgttactgt	tttgcatttgc	aaagaagca	tttgcatttgc	tttgcatttgc	tttgcatttgc	2460
	tccagggttt	cgcccgatgt	catccgcgtt	gtcagggtt	tgagacttta	ggtacttata	2520
	ggaatttcaac	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	2580
55	catttatttgc	ggatttgcatttgc	tgttatttttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	2640
	taggaacaaat	atatagttca	tcacttgcatttgc	caaatcg	tccagcttgc	agggtgttac	2700
	ttcttgcatttgc	cataggatgt	tctacaaata	cagggtt	tttgcatttgc	tttgcatttgc	2760
	gttgcatttgc	tttgcatttgc	atgatttgc	tatccgtt	tttgcatttgc	tttgcatttgc	2820
	gactgttca	ataaaatcttgc	tactatgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	2880
	cagatagcct	atatttttgc	tttgcatttgc	cttgcatttgc	tttgcatttgc	tttgcatttgc	2940
	tttgcatttgc	aaagtgcatt	gttgcatttgc	gttgcatttgc	gttgcatttgc	gttgcatttgc	3000
	aatcttcaat	cagtggatcc	tttgcatttgc	ggcgtccggg	tatggccaca	agccataacca	3060
	cagggttca	tttgcatttgc	tgttgcatttgc	gagggttgc	aatcttgc	ggcgttgc	3120
	gttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	3180
	gttgcatttgc	catatgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	tttgcatttgc	3240
	aagtgc						3246
55	<210>	9					
	<211>	587					

&lt;212&gt; PRT

&lt;213&gt; Oryza sativa L. TIR1 protein amino acid sequence (amino acid sequence)

&lt;400&gt; 9

5	Met	Gly	Arg	Gly	Gly	Ser	Arg	Ala	Ala	Cys	Ala	Ala	Ala	Ala	Pro	Pro
	1			5				10							15	
	Trp	His	Ser	Leu	Pro	Asp	Glu	Val	Trp	Glu	His	Ala	Phe	Ser	Phe	Leu
				20				25							30	
10	Pro	Ala	Ala	Ala	Asp	Arg	Gly	Ala	Ala	Ala	Gly	Ala	Cys	Ser	Ser	Trp
				35				40							45	
	Leu	Arg	Ala	Glu	Arg	Arg	Ser	Arg	Arg	Arg	Leu	Ala	Val	Ala	Asn	Cys
				50				55							60	
15	Tyr	Ala	Ala	Ala	Pro	Arg	Asp	Ala	Val	Glu	Arg	Phe	Pro	Ser	Val	Arg
				65				70							80	
	Ala	Ala	Glu	Val	Lys	Gly	Lys	Pro	His	Phe	Ala	Asp	Phe	Gly	Leu	Val
				85				90							95	
20	Pro	Pro	Ala	Trp	Gly	Ala	Ala	Ala	Pro	Trp	Ile	Ala	Ala	Ala	Ala	
				100				105							110	
	Asp	Gly	Trp	Pro	Leu	Leu	Glu	Glu	Leu	Ser	Phe	Lys	Arg	Met	Val	Val
				115				120							125	
25	Thr	Asp	Glu	Cys	Leu	Glu	Met	Ile	Ala	Ala	Ser	Phe	Arg	Asn	Phe	Gln
				130				135							140	
	Val	Leu	Arg	Leu	Val	Ser	Cys	Asp	Gly	Phe	Ser	Thr	Ala	Gly	Leu	Ala
				145				150							160	
30	Ala	Ile	Ala	Ala	Gly	Cys	Arg	His	Leu	Arg	Glu	Leu	Asp	Leu	Gln	Glu
				165				170							175	
	Asn	Glu	Ile	Glu	Asp	Cys	Ser	Ile	His	Trp	Leu	Ser	Leu	Phe	Pro	Glu
				180				185							190	
35	Ser	Phe	Thr	Ser	Leu	Val	Thr	Leu	Asn	Phe	Ser	Cys	Leu	Glu	Gly	Glu
				195				200							205	
	Val	Asn	Ile	Thr	Val	Leu	Glu	Arg	Leu	Val	Thr	Arg	Cys	His	Asn	Leu
				210				215							220	
40	Lys	Thr	Leu	Lys	Leu	Asn	Asn	Ala	Ile	Pro	Leu	Asp	Lys	Leu	Ala	Ser
				225				230							240	
	Leu	Leu	His	Lys	Ala	Pro	Gln	Leu	Val	Glu	Leu	Gly	Thr	Gly	Lys	Phe
				245				250							255	
45	Ser	Ala	Asp	Tyr	His	Ser	Asp	Leu	Phe	Ala	Lys	Leu	Glu	Ala	Ala	Phe
				260				265							270	
50	Gly	Gly	Cys	Lys	Ser	Leu	Arg	Arg	Leu	Ser	Gly	Ala	Trp	Asp	Ala	Val
				275				280							285	
	Pro	Asp	Tyr	Leu	Pro	Ala	Phe	Tyr	Cys	Val	Cys	Glu	Gly	Leu	Thr	Ser
				290				295							300	
55	Leu	Asn	Leu	Ser	Tyr	Ala	Thr	Val	Arg	Gly	Pro	Glu	Leu	Ile	Lys	Phe
				305				310							320	
	Ile	Ser	Arg	Cys	Arg	Asn	Leu	Gln	Gln	Leu	Trp	Val	Met	Asp	Leu	Ile
				325				330							335	
	Glu	Asp	His	Gly	Leu	Ala	Val	Val	Ala	Ser	Ser	Cys	Asn	Lys	Leu	Gln
				340				345							350	
	Glu	Leu	Arg	Val	Phe	Pro	Ser	Asp	Pro	Phe	Gly	Ala	Gly	Phe	Leu	Thr
				355				360							365	
	Glu	Arg	Gly	Leu	Val	Asp	Val	Ser	Ala	Ser	Cys	Pro	Met	Leu	Glu	Ser
				370				375							380	
	Val	Leu	Tyr	Phe	Cys	Arg	Arg	Met	Thr	Asn	Glu	Ala	Leu	Ile	Thr	Ile
				385				390							400	
	Ala	Lys	Asn	Arg	Pro	Asn	Phe	Thr	Cys	Phe	Arg	Leu	Cys	Ile	Leu	Glu
				405				410							415	
	Pro	His	Thr	Pro	Asp	Tyr	Ile	Thr	Arg	Glu	Pro	Leu	Asp	Ala	Gly	Phe
				420				425							430	
	Ser	Ala	Ile	Val	Glu	Ser	Cys	Arg	Gly	Leu	Arg	Arg	Leu	Ser	Ile	Ser
				435				440							445	
55	Gly	Leu	Leu	Thr	Asp	Leu	Val	Phe	Lys	Ser	Ile	Gly	Ala	His	Ala	Asp
				450				455							460	
	Arg	Leu	Glu	Met	Leu	Ser	Ile	Ala	Phe	Ala	Gly	Asn	Ser	Asp	Leu	Gly

	465	470	475	480
	Leu His Tyr Ile Leu Ser Gly Cys Lys Ser	Leu Lys Lys Leu Glu Ile		
	485	490	495	
5	Arg Asp Cys Pro Phe Gly Asp Lys Pro	Leu Leu Ala Asn Ala Ala Lys		
	500	505	510	
	Leu Glu Thr Met Arg Ser Leu Trp Met Ser Ser Cys	Leu Leu Thr Leu		
	515	520	525	
	Gly Ala Cys Arg Gln Leu Ala Arg Lys Met Pro	Arg Leu Ser Val Glu		
	530	535	540	
10	Ile Met Asn Asp Pro Gly Arg Ser Cys Pro	Leu Asp Ser Leu Pro Asp		
	545	550	555	560
	Glu Thr Pro Val Glu Lys Leu Tyr Val Tyr Arg Thr Ile Ala Gly Pro			
	565	570	575	
	Arg Ser Asp Thr Pro Ala Cys Val Gln Ile Val			
	580	585		
15	<210> 10			
	<211> 5791			
	<212> DNA			
	<213> Oryza sativa L. TIR1 genome DNA sequence (DNA sequence)			
20	<400> 10			
	atggggcgcg gcccgtgc gcccggcgg cgccgcgtg	gcactcgctc	60	
	ccggacgagg tctggagca cgccttc	ttctccccg cggccggga caggggggcc	120	
	gcggcgaaaa cgtgcagctc	gtggctccgc gccgagcgcc ggtcgccgc cgcctcgcc	180	
25	gtcgccaaact gctacgcccgc	cgcggccggg gacggcgtcg agcggtccc gtccgtgcgc	240	
	gcccggagg tcaaggggcaa gccccacttc	gccgacttcg gcctcgccc cccgcctgg	300	
	ggcggccccc cggcggccgtg	gatcgccgccc gccggcgcacg ggtggccgat getcgaggag	360	
	ctcagcttca agcgcatggt	cgtcaccgac gagtgccctcg agatgatcgc cgcgtccctc	420	
	aggaacttcc aggtgctccg	cctcgcttcc tgcgacggct tcagcacgc gggcctcgcc	480	
30	gccattgctg ccgggttgcag	gtgagggttga gagtgttgg tgggtcaaat cttttagct	540	
	aaataagcat gggatctttt	gttgttctt ggatttgaga ggaatctttc acggcccaag	600	
	ctttgattcg ttttctttt	agctaaaaca tgacagattt ctgggtgcca ttctttctt	660	
	cacagtgttc ttttagtgagt	tcatggctca gctgaagtga ttgttcttc catgattcct	720	
	atattaactt ttttttgaac	ttcatttgagt agtatgttga agtagatgtg getttctgct	780	
	aagatttgat ctattcttgg	tgttaattt gcaaaggggac agttttttt tctgtttact	840	
35	tgtgtatgat ctggatttac	ttttgggttgc ttctaaatgg aagaccaaga agtcttctt	900	
	tgacagttt ttagatgttt	ataacgtgtt aatgtcaaaag gaatgattcc catgaaaatt	960	
	ctgctgtgtat tttaggtt	taagtaacag agtgaattga gtgatgtca	1020	
	ctgtttttttt aaatcttagc	tatgtttgtt tccttaggaag ctttgacttg tatttgagt	1080	
40	cctgggttgc gaagatgtga	ttctcttgca cgatctgtg tgattgacag aagaaatatt	1140	
	tctacctttt taagcaatga	tggatccatc ttttgttgc ttttgtgtt ttttacccaa	1200	
	aagaatggga ttttcaagt	ttacaaaacaa ctgggtatgc gcctttataa ttttgggctt	1260	
	ttttaaaaga atgtattttt	cccatcatga gtactatgtat ctgagttgaa ctgttttctt	1320	
45	ttatattttc ctgatattta	agtgaatttg ttgagttcct atcctgattt gaacttccag	1380	
	attctaaaga accaggttctt	ttaagcatgt ttgttgcatac tagattttca gtttatcccc	1440	
	tcccccttctt tttgtgtca	agaatctact accttttatac ttgtttttgg tgagggaaag	1500	
	tgatttttgc tccaatattt	gtgatgttgc catgtatgg acatgtaaaa caaacagatt	1560	
50	tgttctgtcg agtatcttct	ggactcccc cagaaaagaa ttaggaatgt gtgggattcc	1620	
	tatcatgtgg gaattttat	cctaaaccag atagaaggat ctaatagtgg gagtttgcgt	1680	
	cctgattcat ccatttagtag	tagtttataa tggccaaagca gtatcatttgc catcatttttgc	1740	
	aggaactcga aacagtgtaa	ctgtgaagtt taactactgt acacttcaaa caggcagatt	1800	
	agaataggtt gtgagataac	taaatttggt tcgggtatcg tataaggcca tatttttgg	1860	
55	tatttagacg aaatgaggcct	gaggcaccaca tggacttttg ctgatgttgc agaactttagc	1920	
	atagatacta gtttttatgc	tagacttagag ggctacttgc acattcaactt gttgataaaa	1980	
	tccagatgtta ttaagttttt	gattcaaat attttgactc tagattttggag actacatttt	2040	
	agagcattca aaatttacaa	tggaaaggcagg atttttagcaa aactattttc atacatttac	2100	
	tctgttattt agccagaatc	atttttttttgc tttttgttacacttgc cactgttatgc	2160	
	catgttaact cagcgttaca	tttatttcaa ttacaaggat ctacacttgc agctctgggt	2220	
	gcagatggta gttAACATTG	atgtatcttgc gctatatcgaa aatactatct tctaagtaca	2280	
	aacttttttctt gtgtgatttct	tatccttgc ttactgttgc ctacttataat ccattgttgc	2340	
	agcatgtttttaaatttctt	aaatgttgc ttacttgc aacgtcaaaat gtttgcatttgc	2400	
	gaactgtgtt tattttctccc	aaatgttgcatttgc aacgtcaaaat gtttgcatttgc	2460	

	gtagtgggca gcacatttgtt gcccctagctg caaaataaaac cttttttttt tggtaaaaaca	2520
5	atgattatga ggatgtcggt ttacactgtc cactcagaaaa ttgtacttga ttatggatag	2580
	ggtacacccctt gttgttttag atggttccctt cttaaacaaga tgctagattt tgatgatttg	2640
	attctttaca aggagtagtc taaaatactg gaagtgtcaa cttaaaaaagt tcaaacttgc	2700
	cctaggccata cggtgatacat tagacaactg atgtgtctgtt ttatctgcaat ctgacaatta	2760
	ttgttgttgc tatctccctt tttgtctta gagtaggcag gtatgttgc tcaactaaaa	2820
	gtgtgtttca cggtgtgcct ttcttaatac ctttgaata atagggtata agaaaatttt	2880
	catgaatggt ttgtttcca attgttgatg ttgtttccct tcaccttatt tttaaaccca	2940
10	ccatgattaa tctatttatac taggtaatat tgctcatgtc gattcatgtat acgtggtaaa	3000
	aaaataggc tttcaactt tcagaggctg ctggactcca tttatttagt tgatgttgc	3060
	ccctacacccg attttatcca ttggttata ttcaactgtca acatatgata ctcgctatca	3120
	tttacttcac tggatgtatcat caatgttttta attttgacca ttattaactt ttatgttgc	3180
	aatttagtat taatgaaagt tggcggtt tatttgc atgacacaca ttcttaacat	3240
	ggcatacttt taagcataca taactttttt tgtaaaaaagg acaagcgatc aaaggtaaaa	3300
	aggtgaatag catacagtgc gaagtattt agacccaaagg gagtataatc attatctact	3360
15	aaacaaatac aataccatttcc attataagct gagttcgcat agtgaattt aagtggggta	3420
	atgcttagtat agtcacaca tggaaatcaag tccctgtgaa atgatgtaa accctacgtg	3480
	ttcattgttt tagtcttgc atatgttcca gttatgtttt ccatgggtt tgaatgcagt	3540
	agtcaaagt tggcttgc ttgttgcgtt gattcttatt tacctgcct ttccaccat	3600
	atactcatat gggcttgc ttcaacaagtgc taacatttttgc taatggctat atgtatcc	3660
20	tacagttca ttcttctttt aagatatgag tcaccatgtc catttcttca attctgttgc	3720
	ttgtctaata ttgcagacac ctaagagaac ttgacctgtca agagaacgag attgaggatt	3780
	gttctattca ttggctcagc ctcttccccgg aatcggtcac ttctctagta actctaaact	3840
	tttcatgctt agagggggag gtcaatatac ctgtacttgc acgggttagt accagatgtc	3900
	acaacctgaa gactcttcaag ctcaacaatgc ctatccccct tgacaagctt gctagcctcc	3960
25	ttcataaggc tcctcagctt gttgaactcg gaaactggcaaa attctctgtt gattaccatt	4020
	ccgatctgtt tgcaaaagctg gaggcgccgt ttggagggtt taaaagctt tggaggctt	4080
	ctggggctt ggatgtcggtt ccagattatc tgccagcattt ctattgttgc tggtaaggcc	4140
	tcacatcact taatctgagt tatgtctactg tggcgagggtt tggacttgc aaattcatta	4200
	gtagatgcag aaatttgc当地 caattatggg tatgcaactc ttgtcaattt tttttacta	4260
	ctatgttattc atcccatttcg aatctgtttaa cttttgaaaa actgttgcattt atctctttgc	4320
30	aatacaagca ttgtattccc tacaaggtaa ttttgc当地 tttttttt ctctgttgc	4380
	ctccagtaaa aaagtttccgc atgttttccattt attatcatg gtcaactgtt aaagtccgac	4440
	tacatgttattt ctttttccac aatgttagttt tatacaattt attgttgc当地 ggctgcaata	4500
	gcattttatg atgaaactgtt ctgcactttt cccatgacaa caagaatttt tggttgc当地	4560
	taatcttattt agtggtaat catgcatttgc ttttgc当地 ttgtcaatac agttttgttgc	4620
	gatgcattttt acattcaatgc gcatgtcggtt ggggtcccttcc aatgaataga ttgaaattac	4680
35	cactcttattt ttttggggaa gtaggatatt tataatggaaa aaacatagat gacaatggta	4740
	tcctatcctt attagaatca ttgtgaaaag actgcccattt gatccattt gctctaggcc	4800
	agcatttttgc gcctttcttg gcaaaactgag ttattgtctt aacccttcata gttccatcc	4860
	tttaactgtt aactgtcttgc tggtttgc当地 gctctacggc agatatttttta aactgttgc	4920
	tgatctgttgc cccaatggctt attaacctgg tcatgttgc当地 tctttatgc当地 atcttttttgc	4980
40	tctatattttt tcttctt当地 gttaaagctgc aatttttgc当地 gtgtatggacc tcattgggaa	5040
	tcatgggttgc gctgttgc当地 catcatcttgc caataaaactt caagagggttgc gggcttccc	5100
	ttctgaccctt ttttgc当地 gattcttgc当地 tggaaagggat ttgttgc当地 tctctgttgc	5160
	ttgtccaaatgc ttggagtc当地 tgctctactt ctgcagacgg atgacaaatgc aggcaacttgc	5220
	taccatttgc aagaaccgtt ccaacttcac ttgttgc当地 ctatgcatttcc ttgagccaca	5280
	cactccagac tacatcacac gggagccttgc tgatgc当地 ttcaagccca ttgtggggcc	5340
45	atgcaggggc tttaggc当地 tctctatcttgc aggcttctc acagatcttgc tggtaatgc当地	5400
	cattggggca catgtgttgc当地 gtcttgc当地 gcttcaatgc gccttc当地 ggaacagccg	5460
	cttggccctt cattacatcc tctcaggcttgc caagaggcttgc aagaaactgg agatcaggga	5520
	ctgcccattt ggtgataaggc cattgttgc当地 gaacgc当地 aagctgggaa caatgc当地	5580
	cctttggatgc tggctgttgc当地 tggttgc当地 gggcgcatgc cgacagcttgc cacgc当地	5640
	gccccggc当地 agtgggagaa tcatgaacga ttctggaaagg tcatgc当地 tggatttc当地	5700
50	tccggatgaa acacctgttgc当地 agaaactgttgc当地 cgttctaccgg acgatgc当地 gtccaaagg	5760
	cgacacacca gccttc当地 agattgttgc当地 g	5791
	<210> 11	
	<211> 644	
	<212> PRT	
55	<213> Oryza sativa L. ALS protein amino acid sequence (amino acid sequence)	
	<400> 11	

## EP 3 889 266 A1

	Met Ala Thr Thr Ala Ala Ala Ala Ala Ala Thr Leu Ser Ala Ala Ala			
1	5	10	15	
	Thr Ala Lys Thr Gly Arg Lys Asn His Gln Arg His His Val Leu Pro			
	20	25	30	
5	Ala Arg Gly Arg Val Gly Ala Ala Ala Val Arg Cys Ser Ala Val Ser			
	35	40	45	
	Pro Val Thr Pro Pro Ser Pro Ala Pro Pro Ala Thr Pro Leu Arg Pro			
	50	55	60	
	Trp Gly Pro Ala Glu Pro Arg Lys Gly Ala Asp Ile Leu Val Glu Ala			
	65	70	75	80
10	Leu Glu Arg Cys Gly Val Ser Asp Val Phe Ala Tyr Pro Gly Gly Ala			
	85	90	95	
	Ser Met Glu Ile His Gln Ala Leu Thr Arg Ser Pro Val Ile Thr Asn			
	100	105	110	
	His Leu Phe Arg His Glu Gln Gly Glu Ala Phe Ala Ala Ser Gly Tyr			
	115	120	125	
15	Ala Arg Ala Ser Gly Arg Val Gly Val Cys Val Ala Thr Ser Gly Pro			
	130	135	140	
	Gly Ala Thr Asn Leu Val Ser Ala Leu Ala Asp Ala Leu Leu Asp Ser			
	145	150	155	160
20	Val Pro Met Val Ala Ile Thr Gly Gln Val Pro Arg Arg Met Ile Gly			
	165	170	175	
	Thr Asp Ala Phe Gln Glu Thr Pro Ile Val Glu Val Thr Arg Ser Ile			
	180	185	190	
	Thr Lys His Asn Tyr Leu Val Leu Asp Val Glu Asp Ile Pro Arg Val			
	195	200	205	
25	Ile Gln Glu Ala Phe Phe Leu Ala Ser Ser Gly Arg Pro Gly Pro Val			
	210	215	220	
	Leu Val Asp Ile Pro Lys Asp Ile Gln Gln Gln Met Ala Val Pro Val			
	225	230	235	240
	Trp Asp Thr Ser Met Asn Leu Pro Gly Tyr Ile Ala Arg Leu Pro Lys			
	245	250	255	
30	Pro Pro Ala Thr Glu Leu Leu Glu Gln Val Leu Arg Leu Val Gly Glu			
	260	265	270	
	Ser Arg Arg Pro Ile Leu Tyr Val Gly Gly Gly Cys Ser Ala Ser Gly			
	275	280	285	
	Asp Glu Leu Arg Arg Phe Val Glu Leu Thr Gly Ile Pro Val Thr Thr			
	290	295	300	
35	Thr Leu Met Gly Leu Gly Asn Phe Pro Ser Asp Asp Pro Leu Ser Leu			
	305	310	315	320
	Arg Met Leu Gly Met His Gly Thr Val Tyr Ala Asn Tyr Ala Val Asp			
	325	330	335	
	Lys Ala Asp Leu Leu Ala Phe Gly Val Arg Phe Asp Asp Arg Val			
	340	345	350	
40	Thr Gly Lys Ile Glu Ala Phe Ala Ser Arg Ala Lys Ile Val His Ile			
	355	360	365	
	Asp Ile Asp Pro Ala Glu Ile Gly Lys Asn Lys Gln Pro His Val Ser			
	370	375	380	
	Ile Cys Ala Asp Val Lys Leu Ala Leu Gln Gly Leu Asn Ala Leu Leu			
	385	390	395	400
45	Asp Gln Ser Thr Thr Lys Thr Ser Ser Asp Phe Ser Ala Trp His Asn			
	405	410	415	
	Glu Leu Asp Gln Gln Lys Arg Glu Phe Pro Leu Gly Tyr Lys Thr Phe			
	420	425	430	
50	Gly Glu Glu Ile Pro Pro Gln Tyr Ala Ile Gln Val Leu Asp Glu Leu			
	435	440	445	
	Thr Lys Gly Glu Ala Ile Ile Ala Thr Gly Val Gly Gln His Gln Met			
	450	455	460	
	Trp Ala Ala Gln Tyr Tyr Thr Tyr Lys Arg Pro Arg Gln Trp Leu Ser			
	465	470	475	480
55	Ser Ala Gly Leu Gly Ala Met Gly Phe Gly Leu Pro Ala Ala Ala Gly			
	485	490	495	
	Ala Ser Val Ala Asn Pro Gly Val Thr Val Val Asp Ile Asp Gly Asp			

	500	505	510
	Gly Ser Phe Leu Met Asn Ile Gln Glu Leu Ala Leu Ile Arg Ile Glu		
5	515	520	525
	Asn Leu Pro Val Lys Val Met Val Leu Asn Asn Gln His Leu Gly Met		
	530	535	540
	Val Val Gln Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg Ala His Thr		
	545	550	555
	Tyr Leu Gly Asn Pro Glu Cys Ser Glu Ile Tyr Pro Asp Phe Val		
	565	570	575
10	Thr Ile Ala Lys Gly Phe Asn Ile Pro Ala Val Arg Val Thr Lys Lys		
	580	585	590
	Ser Glu Val Arg Ala Ala Ile Lys Lys Met Leu Glu Thr Pro Gly Pro		
	595	600	605
	Tyr Leu Leu Asp Ile Ile Val Pro His Gln Glu His Val Leu Pro Met		
	610	615	620
15	Ile Pro Asn Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp Gly		
	625	630	635
	Arg Thr Met Tyr		640

20 <210> 12  
 <211> 1935  
 <212> DNA  
 <213> Oryza sativa L. ALS gene DNA sequence (DNA sequence)

	<400> 12	
25	atggctacga ccgcccgcgc cgccggcgcc accttgtccg ccggccgcac ggccaagacc	60
	ggccgttaaga accaccagcg acaccacgtc ctccccgtc gaggccgggt gggggcggcg	120
	gcgggtcaggt gctccgggt gtcccccgtc accccgcgt ccccgccgc gccggccacg	180
	ccgctccggc cgtgggggccc ggccgagccc cgcaaggccg cggacatcct cgtggaggcg	240
	ctggagcggt gcccgcgtcag cgacgtgttc gcctaccggc gccgcgcgtc catggagatc	300
30	caccaggcgc tgacgcgcgc cccggtcatac accaaccacc tcttccgcga cgagcaggcc	360
	gaggcgttgc cgccgtccgg gtacgcgcgc ggtccggcc ggtccgggt ctgcgtcgcc	420
	acctccggcc ccggggcaac caacctcgta tcgcgcgtcg ccgacgcgtc gtcgactcc	480
	gtccccatgg tcgccatcac gggccaggc cccgcgcga tgatcgac cgcgccttc	540
	caggagacgc ccatagtcga ggtcacccgc tccatcacca agcacaatta ctttgcctt	600
	gatgtggagg acatcccccg cgtcatacag gaagccttct tcctcggtc ctcggccgt	660
35	cctggccccc tgctggtcga catcccaag gacatccagc agcagatggc tgtgccagtc	720
	tgggacaccc cgtatgaatct accgggttac attgcacgc tgcccaagcc acccgccaca	780
	gaattgttg agcaggtctt ggtctgtt ggcagtcac ggcgcggat tctctatgtc	840
	ggtgtggct gctctgcata tggatgttgc ttgcgcgggt ttgtttagt gaccggcatc	900
	ccagttacaa ccaactctgtat gggcctccgc aatttccca gtatgtatcc gtgtccctg	960
40	cgcacatgtt ggtatgtatgg cacgggtgtac gcaaattatg cgggtggataa ggctgacact	1020
	ttgtttgcata ttggcgtgcg ttttgcata cgtgtgacag gaaaaattga ggctttgcata	1080
	agcaggccca agattgtgca cattgacatt gatccagccg agattggaaa gaacaagcaa	1140
	ccacatgtt caatttgcgc agatgttaag ctgttttac agggctgaa tgctctgtca	1200
	gaccagagca caacaaagac aagtctgtat ttiagtgcata ggcacaatga gttggaccag	1260
	cagaagaggg agtttccctt ggggtacaag acttttgggt aagagatccc accgcaatata	1320
45	gctattcagg tgctggatga gctgacaaa ggggaggccaa tcatcgctac tggtgttggaa	1380
	cagcaccaga tggggccgc acaatattac acctacaagc ggccacggca gtggctgtct	1440
	tcggctggtc tgggcgcata gggatttggg ctgcctgtc cagctgtgc ttctgtggct	1500
	aacccagggtg tcacagttgt tgatattgtat gggatggta gtttcctcat gaacattcag	1560
	gagttggcat tgatccgcata tgagaacctc cccgtgaagg tgatgggtt gaacaaccaa	1620
50	catttggta tggatgtgca atgggaggat aggtttaca aggcaaata ggcgcataaca	1680
	tacttggca acccagaatg tgagaggcag atatatccag attttgcata tattgtctaaa	1740
	gggttcaata ttccctgcata ccgtgtaaaca aagaagagt g aagtcgtgc cccatcaag	1800
	aagatgtcg agacccagg gccatacttg ttggatata tcgtcccaaca ccaggagcat	1860
	gtgctgcata tgatcccaa tggggccca ttcaaggaca tgatcctgga tggtgtatggc	1920
	aggactatgt attaa	1935

55 <210> 13  
 <211> 2327  
 <212> PRT

<213> Oryza sativa L. ACCase2 protein amino acid sequence (amino acid sequence)

<400> 13

5	Met Thr Ser Thr His Val Ala Thr Leu Gly Val Gly Ala Gln Ala Pro	
	1 5	10 15
	Pro Arg His Gln Lys Lys Ser Ala Gly Thr Ala Phe Val Ser Ser Gly	
	20 25	30
10	Ser Ser Arg Pro Ser Tyr Arg Lys Asn Gly Gln Arg Thr Arg Ser Leu	
	35 40	45
	Arg Glu Glu Ser Asn Gly Gly Val Ser Asp Ser Lys Lys Leu Asn His	
	50 55	60
	Ser Ile Arg Gln Gly Leu Ala Gly Ile Ile Asp Leu Pro Asn Asp Ala	
	65 70	75 80
15	Ala Ser Glu Val Asp Ile Ser His Gly Ser Glu Asp Pro Arg Gly Pro	
	85 90	95
	Thr Val Pro Gly Ser Tyr Gln Met Asn Gly Ile Ile Asn Glu Thr His	
	100 105	110
	Asn Gly Arg His Ala Ser Val Ser Lys Val Val Glu Phe Cys Thr Ala	
	115 120	125
20	Leu Gly Gly Lys Thr Pro Ile His Ser Val Leu Val Ala Asn Asn Gly	
	130 135	140
	Met Ala Ala Ala Lys Phe Met Arg Ser Val Arg Thr Trp Ala Asn Asp	
	145 150	155 160
	Thr Phe Gly Ser Glu Lys Ala Ile Gln Leu Ile Ala Met Ala Thr Pro	
	165 170	175
25	Glu Asp Leu Arg Ile Asn Ala Glu His Ile Arg Ile Ala Asp Gln Phe	
	180 185	190
	Val Glu Val Pro Gly Gly Thr Asn Asn Asn Asn Tyr Ala Asn Val Gln	
	195 200	205
	Leu Ile Val Glu Ile Ala Glu Arg Thr Gly Val Ser Ala Val Trp Pro	
	210 215	220
30	Gly Trp Gly His Ala Ser Glu Asn Pro Glu Leu Pro Asp Ala Leu Thr	
	225 230	235 240
	Ala Lys Gly Ile Val Phe Leu Gly Pro Pro Ala Ser Ser Met His Ala	
	245 250	255
	Leu Gly Asp Lys Val Gly Ser Ala Leu Ile Ala Gln Ala Ala Gly Val	
	260 265	270
35	Pro Thr Leu Ala Trp Ser Gly Ser His Val Glu Val Pro Leu Glu Cys	
	275 280	285
	Cys Leu Asp Ser Ile Pro Asp Glu Met Tyr Arg Lys Ala Cys Val Thr	
	290 295	300
	Thr Thr Glu Glu Ala Val Ala Ser Cys Gln Val Val Gly Tyr Pro Ala	
	305 310	315 320
40	Met Ile Lys Ala Ser Trp Gly Gly Gly Lys Gly Ile Arg Lys Val	
	325 330	335
	His Asn Asp Asp Glu Val Arg Thr Leu Phe Lys Gln Val Gln Gly Glu	
	340 345	350
	Val Pro Gly Ser Pro Ile Phe Ile Met Arg Leu Ala Ala Gln Ser Arg	
	355 360	365
45	His Leu Glu Val Gln Leu Leu Cys Asp Gln Tyr Gly Asn Val Ala Ala	
	370 375	380
	Leu His Ser Arg Asp Cys Ser Val Gln Arg Arg His Gln Lys Ile Ile	
	385 390	395 400
	Glu Glu Gly Pro Val Thr Val Ala Pro Arg Glu Thr Val Lys Glu Leu	
	405 410	415
50	Glu Gln Ala Ala Arg Arg Leu Ala Lys Ala Val Gly Tyr Val Gly Ala	
	420 425	430
	Ala Thr Val Glu Tyr Leu Tyr Ser Met Glu Thr Gly Glu Tyr Tyr Phe	
	435 440	445
55	Leu Glu Leu Asn Pro Arg Leu Gln Val Glu His Pro Val Thr Glu Trp	
	450 455	460
	Ile Ala Glu Val Asn Leu Pro Ala Ala Gln Val Ala Val Gly Met Gly	

## EP 3 889 266 A1

	465	470	475	480
	Ile Pro Leu Trp Gln Ile Pro Glu Ile Arg Arg Phe Tyr Gly Met Asn			
	485	490		495
5	His Gly Gly Tyr Asp Leu Trp Arg Lys Thr Ala Ala Leu Ala Thr			
	500	505		510
	Pro Phe Asn Phe Asp Glu Val Asp Ser Lys Trp Pro Lys Gly His Cys			
	515	520		525
	Val Ala Val Arg Ile Thr Ser Glu Asp Pro Asp Asp Gly Phe Lys Pro			
	530	535		540
10	Thr Gly Gly Lys Val Lys Glu Ile Ser Phe Lys Ser Lys Pro Asn Val			
	545	550	555	560
	Trp Ala Tyr Phe Ser Val Lys Ser Gly Gly Ile His Glu Phe Ala			
	565	570		575
	Asp Ser Gln Phe Gly His Val Phe Ala Tyr Gly Thr Thr Arg Ser Ala			
	580	585		590
15	Ala Ile Thr Thr Met Ala Leu Ala Leu Lys Glu Val Gln Ile Arg Gly			
	595	600		605
	Glu Ile His Ser Asn Val Asp Tyr Thr Val Asp Leu Leu Asn Ala Ser			
	610	615		620
	Asp Phe Arg Glu Asn Lys Ile His Thr Gly Trp Leu Asp Thr Arg Ile			
	625	630	635	640
20	Ala Met Arg Val Gln Ala Glu Arg Pro Pro Trp Tyr Ile Ser Val Val			
	645	650		655
	Gly Gly Ala Leu Tyr Lys Thr Val Thr Ala Asn Thr Ala Thr Val Ser			
	660	665		670
	Asp Tyr Val Gly Tyr Leu Thr Lys Gly Gln Ile Pro Pro Lys His Ile			
	675	680		685
25	Ser Leu Val Tyr Thr Thr Val Ala Leu Asn Ile Asp Gly Lys Lys Tyr			
	690	695		700
	Thr Ile Asp Thr Val Arg Ser Gly His Gly Ser Tyr Arg Leu Arg Met			
	705	710	715	720
	Asn Gly Ser Thr Val Asp Ala Asn Val Gln Ile Leu Cys Asp Gly Gly			
30	725	730		735
	Leu Leu Met Gln Leu Asp Gly Asn Ser His Val Ile Tyr Ala Glu Glu			
	740	745		750
	Glu Ala Ser Gly Thr Arg Leu Leu Ile Asp Gly Lys Thr Cys Met Leu			
	755	760		765
35	Gln Asn Asp His Asp Pro Ser Lys Leu Leu Ala Glu Thr Pro Cys Lys			
	770	775		780
	Leu Leu Arg Phe Leu Val Ala Asp Gly Ala His Val Asp Ala Asp Val			
	785	790	795	800
	Pro Tyr Ala Glu Val Glu Val Met Lys Met Cys Met Pro Leu Leu Ser			
	805	810		815
40	Pro Ala Ser Gly Val Ile His Val Val Met Ser Glu Gly Gln Ala Met			
	820	825		830
	Gln Ala Gly Asp Leu Ile Ala Arg Leu Asp Leu Asp Asp Pro Ser Ala			
	835	840		845
	Val Lys Arg Ala Glu Pro Phe Glu Asp Thr Phe Pro Gln Met Gly Leu			
	850	855		860
45	Pro Ile Ala Ala Ser Gly Gln Val His Lys Leu Cys Ala Ala Ser Leu			
	865	870	875	880
	Asn Ala Cys Arg Met Ile Leu Ala Gly Tyr Glu His Asp Ile Asp Lys			
	885	890		895
	Val Val Pro Glu Leu Val Tyr Cys Leu Asp Thr Pro Glu Leu Pro Phe			
	900	905		910
50	Leu Gln Trp Glu Glu Leu Met Ser Val Leu Ala Thr Arg Leu Pro Arg			
	915	920		925
	Asn Leu Lys Ser Glu Leu Glu Gly Lys Tyr Glu Glu Tyr Lys Val Lys			
	930	935		940
	Phe Asp Ser Gly Ile Ile Asn Asp Phe Pro Ala Asn Met Leu Arg Val			
	945	950	955	960
55	Ile Ile Glu Glu Asn Leu Ala Cys Gly Ser Glu Lys Glu Lys Ala Thr			
	965	970		975

	Asn Glu Arg Leu Val Glu Pro Leu Met Ser Leu Leu Lys Ser Tyr Glu			
	980	985	990	
	Gly Gly Arg Glu Ser His Ala His Phe Val Val Lys Ser Leu Phe Glu			
	995	1000	1005	
5	Glu Tyr Leu Tyr Val Glu Glu Leu Phe Ser Asp Gly Ile Gln Ser Asp			
	1010	1015	1020	
	Val Ile Glu Arg Leu Arg Leu Gln His Ser Lys Asp Leu Gln Lys Val			
	1025	1030	1035	1040
	Val Asp Ile Val Leu Ser His Gln Ser Val Arg Asn Lys Thr Lys Leu			
	1045	1050	1055	
10	Ile Leu Lys Leu Met Glu Ser Leu Val Tyr Pro Asn Pro Ala Ala Tyr			
	1060	1065	1070	
	Arg Asp Gln Leu Ile Arg Phe Ser Ser Leu Asn His Lys Ala Tyr Tyr			
	1075	1080	1085	
	Lys Leu Ala Leu Lys Ala Ser Glu Leu Leu Glu Gln Thr Lys Leu Ser			
	1090	1095	1100	
15	Glu Leu Arg Ala Arg Ile Ala Arg Ser Leu Ser Glu Leu Glu Met Phe			
	1105	1110	1115	1120
	Thr Glu Glu Ser Lys Gly Leu Ser Met His Lys Arg Glu Ile Ala Ile			
	1125	1130	1135	
20	Lys Glu Ser Met Glu Asp Leu Val Thr Ala Pro Leu Pro Val Glu Asp			
	1140	1145	1150	
	Ala Leu Ile Ser Leu Phe Asp Cys Ser Asp Thr Thr Val Gln Gln Arg			
	1155	1160	1165	
	Val Ile Glu Thr Tyr Ile Ala Arg Leu Tyr Gln Pro His Leu Val Lys			
	1170	1175	1180	
25	Asp Ser Ile Lys Met Lys Trp Ile Glu Ser Gly Val Ile Ala Leu Trp			
	1185	1190	1195	1200
	Glu Phe Pro Glu Gly His Phe Asp Ala Arg Asn Gly Gly Ala Val Leu			
	1205	1210	1215	
	Gly Asp Lys Arg Trp Gly Ala Met Val Ile Val Lys Ser Leu Glu Ser			
	1220	1225	1230	
30	Leu Ser Met Ala Ile Arg Phe Ala Leu Lys Glu Thr Ser His Tyr Thr			
	1235	1240	1245	
	Ser Ser Glu Gly Asn Met Met His Ile Ala Leu Leu Gly Ala Asp Asn			
	1250	1255	1260	
	Lys Met His Ile Ile Gln Glu Ser Gly Asp Asp Ala Asp Arg Ile Ala			
	1265	1270	1275	1280
35	Lys Leu Pro Leu Ile Leu Lys Asp Asn Val Thr Asp Leu His Ala Ser			
	1285	1290	1295	
	Gly Val Lys Thr Ile Ser Phe Ile Val Gln Arg Asp Glu Ala Arg Met			
	1300	1305	1310	
	Thr Met Arg Arg Thr Phe Leu Trp Ser Asp Glu Lys Leu Ser Tyr Glu			
	1315	1320	1325	
40	Glu Glu Pro Ile Leu Arg His Val Glu Pro Pro Leu Ser Ala Leu Leu			
	1330	1335	1340	
	Glu Leu Asp Lys Leu Lys Val Lys Gly Tyr Asn Glu Met Lys Tyr Thr			
	1345	1350	1355	1360
	Pro Ser Arg Asp Arg Gln Trp His Ile Tyr Thr Leu Arg Asn Thr Glu			
	1365	1370	1375	
45	Asn Pro Lys Met Leu His Arg Val Phe Phe Arg Thr Leu Val Arg Gln			
	1380	1385	1390	
	Pro Ser Val Ser Asn Lys Phe Ser Ser Gly Gln Ile Gly Asp Met Glu			
	1395	1400	1405	
50	Val Gly Ser Ala Glu Glu Pro Leu Ser Phe Thr Ser Thr Ser Ile Leu			
	1410	1415	1420	
	Arg Ser Leu Met Thr Ala Ile Glu Glu Leu Glu Leu His Ala Ile Arg			
	1425	1430	1435	1440
	Thr Gly His Ser His Met Tyr Leu His Val Leu Lys Glu Gln Lys Leu			
	1445	1450	1455	
55	Leu Asp Leu Val Pro Val Ser Gly Asn Thr Val Leu Asp Val Gly Gln			
	1460	1465	1470	
	Asp Glu Ala Thr Ala Tyr Ser Leu Leu Lys Glu Met Ala Met Lys Ile			

	1475	1480	1485
	His Glu Leu Val Gly Ala Arg Met His His	Leu Ser Val Cys Gln Trp	
	1490	1495	1500
5	Glu Val Lys Leu Lys Leu Asp Cys Asp Gly Pro Ala Ser Gly Thr Trp		
	1505	1510	1515
	Arg Ile Val Thr Thr Asn Val Thr Ser His Thr Cys Thr Val Asp Ile		1520
	1525	1530	1535
	Tyr Arg Glu Met Glu Asp Lys Glu Ser Arg Lys Leu Val Tyr His Pro		
	1540	1545	1550
10	Ala Thr Pro Ala Ala Gly Pro Leu His Gly Val Ala Leu Asn Asn Pro		
	1555	1560	1565
	Tyr Gln Pro Leu Ser Val Ile Asp Leu Lys Arg Cys Ser Ala Arg Asn		
	1570	1575	1580
	Asn Arg Thr Thr Tyr Cys Tyr Asp Phe Pro Leu Ala Phe Glu Thr Ala		
	1585	1590	1595
15	Val Arg Lys Ser Trp Ser Ser Ser Thr Ser Gly Ala Ser Lys Gly Val		
	1605	1610	1615
	Glu Asn Ala Gln Cys Tyr Val Lys Ala Thr Glu Leu Val Phe Ala Asp		
	1620	1625	1630
	Lys His Gly Ser Trp Gly Thr Pro Leu Val Gln Met Asp Arg Pro Ala		
	1635	1640	1645
20	Gly Leu Asn Asp Ile Gly Met Val Ala Trp Thr Leu Lys Met Ser Thr		
	1650	1655	1660
	Pro Glu Phe Pro Ser Gly Arg Glu Ile Ile Val Val Ala Asn Asp Ile		
	1665	1670	1675
	Thr Phe Arg Ala Gly Ser Phe Gly Pro Arg Glu Asp Ala Phe Phe Glu		
	1685	1690	1695
25	Ala Val Thr Asn Leu Ala Cys Glu Lys Lys Leu Pro Leu Ile Tyr Leu		
	1700	1705	1710
	Ala Ala Asn Ser Gly Ala Arg Ile Gly Ile Ala Asp Glu Val Lys Ser		
	1715	1720	1725
	Cys Phe Arg Val Gly Trp Ser Asp Asp Gly Ser Pro Glu Arg Gly Phe		
30	1730	1735	1740
	Gln Tyr Ile Tyr Leu Ser Glu Glu Asp Tyr Ala Arg Ile Gly Thr Ser		
	1745	1750	1755
	Val Ile Ala His Lys Met Gln Leu Asp Ser Gly Glu Ile Arg Trp Val		
	1765	1770	1775
35	Ile Asp Ser Val Val Gly Lys Glu Asp Gly Leu Gly Val Glu Asn Ile		
	1780	1785	1790
	His Gly Ser Ala Ala Ile Ala Ser Ala Tyr Ser Arg Ala Tyr Lys Glu		
	1795	1800	1805
	Thr Phe Thr Leu Thr Phe Val Thr Gly Arg Thr Val Gly Ile Gly Ala		
	1810	1815	1820
40	Tyr Leu Ala Arg Leu Gly Ile Arg Cys Ile Gln Arg Leu Asp Gln Pro		
	1825	1830	1835
	Ile Ile Leu Thr Gly Tyr Ser Ala Leu Asn Lys Leu Leu Gly Arg Glu		
	1845	1850	1855
	Val Tyr Ser Ser His Met Gln Leu Gly Gly Pro Lys Ile Met Ala Thr		
	1860	1865	1870
45	Asn Gly Val Val His Leu Thr Val Ser Asp Asp Leu Glu Gly Val Ser		
	1875	1880	1885
	Asn Ile Leu Arg Trp Leu Ser Tyr Val Pro Ala Tyr Ile Gly Gly Pro		
	1890	1895	1900
	Leu Pro Val Thr Thr Pro Leu Asp Pro Pro Asp Arg Pro Val Ala Tyr		
	1905	1910	1915
50	Ile Pro Glu Asn Ser Cys Asp Pro Arg Ala Ala Ile Arg Gly Val Asp		
	1925	1930	1935
	Asp Ser Gln Gly Lys Trp Leu Gly Gly Met Phe Asp Lys Asp Ser Phe		
	1940	1945	1950
	Val Glu Thr Phe Glu Gly Trp Ala Lys Thr Val Val Thr Gly Arg Ala		
	1955	1960	1965
55	Lys Leu Gly Gly Ile Pro Val Gly Val Ile Ala Val Glu Thr Gln Thr		
	1970	1975	1980

## EP 3 889 266 A1

	Met Met Gln Thr Ile Pro Ala Asp Pro Gly Gln Leu Asp Ser Arg Glu			
1985	1990	1995	2000	
	Gln Ser Val Pro Arg Ala Gly Gln Val Trp Phe Pro Asp Ser Ala Thr			
	2005	2010	2015	
5	Lys Thr Ala Gln Ala Leu Leu Asp Phe Asn Arg Glu Gly Leu Pro Leu			
	2020	2025	2030	
	Phe Ile Leu Ala Asn Trp Arg Gly Phe Ser Gly Gly Gln Arg Asp Leu			
	2035	2040	2045	
	Phe Glu Gly Ile Leu Gln Ala Gly Ser Thr Ile Val Glu Asn Leu Arg			
	2050	2055	2060	
10	Thr Tyr Asn Gln Pro Ala Phe Val Tyr Ile Pro Met Ala Ala Glu Leu			
	2065	2070	2075	2080
	Arg Gly Gly Ala Trp Val Val Val Asp Ser Lys Ile Asn Pro Asp Arg			
	2085	2090	2095	
	Ile Glu Cys Tyr Ala Glu Arg Thr Ala Lys Gly Asn Val Leu Glu Pro			
	2100	2105	2110	
15	Gln Gly Leu Ile Glu Ile Lys Phe Arg Ser Glu Glu Leu Gln Asp Cys			
	2115	2120	2125	
	Met Ser Arg Leu Asp Pro Thr Leu Ile Asp Leu Lys Ala Lys Leu Glu			
	2130	2135	2140	
20	Val Ala Asn Lys Asn Gly Ser Ala Asp Thr Lys Ser Leu Gln Glu Asn			
	2145	2150	2155	2160
	Ile Glu Ala Arg Thr Lys Gln Leu Met Pro Leu Tyr Thr Gln Ile Ala			
	2165	2170	2175	
	Ile Arg Phe Ala Glu Leu His Asp Thr Ser Leu Arg Met Ala Ala Lys			
	2180	2185	2190	
25	Gly Val Ile Lys Lys Val Val Asp Trp Glu Glu Ser Arg Ser Phe Phe			
	2195	2200	2205	
	Tyr Lys Arg Leu Arg Arg Arg Ile Ser Glu Asp Val Leu Ala Lys Glu			
	2210	2215	2220	
	Ile Arg Ala Val Ala Gly Glu Gln Phe Ser His Gln Pro Ala Ile Glu			
	2225	2230	2235	2240
30	Leu Ile Lys Trp Tyr Ser Ala Ser His Ala Ala Glu Trp Asp Asp			
	2245	2250	2255	
	Asp Asp Ala Phe Val Ala Trp Met Asp Asn Pro Glu Asn Tyr Lys Asp			
	2260	2265	2270	
	Tyr Ile Gln Tyr Leu Lys Ala Gln Arg Val Ser Gln Ser Leu Ser Ser			
	2275	2280	2285	
35	Leu Ser Asp Ser Ser Ser Asp Leu Gln Ala Leu Pro Gln Gly Leu Ser			
	2290	2295	2300	
	Met Leu Leu Asp Lys Met Asp Pro Ser Arg Arg Ala Gln Leu Val Glu			
	2305	2310	2315	2320
	Glu Ile Arg Lys Val Leu Gly			
40	2325			
	<210> 14			
	<211> 11927			
	<212> DNA			
	<213> Oryza sativa L. ACCase2 genome DNA sequence (DNA Sequence)			
45	<400> 14			
	atgacatcca cacatgtggc gacattggga gtgggtgcc aggcacctcc tcgtcaccag	60		
	aaaaagtca gtcggactgc atttgtatca tctgggtcat caagaccctc ataccgaaag	120		
	aatggtcaggc gtactcggtc acttagggaa gaaagcaatg gaggagtgtc tgattccaaa	180		
50	aagcttaacc actctattcg ccaagggtgac cactagctac ttatcatatg ctataatttg	240		
	tgcctaaacat aaacatgcaa tggctgttat tttaaacg ttaatgttga aatagctgct	300		
	ataggataca gcaaaaatat ataatttgact gggcaagatg caacaattgt ttttcactaa	360		
	agtttagttat cttttgtgt aaaagacaac tggtttttac ataaaaatggt attaataacc	420		
	ttgttaatatt caatgcaaca tggctcaag taaaaaaaaaa cattgcctgg ttgtataagc	480		
55	aatatgtgtcg tttagacat cttttaaac cttttgtga tatctattac ctagggAAC	540		
	aggggagctg tttaaatctg ttatcataga gtaatatgag aaaagtggat tggcgactt	600		
	tggcatgtat acctgctcaa ttcaaatat atgtctatgt gcaggtcttgc tggcatcat	660		
	tgacccccca aatgacgcag cttcagaagt tgatattca cagtaaggac ttatatttt	720		

	ataataatta ttatataatt ttctgacatg ttttggaaac ctc当地acat gtgattgcac	780
	cttc当地ttt tatgtctgg tcagaaactg ataagtttt acagtgtta ggtggatct	840
	ttgatgc当地 cagtgctttc taatgtttc attttgaaa gtaatgttt agaagaaat	900
	atctgattaa atttatactt tatcttaca aaagtcaa at gcttctgta tcaattgc当地	960
5	tttgaatat ggcaagaaca tgcttcaga atttgttcat acaatgttt ct当地tattta	1020
	ttatgttagaa caaataccta atacttggt cacctttt agtggacacc tctcacagct	1080
	tttcaagtaa gtgatgcaat ttgtacatt tgaagatgt gttccagaaa cctttctcc	1140
	t当地attcttca atgtaccac tcaaactggt atcaccaaag atctccatct gattgaaaaa	1200
	aagctgc当地 aagtatgctt atttatactt accatacatg atttatactg tttatagta	1260
10	caatgctt at ttagctaac catacataat ttattctgt tttcttagtac attttgtg	1320
	ccc当地gacca taaatgatcc ttcttttac agtggtccg aagatcccag gggcctacg	1380
	gtccc当地gggtt cctaccaa at gatgggatt atcaatgaaa cacataatgg gaggcatgct	1440
	t当地agtc当地ca aggttggta gtttggta gacttgggt gcaaaacacc aattcacag	1500
	gtattatgtgg ccaacaatgg aatggcagca gctaagttca tgc当地gggt gccaacatgg	1560
	gctaatgata ct当地ggatc agagaaggca atcagctga tagctatggc aactccggag	1620
	gatctgagga taaatgc当地 gcaatc当地atgca atgccc当地tca aattttaga ggtacctggt	1680
15	ggaacaaaca acaacaacta tgcaatgtc caactcatag tggaggtag ttagctcat	1740
	ccctcaacac aacatttcg ttcttattta agttagggaa aaatcttac gaccctccaa	1800
	tttctgaaaca tccaattttc accatcaact gcaatcacag atagcagaga gacagggt	1860
	ttctgctgtt tggc当地gggtt ggggtcatgc atctgagaat cctgaaacttc cagatgc当地	1920
	gactgcaaaa ggaatttggtt ttcttggcc accagcatca tcaatgcatg cattaggaga	1980
20	caagggttggc tca gctctca ttgctcaagc agctggagtt ccaacactt cttggagtg	2040
	atcacatgtg agccttgc当地 tcttattttt agcttacat cttatctt cgggtatgca	2100
	ttatccaaat gacactaaac catagggtga agttcctctg gagggttgc当地 tgactcaat	2160
	acctgatgag atgtatagaa aagcttggta tactaccaca gaggaa gacag tgc当地	2220
	tcaagggttggt ggttattctg ccatgattaa ggc当地tctgg ggtgggtt gtaaaggaa	2280
25	aaggaagggtt tggcttctt ttagttatca agagattgtt tggattgca gttttagt	2340
	cccatagtttta actctggtct ttcttacatg agtaactcaa ctttctgca gtttcatata	2400
	gatgatgagg ttaggacatt attaagcaa gtc当地aggcg aagtacctt gtttccaaata	2460
	tttatcatga ggttagtgc tcaagggtggg cttttatgg aagtttacacc tt当地ccctta	2520
	atgtttaggtt atccggagt tattatgtt atgttctgt tggattgatct gtaaattatt	2580
	gaaatttacc tccatttggg tccatgatca gcaatc当地tac aattctat atgttata	2640
30	ctttataat actaggattt agggtatctt atatgttta tacatgtt atagatttca	2700
	tttgaaccc tattgaagac atcctgattt ttgtctttagt tagagtgc当地 atcttgaagt	2760
	tca gtttgc当地 tggatcaat atggcaacgt agcagcactt cacagtc当地 attgc当地gt	2820
	acaacggc当地 caccaaaagg tctgctgtct cagtttatac acccctctgat gatctact	2880
	tcttgc当地tgc tgc当地gggtc agaggaataa tggatgtt ctagtgc当地 gataatcgag	2940
35	gaaggaccag ttactgtgc tccctgtgag actgtgaaag agcttgc当地 ggc当地cacgg	3000
	aggcttgc当地 aagctgtggg ttatgttggt gctgctactg ttgatatactt ttagc当地atg	3060
	gaaacttggg aatattttt tctggaaactt aatccatcgcc tacagggtgg cttcttgc当地	3120
	attcttc当地ggg aattaatttc tggatgttcc atgatttaca ttgtcaatg gtc当地cacagg	3180
	ttgagcatcc tgc当地acttggg tggatagctg aagtttacc gcttgc当地tctg caagggttgc当地	3240
	tttggatggg tataccccctt tggc当地tcc caggtaatgc ttcttcatat agtttcttgc当地	3300
40	ctttgttaat tgaatgagct ttatatacaga ccatgagaca cattctactg ttaatttca	3360
	gtatccctg actgtttagt tttagagata cagagatgtt tcaatc当地tac atgttatctc	3420
	ctcaaggact gtaaaaatcc tataattaaa ttctgaaaaa ttgttcttta agtgc当地	3480
	aaaaatcttctt aaattatctt cctgtatatac gatgtc当地ggc gtttctacgg aatgtacccat	3540
	ggaggaggct atgaccttgg gaggaaaaca gcaatc当地tac cgactccatt taacttgc当地	3600
	gaatgtat ctaatggcc aaaaggccac tgc当地tagtgc ttgatataac tagc当地ggat	3660
45	ccagatgatg ggtttaagcc tactgttggg aaagttaaagg tgc当地gttcc ttagtgc当地	3720
	tgtatgattt gaaacacattt ctatatttgc gctactgtt gactgttcc atggttctct	3780
	tattttc当地ggg agataatgtt caagatggg ccaatgtt gggcttattt ctcaatgtt当地	3840
	gtactctca atattgttgc actgtccatc tatttgc当地 gtc当地taatc ttgttgc当地	3900
	attgtttagt ttcaatattt tggatgttgc tttgttgc当地 tggatgttcc tcttgc当地	3960
50	ctggatgttggg catccatgaa ttctgttcc gtc当地tcc ctagtgc当地 tattgtt当地	4020
	atattgttctt tgc当地tcc acttttggg tttgttcc local gatatttgc当地 tccatc当地	4080
	gacatgttggg tgc当地atggg actactatg cgccatc当地tac aactaccatg gcttgc当地	4140
	taaaagggat tcaatgtt当地 ggagaatcc atcaatgtt当地 agactacaca gttgatctt	4200
	taaatgtt当地 gactaaatgtt ctgcttattt aaccttgc当地 tttgttcc local taatgtt当地	4260
	tttagtgc当地 tactgtt当地 cttatccatc atgccc当地tcc tggatgttcc aattcaggcc	4320
55	tcaatgtt当地 gagaatataa gattcatact gttggctgg ataccaggat accatgc当地	4380
	gttcaatgtt当地 agaggccctcc atggttattt tcaatgtt当地 gaggggtt当地 atatgtt当地	4440
	caaactatgc cactcattatg aagcaatgtt当地 ggaaaacatg atcaatgtt当地	4500

5	cgtcttattt aaattttattt atttttgtgc tgcatggaaaac agtaactgcc aacacggcca ctgtttctga ttatgttggt tatcttacca agggccagat tccacccaaag gtactattct gttttttcag gatataatg ctgtttgaat gtaaaaacca ttgaccataa atccttgttt gcagcatata tcccttgtct atacgactgt tgcttgaat atagatggaa aaaaatatac agtaagtgtg acattcttaa tggggaaact taatttggg taaataatca atatcatatt gactcggttga tgctgcatca tagatcgata ctgtgaggag tggacatggt agctacagat tgcgaatgaa tggatcaacg gttgacgaa atgtacaaat attatgtat ggtgggcttt taatgcaggtaat atatcttct tccttagttaa agaagatata tcttgttcaa agaattctga ttattgtatct ttaatgttt tcagctggat gggaaacagcc atgtatattt tgcgttggaa gaggccagtg gtacacgact tcttattgtat gggaaagacat gcatgttaca ggttatgata gccttggcttct ttttagttct agtcacgggt tttgcttgcatttgtat tcttatttata gcattcacta attactataat tagtttgcatttgc catcaagttttaa atatggaaact tctttcttgc agaatgacca tgacccatca aagttattag ctgagacacc atgcaaaactt cttcggttct tgggtgtctga tgggtgtctat gttgatgttgc atgtaccata tgcggaaagtt gaggttatgaa agatgtgtcat gccccttta tcacccgtt ctgggtgtcat acatgttgc atgtctgagg gccaagcaat gcaggtacat tcctacattt cattttttgt gctgtgttgc catgaacatt tcaagtaaat acctgttact tgggttattat tctaggctgg tgatcttata gctaggctgg atcttgcatttgc cccttctgtt gttttagat gggccgtt cgaagatact ttccacaaa tgggtctccc tattgtgtct tctggccaag ttccacaaaattt atgtgttgc atgtgttgc cttgcatttgc gatcatttgc gggatgttgc atgtatatttgc caaggttac acatgttgc cttgcattttt ctttgcatttgc tcatgttcatc ttatgttcatc catgttgc ttttgc gattccgctg tgggttgcatttgc cagatttttgc tcatgttcatc aatgggttgc aacataaaa tgctgttataat ttcatttgcatttgc ttttacttgc atttcttataat ttgttgc ttttgc tttagtacaat ccataatcttgc tcttgcatttgc caacagttaa tccatttataat atatttttgc taaactgttgc tcttgcatttgc gttgttgcatttgc agttgttgcatttgc ctgccttgc acctccggagc ttccttgcatttgc gcaatgggat gggatgttgcatttgc ctgttttgcatttgc aactagactt ccaagaaatc ttaaaaatgttgc ggtatatttgc ggttgcatttgc atagcttgc tcatgttgc ttttgc catttcgcatttgc cataggttgc ttttgcatttgc caagtttgcatttgc ttttgc ctggccatttgc acatgttgcatttgc tgggttgcatttgc ttttgc ttcatgttgcatttgc gttgttgcatttgc acttgcatttgc ttttgc ttatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc agtaaaaattt gactctggat ttttgcatttgc ttttgcatttgc ttttgc tgagggttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc agcgggttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gcttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tcactttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tggaaatttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gtcaaaaatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ctatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gttacgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ctacagaaggat ttttgcatttgc ttttgcatttgc ttttgc cttcacttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gagttgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tcctgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc caagggttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc tctgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ttgaacaaat ttttgcatttgc ttttgcatttgc ttttgc agatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc agagcatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ttgattgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc accagggttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gcatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc cattaaat ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc aacatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc taaaggacat ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ctgaagggttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ccatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc agacatgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc ataataat ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gttataat ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc gatgtgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgc	4560 4620 4680 4740 4800 4860 4920 4980 5040 5100 5160 5220 5280 5340 5400 5460 5520 5580 5640 5700 5760 5820 5880 5940 6000 6060 6120 6180 6240 6300 6360 6420 6480 6540 6600 6660 6720 6780 6840 6900 6960 7020 7080 7140 7200 7260 7320 7380 7440 7500 7560 7620 7680 7740 7800 7860 7920 7980 8040 8100 8160 8220 8280
10		
15		
20		
25		
30		
35		
40		
45		
50		
55		

	catgcctctg	gtgtaaaaac	aataagttc	attgttcaaa	gagatgaagc	acggatgaca	8340
5	atgcgtcgta	ccttccttgc	gtctgtatgaa	aagctttctt	atgaggaaga	gcacaattctc	8400
	cggcatgtgg	aacctccctc	ttctgcactt	cttgagttgg	tacgtgatat	catcaaaaatg	8460
	ataatgtttt	ggtatggcat	tgattatctt	ctatgctctt	tgtatttattt	cagcctatttgc	8520
	tggatacagg	acaagttgaa	agtgaaagga	tacaatgaaa	tgaagtatac	cccatcacgg	8580
	gatcgtaat	ggcatatcta	cacacttaga	aatactgaaa	acccc当地	gttgcaccgg	8640
	gtattttcc	gaacccttgc	caggcaaccc	agtgtatcca	acaagttttc	ttcgggcccag	8700
10	attggtgaca	tggaaagttgg	gagtgtgaa	gaacctctgt	catttacatc	aaccagcata	8760
	ttaagatctt	tgtatgactgc	tatagaggaa	ttggagcttc	acgcaattag	aactggccat	8820
	tcacacatgt	atttgcatgt	attgaaagaa	caaagtttc	tttgatcttgc	tccagtttca	8880
	gggtaagtgc	gcataatttct	ttttggaaac	atatgttgc	ttatgagggtt	ggtcttctca	8940
15	atgatcttct	tatcttactc	aggaatatac	ttttggatgt	tggtcaagat	gaagctactg	9000
	catattcact	tttaaaaagaa	atggctatga	agatacatga	acttgttgg	gcaagaatgc	9060
	accatcttc	tgtatgcca	tggaaagttgaa	aacttaagtt	ggactgcgt	ggtcctgcca	9120
	gtggtaacctg	gaggattgt	acaaccaatg	ttactagtca	cacttgcact	gtggatgtaa	9180
20	gtttaatcc	ctagcatttt	ttttcttgc	gaaaagcatg	tgattttaag	ccggctggtc	9240
	ctcataccca	gacctagtga	tctttatata	gttagacat	ttttcttaact	gcttttaatt	9300
	gttttagatc	taccgtgaga	tggaaagataa	agaatcacgg	aagtttagtat	accatcccgc	9360
	cactccggcg	gctggcctc	tgcattgtgt	ggcactgtaat	aatccatatc	agcctttgag	9420
25	tgtcattgtat	ctcaaacgct	gttctgttag	gaataataga	actacataact	gctatgattt	9480
	tccactgggt	agttgactgc	tcccttatata	tcaatgcatt	accatagcaa	attcatattc	9540
	gttcatgttg	tcaaaaataag	ccgatgaaaa	ttcaaaaactg	taggcatttgc	aaactgcagt	9600
	gaggaagtca	tggcctctc	gtacctctgg	tgcttctaaa	ggtggtaaa	atgccccatg	9660
30	ttatgttaaa	gctacagagt	tggtatttgc	ggacaaaacat	gggtcatggg	gcactcctt	9720
	agttcaatgc	gaccggcctg	ctgggctcaa	tgacattgg	tatgttagctt	ggaccttgaa	9780
	gatgtccact	cctgaatttc	ctagtgtag	ggagattatt	tttggtaaa	atgatattac	9840
	gttcagagct	ggatcatttgc	gcccaaggaa	agatgcattt	tttggaaactg	ttaccaacct	9900
35	agcctgtgag	aagaaaacttc	ctcttattta	tttggcagca	aattcttgc	ctcgaatttgc	9960
	catagcagat	gaagtgaaat	tttgctccg	tgggggtgg	tctgtatgt	gcagccctg	10020
	acgtgggtt	cagtcattt	atctaagcga	agaagactat	gctcgatttgc	gcacttctgt	10080
	catagcacat	aagatgcage	tagacagtgg	tggaaatttagg	tgggttatttgc	attctgttgc	10140
40	gggcaaggaa	gttggacttg	gtgtggagaa	tatacatgg	agtgtctgt	ttggccagtgc	10200
	ttattcttag	gcatataagg	agacatttac	attacatttgc	gtgactggaa	gaactgttgg	10260
	aataggagct	tatcttgctc	gacttgcatt	ccgggtgcata	cagcgttgc	accagcctat	10320
	tattcttaca	ggctattctg	cactgaacaa	gtttcttggg	cgggaaagtgt	acagctccca	10380
45	catgcagtttgc	gggtggccca	aaatcatggc	aactaatgg	gttgcatttc	ttactgtttc	10440
	agatgaccctt	gaaggcggtt	ctaataatatt	gaggtggctc	agttatgttc	ctgcctacat	10500
	tggtggacca	cttccagtaa	caacaccgtt	ggaccccaccc	gacagaccc	ttgcatacat	10560
	tcctgagaac	tctgtgtatc	ctcgagccgc	tatccgttgc	gttgcatttgc	gccaaggggaa	10620
50	atggtaggt	ggatgttttgc	ataaagacag	tttgcatttgc	acatttgc	gttgggctaa	10680
	gacagtgggtt	actggcagag	caaagcttgc	tggaaatttgc	gtgggtgt	tagctgtgg	10740
	gactcagacc	atgatgcata	ctatccctgc	tgacccttgc	cagcttgc	cccgctgagca	10800
	atctgttcc	cgtgttgc	aagtgttgc	tccagatttgc	gcaaccaaga	ctgcgcaggc	10860
55	attgtctggac	ttcaaccgtc	aaggattacc	tctgttgc	ctcgctact	ggagaggctt	10920
	ctctgttgc	caaagagatc	ttttgttgc	aattcttgc	gttgcatttgc	ctattgttgc	10980
	gaaccttgc	acatacaatc	agcctgcctt	tgtctacatt	cccatgc	cagagctac	11040
	aggagggct	tgggttgc	ttgatagca	gataaacc	gaccgcatttgc	agtgctatgc	11100
	tgagaggact	gcaaaaaggca	atgttcttgc	accgcaagg	ttaatttgc	tcaagtttgc	11160
	gtcagaggaa	ctccaggatt	gcatgatgt	gttgcatttgc	acattaatttgc	atctgaaagc	11220
	aaaactcgaa	gtacaaata	aaaatgttgc	tgcgttgc	aaatcgcttgc	aagaaaat	11280
	agaagctcgaa	acaaaacatgt	tgtatgttgc	atatacttgc	attgcatttgc	ggtttgc	11340
	attgcattat	acatccctca	gaatggctgc	gaaagggtgt	attaagaaat	tttgcatttgc	11400
	ggaagaatca	cgatcttgc	tctataatgc	attacggagg	aggatcttgc	aggatgttgc	11460
	tgcataatgc	attagatgttgc	tagcattgt	gcagtttgc	caccaacc	caatcgatgt	11520
	gatcaatgc	tggtagtttgc	tttgcatttgc	agctgtatgc	gatgtatgt	atgttttgc	11580
	tgcttgcatt	gataacccttgc	aaaacttgc	ggattatatttgc	caatatgc	aggctcaatgc	11640
	agatccaa	tccctcttca	gtcttgc	ttccagtttgc	gatttgc	ccctggcc	11700
	gggtcttgc	atgttacttgc	ataaggttgc	tagcttacttgc	atgtttatgc	aaatttgc	11760
	tcattacata	tggcttgc	actatgttgc	caaataatgc	ttataatttgc	aatcgatgttgc	11820
	tttgcatttgc	tatgtatgttgc	tgaaatttgc	tttgcatttgc	acttatttgc	atggatcc	11880
	ctagaagagc	tcaacttgc	gaagaaatca	ggaaggttgc	tgggttgc		11927

<211> 4203  
<212> DNA  
<213> SpCas9 DNA sequence (DNA sequence)

5	<400> 15	
	atggctccga agaagaagag gaagggtggc atccacgggg tgccagctgc tgacaagaag	60
	tactcgatcg gcctcgatat tgggactaac tctgttggct gggccgtat caccgacgag	120
	tacaagggtgc cctcaaagaa gttcaaggtc ctgggcaaca ccgatcgca ttccatcaag	180
	aagaatctca ttggcgctct cctgttcgac agcggcgaga cggctgaggc tacgcccgtc	240
	aagcgccccg ccccgaggcg gtacacgcg aggagaatac gcatactcta cctgcaggag	300
10	attttctcca acgagatggc gaagggttgac gattcttct tccacaggct ggaggagtca	360
	ttcctcggtgg aggaggataa gaagcacgag cgccatccaa tcttcggcaa cattgtcgac	420
	gagggtgcct accacgagaa gtaccctacg atctaccatc tgccgaaagaa gtcgtggac	480
	tccacagata aggccgaccc tcgcctgatc tacctcgctc tggcccat gattaagttc	540
	agggggcatt tcctgatcga ggggatctc aacccggaca atagcgatgt tgacaagctg	600
15	ttcatccagc tcgtcgagac gtacaaccag ctcttcgagg agaaccctat taatgcgtca	660
	ggcgtcgacg cgaaggctat cctgtccgct aggctctcga agtctcgccg cctcgagaac	720
	ctgatcgccc agctgcccgg cgagaagaag aacggcctgt tcgggaatct cattgcgtc	780
	agcctggggc tcacgccccaa cttcaagtcg aatttcgatc tcgctgagga cgccaaagctg	840
	cagctctcca aggacacata cgacgtgac ctggataacc tcctggccca gatcgccgt	900
20	cagtacgccc acctgttccct cgctgccaag aatctgtcggt acgccccctt cctgtctgtat	960
	attctcaggg tgaacaccga gattacaag gctccgctct cagcctccat gatcaagcgc	1020
	tacgacgagc accatcagga tctgaccctc ctgaaggcgc tggtcaggca gcagctcccc	1080
	gagaagtaca aggagatctt cttcgatcag tgaagaacg gctacgtgg gtacattgac	1140
	ggcggggcct ctcaggagga gttctacaag ttcatcaagc cgattctgga gaagatggac	1200
25	ggcacggagg agctgctggt gaagctcaat cgcgaggacc tcctgagga gcaaggccgaca	1260
	ttcgataacg gcagcatccc acaccagatt catctcgggg agctgcacgc tattctgagg	1320
	aggcaggagg acttctaccc tttcctcaag gataaccgcg agaagatcga gaagattctg	1380
	actttcagga tcccgacta cgtcgccccca ctcgcttaggg gcaactcccg cttcgcttgg	1440
	atgaccgcga agtcagagga gacgatcacg ccgtgaaact tcgaggaggt ggtcgacaag	1500
30	ggcgtctagcg ctcaagtcgtt catcgagagg atgacgaatt tcgacaagaa cctgccaaat	1560
	gagaagggtgc tccctaagca ctcgctctg tacgagtaact tcacagtcta caacgagctg	1620
	actaagggtga agtatgtgac cgaggccatg aggaaggccgg ctttcctgtc tggggagcag	1680
	aagaaggccca tcgtggaccc cctgttcaag accaaccggaa aggtcacggg taagcagctc	1740
	aaggaggact acttcaagaa gattgagtgc ttcgattcgg tcgagatctc tggcggtttag	1800
	gaccgcttca acgcctccct ggggacttac cacgatctcc tgaagatcat taaggataag	1860
	gacttcctgg acaaacgagga gaatgaggat atcctcgagg acattgtgct gacactca	1920
35	ctgttcgagg accgggagat gatcgaggag cgccctgaaga cttacgccccca tctttcgat	1980
	gacaagggtca tgaagcagct caagaggagg aggtacaccg gctggggag gctgaggcagg	2040
	aagctcatca acggcatttcg ggacaagcag tccggaaaga cgatcccgaa cttccctgaag	2100
	agcgtatggct tcgcgaaccg caatttcatc cagctgattc acgtacgacg cctcacattc	2160
	aaggaggata tccagaaggc tcaggtggc gggcaggggg actcgctgca cgagcatatc	2220
40	gcgaacctcg ctggctcgcc agtatacaag aaggggattc tgcagaccgt gaagggttgt	2280
	gacgagctgg tgaagggtcat gggcaggcac aacgctgaga acatcgatc ttagatggcc	2340
	cgggagaatc agaccacgcga aagggggcag aagaactcac gcgagaggat gaagaggatc	2400
	gaggaggcgtttaaggagat ggggtcccgat atcctcaagg agcaccggg gggaaacacg	2460
	cagctcgaga atgagaagct ctacctgtac tacctccaga atggccgcga tatgtatgt	2520
	gaccaggagc tggatattaa caggctcagc gattacgacg tcgatcatat cttccacag	2580
45	tcatttcgttca aggtacttc cattgacaac aagggtctca ccaggctggca caagaaccgg	2640
	ggcaagtctg ataatgttcc ttcatggggat gtcgtttaaga agatgaagaa ctactggcgc	2700
	cagctctgttca atgccaagct gatcacgcg cgaagggtcg ataaccctcac aaaggctgag	2760
	agggggccggc tctctgagct ggacaaggcg ggcttcatca agggcagct ggtcgagac	2820
	cgccagatca ctaaggcacgt tgcgcgat ctcgactcac ggatgaaacac taagtacgt	2880
	gagaatgaca agtgcaccccg cgagggtggaa gtcacccatc tgaaggtaaaa gtcgtctcc	2940
50	gacttcaggaa aggatttcca gttctacaag gttcgggaga tcaacaattt ccaccatgcc	3000
	catgacgcgt acctgaacgc ggtggctggc acagctctga tcaagaagta cccaaagctc	3060
	gagagcgagt tcgtgtacgg ggactacaag gttacgatg tgaggaagat gatcgccaa	3120
	tcggaggcagg agattggcaa ggctaccggc aagtacttct tctactctaa cattatgaat	3180
	ttcttcataaga cagagatcac tctggccat ggcgagatcc ggaaggccccc cctcatcgag	3240
	acgaacggcg agacggggga gatcggttgg gacaaggggca gggatttgc gaccgtcagg	3300
55	aagggttcttccatgccaca agtgaattatc gtcagaagaagaa cagaggctca gactggcg	3360
	ttcttcataagg agtcaatttccat gcttaaggcgg aacagcgaca agtcatcgcc ccccaagaag	3420
	gactgggatc cgaagaagta cggcggttcc gacagccccca ctgtggctctg	3480

	gttgtggcga aggtttagaa gggcaagtcc aagaagctca agagcgtgaa ggagctgctg	3540
	gggatcacga ttatggagcg ctccagctc gagaagaacc cgatcgattt cctggaggcg	3600
	aagggttaca aggaggtgaa gaaggacctg atcattaagc tccccaaagta ctcactcttc	3660
	gagctggaga acggcaggaa gcggatgctg gttccgcgtg gcgagctgca gaaggggaaac	3720
5	gagctggctc tgccgtccaa gtatgtaac ttcttctacc tggcctccca ctacgagaag	3780
	ctcaaggcga gcccccggaa caacgagcag aagcagctgt tcgtcgagca gcacaaggcat	3840
	tacctcgacg agatcattga gcagattcc gagttctcca agcgcgtgat cctggccgac	3900
	gcgaatctgg ataaggcttct ctcgcgtac aacaaggcacc gcgacaagcc aatcaggggag	3960
10	caggctgaga atatcattca tcttttacc ctgacgaaacc tcggcgcccc tgctgtttc	4020
	aagtacttcg acacaactat cgatcgaaag aggtacacaa gcaactaagga gtcctggac	4080
	gcgaccctca tccaccagtc gattaccggc ctctacgaga cgcgcatcga cctgtctcag	4140
	ctcggggcg acaagcggcc agcggcagc aagaaggcgg ggcaggcgaa gaagaagaag	4200
	tga	4203
15	<210> 16	
	<211> 4203	
	<212> DNA	
	<213> NGA-Cas9 DNA sequence (DNA sequence)	
	<400> 16	
20	atggctccga agaagaagag gaagggtggc atccacgggg tgccagctgc tgacaagaag	60
	tactcgatcg gcctcgatat tggactaac tctgttgct gggccgtgat caccgacgag	120
	tacaagggtgc cctcaaagaa gttcaaggc ctggcaaca ccgatcgca ttccatcaag	180
	aagaatctca ttggcgctc cctgttcgac agcggcggaa cggctgaggc tacgcggctc	240
	aagcgcaccg cccgcaggcg gtacacgcg aggaagaatc gcacatgtcta cctgcaggag	300
	attttctcca acgagatggc gaagggtgac gatttttct tccacaggct ggaggagtca	360
25	ttcctcggtt aggaggataa gaagcacgag cgccatccaa ttttcgcaaa cattgtcgac	420
	gaggttgccct accacgagaa gtaccctacg atctaccatc tgccgaagaa gtcgtggac	480
	tccacagata aggccgaccc cccgctgatc tacctcgctc tggccacat gattaagttc	540
	agggggcatt tcctgatcga ggggatctc aacccggaca atagcgatgt tgacaagctg	600
	ttcatcgacg tcgtcgagac gtacaaccag ctcttcgagg agaaccctat taatgcgtca	660
	ggcgtcgacg cgaaggctat cctgtccgct aggtctctga agtctcgcc cctcgagaac	720
30	ctgatcgccc agctgccggg cgagaagaag aacggcctgt tcggaatct cattgcgctc	780
	agcctggggc tcacgccccaa ttcaagtcg aatttcgatc tcgctgagga cgccaagctg	840
	cagctctcca aggacacata cgacgtgac ctggataacc ttctggccca gatcgccgat	900
	cagtacgccc acctgttccct cgctgccaag aatctgtcg agccatctt cctgtctgat	960
	attctcaggg tgaacaccga gattacaag gtcggctct cagcctccat gatcaagcgc	1020
35	tacgacgagc accatcagga tctgaccctc ctgaaggcgc tggtcaggca gcagctcccc	1080
	gagaagtaca aggagatctt ttctcgatcag tcgaagaacg gtcacgctgg gtacattgac	1140
	ggcggggcct ctcaggagga ttctacaag ttcatcaagc cgattctgga gaagatggac	1200
	ggcacggagg agtgcgtggt gaagctcaat cgcgaggacc ttctgaggaa gcagcggaca	1260
	ttcgatcaccg gcacgtatccc acaccagatt catctcgaaa agtcgcacgc tatcctgagg	1320
	aggcaggagg acttctaccc ttcttcaga gataaccgcg agaagatcga gaagattctg	1380
40	actttcagga tcccgacta ctcgcggccca ctcgcctagg gcaactcccc ctcgccttgg	1440
	atgaccctcgca agtcagagga gacgatcaccg ccgtggaaact tcgaggaggt ggtcgacaag	1500
	ggcgtctagcg ctcagtcgtt catcgagagg atgacgaaatt tcgacaagaa ctcgcctaaat	1560
	gagaagggtgc tccctaagca ctcgcctctg tacgagtact tcacagtcta caacgagctg	1620
	actaagggtga agtatgtgac cgaggggatc agaaggccgg ttccctgtc tggggagcag	1680
	aagaaggcca tctgtggaccct cctgttcaag accaaccggc aggtcacggt taagcagctc	1740
45	aaggaggact acttcaagaa gattgagtgc ttcgattcgg tcgagatctc tgctgttgag	1800
	gaccgttca acgcctccct ggggacttac cacgatctcc tgaagatcat taaggataag	1860
	gacttcctgg acaacgagga gaatgaggat atcctcgagg acattgtgct gacactca	1920
	ctgttcgagg accggggat gatcgaggag cccctgaaga cttacgcccc tctcttcgat	1980
	gacaagggtca tgaagcagct caaggaggagg aggtacacccg gtcggggag gctgagcagg	2040
50	aagctcatca acggcattcg ggacaaggcag tccggaaaga cgatccctcgat ctcctcgaa	2100
	agcgatggct tcgcgaaaccg caatctcgat cagctgattc acgtacgacag ctcacattc	2160
	aaggaggata tccagaaggc tcaggtgagc gcccagggggg actcgctgca cgacgatata	2220
	gccaacctcg ctggctcgcc agtatacaag aaggggatc tgcagaccgt gaagggttg	2280
	gacgagctgg tgaagggtcat gggcaggcgc aacccctgaga acatcgatcat tgagatggcc	2340
	cgggagaatc agaccacgca gaaggggccag aagaactcac gcgagaggat gaagaggatc	2400
55	gaggaggcga ttaaggagct ggggtcccaatcctcaagg agcaccctggt ggagaacacg	2460
	cagctcgaga atgagaagct ttcacgttac tacctccaga atggcccgca tatgtatgt	2520
	gaccaggagc tggatattaa caggctcagc gattacgacg tcgatcatat cttccacac	2580

	tcattcctga aggatgactc cattgacaac aaggtcctca ccaggtcgga caagaaccgg	2640
5	ggcaagtctg ataatgttcc ttcagaggag gtcgttaaga agatgaagaa ctactggcgc	2700
	cagctcctga atgccaagct gatcacgcag cggaaaggctcg ataacctcac aaaggctgag	2760
	agggggcgggc tcctctgagct ggacaaggcg ggcttcatca agaggcagct ggtcgagaca	2820
	cgccagatca ctaagcacgt tgccgagatt ctgcactcac ggatgaacac taagtacgat	2880
	gagaatgaca agctgatccg cgaggtgaag gtcatcaccc tgaagtc当地 gctcgctcc	2940
10	gacttcagga aggatttcca gttctacaag gtcgggaga tcaacaatta ccaccatgcc	3000
	catgacgcgt acctgaacgc ggtggcggc acagctctga tcaagaagta cccaaagctc	3060
	gagagcgagt tcgtgtacgg ggactacaag gttacgatg tgaggaagat gatgc当地	3120
	tcggagcagg agattggcaa ggctaccggc aagtacttct tctactctaa cattatgaat	3180
	ttcttcaaga cagagatcac tctggccaat ggcgagatcc ggaagc当地 cctcatcgag	3240
	acgaacccggc agacggggga gatcggttgg gacaaggcg gggatttgc gaccgtcagg	3300
15	aaggttctct ccatgccaca agtgaatatac gtcaagaaga cagaggc当地 gactggcggg	3360
	ttctcttaagg agtcaattct gcctaaggcg aacagcgaca agctcatcgcc ccccaagaag	3420
	gactggatc cgaagaagta cggcgggttc gtcagcccc ctgtggcctc ctcggc当地	3480
	gttgtggcga aggttgagaa gggcaagtcc aagaagctca agagc当地 gtagctgctg	3540
	gggatcacga ttatggagcg ctccagcttc gagaagaacc cgatcgat当地 cctggaggcgg	3600
	aagggtaca aggaggtgaa gaaggacccg atcattaagc tccccaaagta ctcactcttc	3660
	gagctggaga acggcaggaa gcggatgtcg gttccgctg gcgagctgca gaaggggaaac	3720
20	gagctggctc tgccgtccaa gtatgtAAC ttcctctacc tggc当地 ctaggagaag	3780
	ctcaaggc当地 gccccgagga caacgagcag aagcagctgt tcgtcgagca gcacaagcat	3840
	tacctcgacg agatcattga gcagattcc gagttctccaa agcgc当地 gtat cctggccgac	3900
	gcgaactctgg ataaggtccct ctccgc当地 taccaaggcc accgacaagcc aatcaggggag	3960
	caggctgaga atatcattca tctcttcacc ctgacgaaacc tcggc当地 cccccc tgctgtttc	4020
25	aagtacttcg acacaactat cgatcgcaag cagtaccggc gcactaaggaa ggtc当地 ctgggac	4080
	gcgaccctca tccaccaggc gattaccggc ctctacgaga cgccgcatcgaa cctgtctcag	4140
	ctcgccccggc acaagcggcc acgccc当地 gag aaggcgg ggcaggc当地 gaagaagaag	4200
	tga	4203
	<210> 17	
	<211> 5860	
	<212> DNA	
30	<213> Oryza sativa L. OsBADH2 genome DNA sequence (DNA sequence)	
	<400> 17	
	atggccacgg cgatcccgca gcggcagctc ttctcgcccg gcgagttggcg cgccccccggc	60
	ctcgccgccc gcctccccgt cgtcaacccccc gcacccgagt cccccc当地 taccctcctc	120
35	ttcacccctt ccacccctctg ttctcgccctc tgattagcct ttttgttgg ttgtttgtt	180
	ctgtgtttt ttgcgtgtcg gtgc当地 gagatcccg gggcacggc ggaggacgtg	240
	gacgc当地 ggccggccgg gggggaggcg ctgaagagga accggggccg cgaactggc当地	300
	cgccgcccgg gcggcgtcccg ggcaagttac ctccgc当地 gcaaa tgccggccaa gtaggggtgg	360
	tgactacccc cccccccccc cccccccccc acgc当地 gaccccg cgtgc当地 gtgttcc	420
	ggggaggaggc tccgc当地 gtggc tctccagtag gtttttggc cccaaatcgca tgc当地 atgtatgct	480
40	ctagtttaa gtttgc当地 gct taaaatccctc aagggtttt tag tttgcaacca aatccttatt	540
	ttagctc当地 gg tataaggcccc ccatatgtatg tgc当地 gtggc ggc当地 catcgaa gtgc当地 atatcc	600
	tctgttctgg actaggaatt gccc当地 ataggatg tgc当地 gagat ttgc当地 gagat ctgc当地 ttctgt	660
	ttggaataag ttggaagcat ggctgatggt gtatctggat gctgttttgg tgggatattcg	720
45	tttcaagctc ttgttaattt atgggttcaa gccc当地 gagagg tgc当地 caacca caatgttata	780
	tggctcacgg ccatgggtgt gcacatgtatg ttggc当地 gca acaacaatgt tatattgttt	840
	gtgtgcttgg ttagttggca ggtcttagtc actaaatcac tattttgattt gtagctatgtta	900
	cttttgc当地 ttgacgatgg gactggatta ctgc当地 ctttggatggt gttgc当地 ctttggatgg	960
	gttgttatgg gc当地 tggatggatcc cttaatttc tagtgc当地 aaa tgc当地 atgtatgt	1020
50	atttctcaca gttttctctc tc当地 aggttata ttctcgat ttc当地 ctttcc taaaggattt	1080
	cttttccatg tattttctgg catatataagg ttattttat tattatctc cagaacaaga	1140
	ttacccat tattttat tatggatcac tagtgc当地 ac tttttggat gaaaaaccta ctactgtaaa	1200
	gtaaaacagt gaccgtgca cacttactt gactgtcaa accatcaatt ttctatgtaaa	1260
	gcaggggatg ctgc当地 ctttcc agtctaaatg acatgttactt actataactt tgc当地 ctttgc当地	1320
	tttggaaata tgc当地 ttttcatgg catatgtgg ctttgc当地 ctttgc当地	1380
55	tcacccccc ccttttctt cagataatcg agggaaatc tgactgtggct agactagaga	1440
	cgcttgatgg tggaaaggctt cttgatgc当地 ag cggatggc catggatgt ggc当地 ctttgc当地	1500
	ccactgtatg aatatgtatg tgc当地 tacaca gcaatcttc ctgc当地 acatgtatgt	1560
	atgatattcc atttgc当地 cggatgttgc tgatgttcc gtagtcttgg cagatcttgc	1620
	agaatccctg gacaaaaggc aaaatgc当地 acc ttttgc当地 ccaatggaaa actttaaatg	1680

	ctatcttcgg aaagagccta tcggtgttagt tgggttgatc acacccctgg attcacatt	1740
5	tttctctcat cctgcgctta tattttat tgaccacaagg atggactaa atagtagactag	1800
	taacatgcat atactgaatg agtttacaac ttacatgat ttttttgaac tatgaaaagtt	1860
	gaagacattt gagattttat ccctcttc tcgtcaaaac atattattgt ctccacaattt	1920
	gtacctagca gctactctct ccgtttcata ttataagtctg tttgactttt tcccttagtca	1980
	aatgtgtta agtttgcacca agtttataga aaaatttagc aacatctaaa atatcaaagt	2040
	catgttttag tgggtttca ggctctcatg taagcaattt tgatgtgccc tctccttct	2100
10	tcttaatata atgatacaca gctctgtgt atccaaggaa aatataatat atatataatg	2160
	atacacaccc ctccctccgtg ttaatgcacg tcattttgttc tgcccggtt caaatatcta	2220
	tttttctcat atgttgtcag catgattcac ttaatttagt atatagaaga tgccattatt	2280
	tatgtctgga atcttactgc agaaggaaa acaatttgata acggaattga ttgcattcta	2340
	atttgttgtt tctttgttat gttcttatcg acaattacaa atttgattct gagaatcatg	2400
15	ttcgggatgt gtatttctac tgcaggaact atcctctcct gatggcaaca tggaaaggtag	2460
	ctccctccct ggctgctggc tgcacagctg tactaaaacc atctgaattt gcttccgtgt	2520
	aagtttaaca tggtaacttg ttaatgtcat accatgcta gttgcaatga catttgattt	2580
	taaaatgttg tggcatgtcc atgctgcaag caatgttaatt tgaaatctct ctctatcatt	2640
	aattaccagg acttgtttgg agcttgcata tgggtgtaaa gaggttggtc tcccttcagg	2700
	tgtgtctaaac atagtgactg gattagtttca tgaagccggt gctccttgcatcacaccc	2760
	tggtagac aaggtacagg tattccctt gtaatcatgt atacccatc aatggaaatg	2820
20	atattctct caatacatatgg tttatgtttt ctgttaggtt gcatttactg ggagttatga	2880
	aactggtaaa aagattatgg cttagtgc tccatgtgtt aagggttggc tccaaatttc	2940
	tgtggatatt ttttgttctc tttctactaa ctctcttata tcaatttctca atgttgtccct	3000
	tttcttttaa ctcccttact ttttagaattt gtgatcaaga cactttgagc atcattctag	3060
	tagccagttc tattctgttt tttttttttt tatgggttgc tttttcttgc cagcctgttt	3120
25	cactgaaact tgggtggaaaa agtccctatag tgggttttgc tgatgttgc tttttttttt	3180
	gtacatgcca ctgtctatgc ttaactaattt ctgtgttgc ggtttttttt aagcactttaa	3240
	ctgagcttgc tgcgttagaccc cccaaagcccc tttttttttt cttttttttt ttgcagaaat	3300
	actgggttttttca gacgagcagg atgcaagaaa attaactact tttttttttt attagtattt	3360
	tttagaaagttt acacccatcaag gatttagcaat actttttttttaa aatgtgttatg tttttttttt	3420
	gatgtctgtt atttttttttgc gcatgttttttgc ttttttttttgc ttttttttttgc tttttttttt	3480
30	gttgaggatttataatgccc gccaacttgc taaagaaaaag agttgttttttgc tttttttttt	3540
	aacatgaata tatatgttatgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3600
	aggaaggactt gatgtgttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3660
	atttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3720
	atttatgttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3780
35	atttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3840
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3900
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	3960
40	atataatccgtt caatcataat ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4020
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4080
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4140
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4200
45	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4260
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4320
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4380
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4440
50	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4500
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4560
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4620
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4680
55	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4740
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4800
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4860
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4920
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	4980
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5040
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5100
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5160
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5220
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5280
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5340
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5400
	tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc tttttttttt	5460

5	ttaagttgct cttcaacttcc acatcatctg atccatgact ttatatttga cccctttttt ttgcaaaaga aaggaaata cttaacaaa atttcctact gcaggagatc gatgccggaa ttatctgggt gaactgctcg caaccctgct tctgccaagc tccatgggc gggacaacagc gcagcgctt tggacgcgag ctggagaag gttggtagc acacaacaat ctcactttaa aacaccattt cgatcgctcg atgatctcg a cctgacatca tgccttggt atttcattc acttttcagg ggcattgaca actacctaag cgtcaagcaa gtgacggagt acgcctccga tgagccgtgg ggatggtaca aatccccttc caagctgtaa	5520 5580 5640 5700 5760 5820 5860
10	<210> 18 <211> 2192 <212> DNA <213> Oryza sativa L.OsSWEET14 genome DNA sequence (DNA sequence)	
15	<400> 18 atggctggca tgcattttca gcatccctgg gccttcgcct ttggtctcct aggtgtgttg cctttgatct gatccaagga attctttga gaattaatct tgcatggta ttacttttg ttgttattat tctctacatt ttaatcatg tactttcca tggccactt ttgttgccaa agctactata ttttcctac caattcatcc aaaactacta tattatagca aggtagctag tggatcgact ttgcactttt ggatgcattt gtgagtggtt tacatttagag gggccatgca tagccaaata aattgtttgg caaaatatga ttcccttgc agttagcaca atattgtgat ctttcattcc tttactcttt ttgttccacc accctcatta ttgcaaaagcc cccttttta gaacccaaact aaagataaga aacattgtat tcatcattag aggagattag atatacatac atactacaca tgcatttgctg tgataaggctt agccaggcac agaaagagcc aagacaaaag gagctcaaaa gaaaaggggc tccaatatct tggccggcct gatcagtgtt acatgacgtc acccacacaa ttattaaaag gatagatgca tatgtgttgc ttacgtatct ctttgttca cggtgtgtg catggttcac tcgcttaatt ataacgtccg tcgcatttt tgcatgcact tcgttttttag ttaatcaaatt taaacaagat cctaacatat gatttctatt ttattcttctt tgcaggcaac atcatcttcc tcatgaccta cttggccccca ctgtaagtcc cacacatccg gatattttct tttcttttagc gaaattttac attgctccag aaagtacctt ttacactaaa atttcttata caccaccata aaattaatgt ggtactattt gtaacctatc aaaaatgata aaattacctt tttcttcaga taaaaacact ccccgctctt gtaactaa gatgttagaca accaatacac atgtcttagag ctaatcttat caattaacat tgcttagaaa tacaaggaaa aatataaaaga attaaccaat tgcttgatg tttgcaggc cgacgttcta caggatctac aagagcaagt cgacgcaggg gttccagtcg gtaaccctacg tggggcgct gttcagcg atgctgtgga tctactacgc gctgctcaag tccgacgagt gcttcctcat caccatcaac tccgctggct ggtcatcga gaccatctac atcgccgtct acctcgctta cggcccccaag aaggccaaaga tggtcaccgc caagcttcc tcctctgtca acgtcgccgt cttggccctc atccctctcc tcacccttcc cctctccgc ggcgaccggc gcatcggtt tcttgggttgg gtctgcgttg gtttctccgt cagcgcttcc gtcgcccccc ttagcatcat cgtaaagaaac cctagaattt cctgtaaaca ataataactt tggctctctg caccactaa ggtgcacat atttacgcac cgcttttgc atctctgcag aggctgggt tgccgcacca gaggctggag ttcatggcgt tctcgctctc ttctccctc accatcagcg ccgtcgctg gttcctctac ggcctctca tcaaggacaa atatgtcgct gtgagtagct ccgattcgac ccgttcttct tcctaaattt tccgctgtc gatttaattt ctatttctaa ttgtgaaaga ttgtttaattt atagtgatta tggcttagctg ataaccgatt ttaatttctg atggtgatcc aaaatggaaac agcttcccaa cgtgctggc ttctccctcg ggtcatcca gatggggctg tacgcccatt acaggaactc gacgcggcaag gccgtgtga ccaaggaggt cgaggccggc acggccaccg ggtacgacga ccactccgc gccggcgtca aggagcactg cgtcaacatc gccaagctct ctggcccggt cgacgtcgta aagaccggc aggtgcaccc cgtcgacgtc gagtccccgc cggtcggccggc gccgcctgag gaggacaca aggccggcgc gcccacccg gccggcgtcg cggtcggccggc cgagaagaag gtagctgtcat ga	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2192
20	<210> 19 <211> 657 <212> PRT <213> Solanum tuberosum L. StALS2 protein amino acid sequence (amino acid sequence)	
25	<400> 19 Met Ala Ala Ala Ser Pro Ser Pro Cys Phe Ser Lys Asn Leu Pro Pro 1 5 10 15 Ser Ser Ser Lys Ser Ser Ile Leu Leu Pro Lys Ser Thr Phe Thr Phe 20 25 30	
30		
35		
40		
45		
50		
55		

## EP 3 889 266 A1

	His Asn His Pro Lys Asn Thr Ser Pro Leu His Leu Thr His Thr Gln			
	35	40	45	
	His His Ser Arg Phe Thr Val Ser Asn Val Ile Leu Ser Thr Thr Thr			
5	50	55	60	
	His Asn Asp Val Ser Glu Pro Glu Ile Phe Val Ser Arg Phe Ala Pro			
	65	70	75	80
	Asp Glu Pro Arg Lys Gly Cys Asp Val Leu Val Glu Ala Leu Glu Arg			
	85	90	95	
10	Glu Gly Val Lys Asp Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu			
	100	105	110	
	Ile His Gln Ala Leu Thr Arg Ser Asn Ile Ile Arg Asn Val Leu Pro			
	115	120	125	
	Arg His Glu Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ala			
15	130	135	140	
	Thr Gly Phe Pro Gly Val Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr			
	145	150	155	160
	Asn Leu Val Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Ile Pro Ile			
20	165	170	175	
	Val Ala Ile Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala			
	180	185	190	
	Phe Gln Glu Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His			
25	195	200	205	
	Asn Tyr Leu Val Met Asp Val Glu Asp Ile Pro Arg Val Val Arg Glu			
	210	215	220	
	Ala Phe Phe Leu Ala Lys Ser Gly Arg Pro Gly Pro Val Leu Ile Asp			
30	225	230	235	240
	Val Pro Lys Asp Ile Gln Gln Leu Val Ile Pro Asn Trp Asp Gln			
	245	250	255	
	Pro Met Arg Leu Pro Gly Tyr Ile Ser Arg Leu Pro Lys Leu Pro Asn			
	260	265	270	
	Glu Met Leu Leu Glu Gln Ile Val Arg Leu Ile Ser Glu Ser Lys Lys			
35	275	280	285	
	Pro Val Leu Tyr Val Gly Gly Cys Thr Gln Ser Ser Glu Glu Leu			
	290	295	300	
	Arg Arg Phe Val Glu Leu Thr Gly Ile Pro Val Ala Ser Thr Leu Met			
	305	310	315	320
	Gly Leu Gly Thr Phe Pro Cys Gly Asp Glu Leu Ser Leu Gln Met Leu			
40	325	330	335	
	Gly Met His Gly Thr Val Tyr Ala Asn Tyr Ala Val Asp Ser Ser Asp			
	340	345	350	
	Leu Leu Leu Ala Phe Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys			
	355	360	365	
	Leu Glu Ala Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp			
45	370	375	380	
	Ser Ala Glu Ile Gly Lys Asn Lys Gln Pro His Val Ser Ile Cys Ala			
	385	390	395	400
	Asp Ile Lys Leu Ala Leu Gln Gly Leu Asn Ser Ile Leu Glu Gly Lys			
	405	410	415	
	Glu Gly Lys Leu Asp Phe Ser Ala Trp Arg Gln Glu Leu Thr			
50	420	425	430	
	Glu Gln Lys Val Lys Tyr Pro Leu Asn Tyr Lys Thr Phe Gly Glu Ala			
	435	440	445	
	Ile Pro Pro Gln Tyr Ala Ile Gln Val Leu Asp Glu Leu Thr Asn Gly			
	450	455	460	
	Asn Ala Ile Ile Ser Thr Gly Val Gly Gln His Gln Met Trp Ala Ala			
55	465	470	475	480
	Gln Tyr Tyr Lys Tyr Lys Pro Arg Gln Trp Leu Thr Ser Gly Gly			
	485	490	495	
	Leu Gly Ala Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ala Val			
	500	505	510	
	Gly Arg Pro Gly Glu Ile Val Val Asp Ile Asp Gly Asp Gly Ser Phe			
	515	520	525	
	Ile Met Asn Val Gln Glu Leu Ala Thr Ile Lys Val Glu Asn Leu Pro			

	530	535	540	
	Val Lys Ile Met Leu Leu Asn Asn Gln His	Leu Gly Met Val Val Gln		
5	545	550	555	560
	Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg	Ala His Thr Tyr Leu Gly		
	565	570	575	
	Asn Pro Ala Asn Glu Glu Ile Phe Pro Asn Met	Leu Lys Phe Ala		
	580	585	590	
	Glu Ala Cys Gly Val Pro Ala Ala Arg Val Ser	His Arg Asp Asp Leu		
10	595	600	605	
	Arg Ala Ala Ile Gln Lys Met Leu Asp Thr Pro	Gly Pro Tyr Leu Leu		
	610	615	620	
	Asp Val Ile Val Pro His Gln Glu His Val	Leu Pro Met Ile Pro Ser		
	625	630	635	640
	Gly Gly Ala Phe Lys Asp Val Ile Thr Glu	Gly Asp Gly Arg Arg Ser		
	645	650	655	
15	Tyr			
	<210> 20			
20	<211> 1974			
	<212> DNA			
	<213> Solanum tuberosum L. StALS2 gene DNA sequence (DNA sequence)			
	<400> 20			
25	atggcggtcg catctccatc tccttgc tttt tccaaaaacc tacctccatc ttcatcaaaa tcttcatcc ttcttcccaa atctacccccc accttccaca atcaccggaa aaataccctca cccccttcacc ttacccacac ccaacatcat agccgttca ctgtctcaaa tgtcatccta tcaaccacga cccataaacga cgtttctgaa cccgaaatct tcgtttcacg ttccggccct gacgaaccca gaaagggttg tgatgttctt gtggaggcac ttgaaaggga aggggttaag gatgtatttg catacccccagg aggtgttcc atggagattc atcaggcttt gacacgttcc aatattatttc gtaatgtgct gccacgtcat gaacagggtg gtgtgttgc tgcaaggggt tacgcacggg ccactgggtt ccctgggtt tgcatgtca catctggtcc gggagctacg aatcttgtta qcggtcttgc ggatgtttt ttggatagta ttccgattgt tgctattacg ggtaagtgc cgaggaggat gattggact gatgcgttcc aggaaactcc tattgtttag gtaacgagat ccattacgaa gcataattat ctgttatgg atgttagagga tattccctagg gttgcgtg aagcgttttt tctagcgaaa tcgggacggc ctggaccggc tctgattgat gttcctaagg atattcagca acaattgggt atacctaatt gggatcagcc aatgagggtt 30		60	
	cctggttaca tctctagggtt gcctaaattt cctaattgaga tgcttttgg acaaattgtt aggctgattt cgagatcgaa gaagcctgtt ttgtatgtgg gtgggtgggt tacacaatcg agtggaggagc tgagacgatt tggagactt acgggtattc ctgtggcgag tactttgatg ggtctggaa ctttccatg tggggatgag ctttcttcc aaatgttggg tatgcatgg actgtgtatg ctaattatgc ggtggatagt agtgattgt tgcttgcatt tgggggtgagg 40		120	
	tttgatgatc gagttactgg taaattggaa gctttgtca gccgagcgaa aattgtccac attgatattt attcggtctga gattggaaag aacaagcaac ctcattttt cattttgtca gatatacaatg tggcattaca gggtttgaat tccatattgg agggtaaaga agtaagctg aagttggact tttctgctt gaggcaggag ttaacggagc agaaggtgaa gtacccattt aattataaga ctttgggtga agccatccct ccacaatatg ctattcaggt tcttgcatt ttaactaacg gaaatgccat tattgtactt ggtgtggggc aacaccaaattt gttttgttgc 45		180	
	caataactata agtacaaaaaa gccacccaa tggttgcatt ctgttggatt aggagcaatg ggattttgtt tgcctgcttgc tataggtgcg gctgttggaa gaccgggtga gattgttgggt gacattgtatg gtgacgggag ttttgcatttgc aatgtgcagg agttagcaac aattaagggt gagaatctcc cagttaaatg tattgtactt aataatcaac acttgggtt ggtgttgc tggggaggatc gattctataa ggctaacaga gcacacactt acttgggtt tcttgcatt 50		240	
	gaggaagaaa tcttccctaa tatgtctaaa ttgcaggagg cttgtggcgt acctgttgc agagtgtcac acaggatgt tcttagagct gcattcaaa agatgtttaga cacttctgg ccataacttgt tggatgttgcatt tgcatttttgcattt gatttttttgcattt ggtggcgctt tcaaagatgtt gattacggag ggtgtatgggaa gacgttccca ttgtt 55		300	
	<210> 21			
	<211> 1424			
	<212> DNA			
	<213> HBB gene DNA sequence (comprising intron) (DNA sequence)			

<400> 21

5	atgggcacc	tgactcctga	ggagaagtct	gccgttactg	ccctgtgggg	caaggtgaac	60
	gtggatgaag	ttgggttgta	ggccctggc	aggctgttat	caaggttaca	agacagggtt	120
	aaggagacca	ataaaaaactg	ggcatgtgga	gacagagaag	actcttgggt	ttctgtatagg	180
	cactgactct	ctctgcctat	tggctattt	tcccaccctt	aggctgtgg	tggcttaccc	240
	ttggaccagg	aggttctttg	agtccttgg	ggatctgtcc	actcctgtat	ctgttatggg	300
	caaccctaag	gtgaaggctc	atggcaagaa	agtgcctcggt	gccttagtg	atggcctggc	360
10	tcacctggac	aacctcaagg	gcaccttgc	cacactgagt	gagctgact	tgacacaagct	420
	gcacgtggat	cctgagaact	tcagggtgag	tctatggac	gcttgatgtt	ttctttcccc	480
	ttcttttcta	tggtaagtt	catgtcatag	gaaggggata	agtaacaggg	tacagtttag	540
	aatggaaaac	agacgaatga	ttgcatacgat	gtggaaagtct	caggatcggt	ttagtttctt	600
	ttatttgctg	ttcataaaca	ttgtttctt	ttgtttaatt	cttgcttct	tttttttctt	660
15	tctccgcaat	ttttactatt	atacttaatg	ccttaacatt	gtgtataaca	aaaggaaata	720
	tctctgagat	acattaagta	actaaaaaa	aaactttaca	cagtctgcct	agtagattac	780
	tatTTGGAAT	atatgtgtc	ttatTTGCT	attcataatc	tccctacttt	atTTTCTTT	840
	atTTTAAATT	gatacataat	cattatacat	atttatgggt	taaagtgtaa	tgttttaata	900
20	tgtgtacaca	tattgacca	atcagggtaa	tttgcat	gtaattttaa	aaaatgcctt	960
	cttcttttaa	tatactttt	tgtttatctt	atttctaata	ctttccctaa	tctctttctt	1020
	tcagggcaat	aatgatacaa	tgtatcatgc	ctcttgcac	cattctaaag	aataacagtg	1080
	ataatttctg	gtttaaggca	atagcaat	ctctgcata	aaatatttct	gcataataat	1140
25	tgttaactgt	gtaagagggtt	tcatattgct	aatagcagct	acaatccagc	taccattctg	1200
	cttttatttt	atgggtggg	taaggctgga	ttattctgag	tccaagctag	gccctttgc	1260
	taatcatgtt	catacctt	atcttcctcc	cacagctcct	gggcaacgtg	ctggctctgtg	1320
	tgctggccca	tcactttggc	aaagaattca	ccccaccagt	gcaggctgcc	tatcagaaag	1380
	tggtggctgg	tgtggctaat	gccctggccc	acaagtatca	ctaa		1424

<210> 22

<211> 444

<212> DNA

<213> HBB gene CDS sequence (DNA sequence)

<400> 22

30	atgggcacc	tgactcctga	ggagaagtct	gccgttactg	ccctgtgggg	caaggtgaac	60
	gtggatgaag	ttgggttgta	ggccctggc	aggctgtgg	tggcttaccc	ttggaccagg	120
	aggttctttg	agtccttgg	ggatctgtcc	actcctgtat	ctgttatggg	caaccctaag	180
	gtgaaggctc	atggcaagaa	agtgcctcggt	gccttagtg	atggcctggc	tcacctggac	240
	aacctcaagg	gcaccttgc	cacactgagt	gagctgact	tgacacaagct	gcacgtggat	300
35	cctgagaact	tcaggctcct	gggcaacgtg	ctggctctgtg	tgctggccca	tcactttggc	360
	aaagaattca	ccccaccagt	gcaggctgcc	tatcagaaag	tggtggctgg	tgtggctaat	420
	gccctggccc	acaagtatca	ctaa				444

<210> 23

<211> 147

<212> PRT

<213> HBB gene amino acid sequence (amino acid sequence)

<400> 23

40	Met Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Trp						
	1	5	10	15			
45	Gly Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu						
	20	25	30				
	Leu Val Val Tyr Pro Trp Thr Gln Arg Phe Phe Glu Ser Phe Gly Asp						
	35	40	45				
50	Leu Ser Thr Pro Asp Ala Val Met Gly Asn Pro Lys Val Lys Ala His						
	50	55	60				
55	Gly Lys Lys Val Leu Gly Ala Phe Ser Asp Gly Leu Ala His Leu Asp						
	65	70	75	80			
	Asn Leu Lys Gly Thr Phe Ala Thr Leu Ser Glu Leu His Cys Asp Lys						
	85	90	95				
60	Leu His Val Asp Pro Glu Asn Phe Arg Leu Leu Gly Asn Val Leu Val						
	100	105	110				
65	Cys Val Leu Ala His His Phe Gly Lys Glu Phe Thr Pro Pro Val Gln						
	115	120	125				

Ala Ala Tyr Gln Lys Val Val Ala Gly Val Ala Asn Ala Leu Ala His  
130 135 140  
Lys Tyr His  
145

5

## Claims

1. A method for generating a new mutation in an organism, which comprises the following steps: sequentially generating two or more DNA breaks at a specific site in a genome of the organism and spontaneously repairing them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair.
2. The method according to claim 1, wherein the "DNA break" is achieved by delivering a nuclease with targeting property into a cell of the organism to contact with the specific site of genomic DNA.
3. The method according to claim 2, wherein the "nuclease with targeting property" is a ZFN, TALEN or CRISPR/Cas system.
4. The method according to any one of claims 1 to 3, wherein the "sequentially generating two or more DNA breaks at a specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by ZFN or TALEN editing, a new ZFN or TALEN protein is designed to cut the site again.
5. The method according to any one of claims 1 to 3, wherein the "sequentially generating two or more DNA breaks at a specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.
6. The method according to any one of claims 1-4, wherein the "two or more DNA breaks" are generated by sequentially delivering different targeted nucleases into recipient cells of different generations, wherein a mutant cell that has completed the previous editing is used as a recipient to receive the delivery of the targeted nuclease for the later editing, thereby performing a second editing to generate a site-specific mutation.
7. The method according to any one of claims 1-3 and 5, wherein the "two or more DNA breaks" are generated by delivering different targeted nucleases for different targets into a same recipient cell.
8. The method according to any one of claims 1-3, 5 and 7, wherein the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.
9. The method according to any one of claims 1-3, 5 and 7, wherein the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA sequentially cut corresponding target sequences.
10. The method according to claim 7, wherein the targeted nuclease is any CRISPR/Cas nuclease capable of performing a genome editing.
11. The method according to claim 6 or 7, wherein the targeted nuclease is in a form of DNA.
12. The method according to any one of claims 7-9, wherein the targeted nuclease is in a form of mRNA or protein instead of DNA.
13. The method according to claim 6 or 7, wherein the method for delivering targeted nuclease into a cell is selected from: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) a microinjection; 5) a gene gun bombardment; or 6) an *Agrobacterium*-mediated transformation method.
14. A new mutation, which is obtained by the method according to any one of claims 1-13.
15. A protein or biologically active fragment thereof, which has the new mutation according to claim 14.

16. A nucleic acid, which comprises a nucleic acid sequence or complementary sequence thereof that encodes the protein or biologically active fragment thereof according to claim 15.

17. A nucleic acid, which comprises:

5 (a) a nucleotide sequence encoding a target RNA, wherein  
the target RNA comprises at least two target RNAs, in which a first target RNA targets a DNA to cause a break  
in the DNA, and a latter target RNA targets a sequence generated from a previous break repair event and  
generates a break again.

10 18. The nucleic acid according to claim 17, further comprising (b) a nucleotide sequence encoding a Cas polypeptide.

19. The nucleic acid according to claim 17 or 18, wherein the target RNA is a sgRNA or gRNA.

15 20. The nucleic acid according to any one of claims 17-19, wherein the Cas polypeptide and the target RNA are present  
in an in vitro cell or an ex vivo cell.

20 21. A recombinant expression vector, comprising the nucleic acid according to any one of claims 16-20, and a promoter  
operably linked thereto.

22. An expression cassette, comprising the nucleic acid according to any one of claims 16-20.

23. A host cell, comprising the expression cassette according to claim 22.

25 24. An organism, which is regenerated from the host cell according to claim 23.

25. A method for lysing a target DNA, comprising contacting the target DNA with a complex, the complex comprising:

30 (a) a Cas polypeptide; and  
(b) at least two target RNAs, in which the first target RNA targets the DNA to cause a break in the DNA, and  
the latter target RNA targets a sequence generated from a previous break repair event and generates a break  
again.

35 26. The method according to claim 25, wherein the target RNA is sgRNA or gRNA.

35 27. The method according to claim 25 or 26, wherein the target DNA is present in a bacterial cell, a eukaryotic cell, a  
plant cell or an animal cell.

40 28. The method according to any one of claims 25-27, wherein the target DNA is a chromosomal DNA.

40 29. The method according to any one of claims 25-28, wherein the Cas polypeptide and the target RNA are present in  
an in vitro cell or an ex vivo cell.

45 30. The method according to any one of claims 25-29, wherein the contacting comprises introducing the following into  
the cell: (a) the Cas polypeptide or a polynucleotide encoding the Cas polypeptide, and (b) the target RNA or a DNA  
polynucleotide encoding the target RNA.

31. A composition, comprising:

50 (a) a Cas polypeptide, or a polynucleotide encoding the Cas polypeptide; and  
(b) at least two target RNAs, or DNA polynucleotides encoding the target RNAs, wherein a first RNA targets a  
DNA to cause a break in the DNA, and a latter target RNA targets a sequence generated from a previous break  
repair event and generates a break again.

55 32. The composition according to claim 31, wherein the target RNA is a sgRNA or gRNA.

33. The composition according to claim 31 or 32, wherein the Cas polypeptide and the target RNA are present in an in  
vitro cell or an ex vivo cell.

34. Use of the composition according to any one of claims 31-33 in manufacture of a medicament for treatment of a disease.

5       35. A kit, comprising:

(a) a Cas polypeptide, or a nucleic acid comprising a nucleotide sequence encoding the Cas polypeptide; and  
 (b) at least two target RNAs, or nucleic acids comprising nucleotide sequences encoding the target RNAs, wherein a first RNA targets a DNA to cause a break in the DNA, and the latter target RNA targets a sequence generated from a previous break repair event and generates a break again;

10      wherein (a) and (b) are in a same or separate containers.

15      36. The kit according to claim 35, wherein the target RNA is a sgRNA or gRNA.

20      37. The kit according to claim 35 or 36, wherein the target RNAs in (b) are in a same or separate containers.

25      38. A method for screening editing events independent of an exogenous transgenic marker, comprising the following steps:

30      1) two or more DNA breaks are sequentially generated in sequence at a specific site of a first target gene of a recipient cell and spontaneously repaired respectively, wherein a later DNA break is generated based on a new sequence generated from a repair of a previous DNA break;

35      2) certain editing events are generated after the specific site of the first target gene is sequentially cut and repaired, which can confer a mutant cell with a resistance to a certain selection pressure to produce a phenotypic selectable trait, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete organism that contains such editing events is isolated;

40      3) optionally, in addition to the first target gene, a targeted nuclease for at least one second target gene is used to edit another target site at the same time, and an editing event of the second target gene is enriched and screened synchronously through the screening of the selectable trait generated by mutations of the first target gene, and a cell, tissue, organ or complete organism that simultaneously contains the editing events of the first target gene and of the at least one second target gene is isolated.

45      39. The method according to claim 38, wherein the "first target gene" is a gene locus encoding at least one phenotypic selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

50      40. The method according to claim 38 or 39, wherein the "specific site of a first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs, which is capable of conferring the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.

55      41. The method according to claim 40, wherein the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.

42. The method according to any one of claims 38-41, wherein the "certain selection pressure" is an environmental pressure or a pressure resulted from an added compound, wherein the environmental pressure is preferably high temperature, low temperature or hypoxia; and the pressure resulted from an added compound is preferably a pressure resulted from a salt ion concentration, antibiotic, cytotoxin or herbicide.

43. The method according to any one of claims 38-42, wherein the "DNA break" is achieved by delivering a nuclease with a targeting property into a cell of the organism to contact with the specific site of genomic DNA.

44. The method according to claim 43, wherein the "nuclease with a targeting property" is any CRISPR/Cas nuclease capable of performing genome editing.

45. The method according to any one of claims 38-44, wherein feature, the "two or more DNA breaks are sequentially generated in sequence at a specific site", refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.

46. The method according to any one of claims 38-45, wherein the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.

5       47. The method according to any one of claims 38-45, wherein the "two or more DNA breaks" are generated when RNP complexes respectively formed by two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences.

10      48. The method according to any one of claims 38-47, wherein the "second target gene" refers to another gene that is different in coding from the first target gene.

15      49. The method according to any one of claims 38-48, wherein the "targeted nuclease for at least one second target gene" and the CRISPR/Cas nuclease used for generating DNA break at a specific site of the first target gene are the same or different.

50. The method according to any one of claims 38-45 and 48-49, wherein the targeted nuclease is in a form of DNA.

20      51. The method according to any one of claims 38-49, wherein the targeted nuclease is in a form of mRNA or protein instead of DNA.

25      52. The method according to any one of claims 43-51, wherein the method for delivering targeted nuclease into a cell is selected from: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) microinjection; 5) a gene gun bombardment; or 6) an *Agrobacterium*-mediated transformation method.

53. A method for non-transgenic transient editing of an organism genome, comprising the following steps:

30      1) a combination of at least two crRNA fragments or a combination of at least two sgRNA fragments is designed and synthesized for a specific site of a first target gene of a recipient cell, wherein the crRNA combination combined with tracrRNA or the sgRNA combination alone is capable of guiding a corresponding Cas protein to sequentially generate two or more DNA breaks at the specific site in the first target gene of the recipient cell and to spontaneously repair them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair;

35      2) an appropriate amount of CRISPR/Cas protein or corresponding mRNA thereof is mixed with the combination of crRNA fragments and tracrRNA fragment or with the combination of sgRNA fragments alone as above designed and synthesized in advance that is capable of guiding a site-specific editing of the first target gene to generate endogenous selection markers, optionally, at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes is further added, and incubation is carried out in vitro to form an RNP complex;

40      3) the above RNP complex is delivered into the recipient cell and contacts with the specific site of genomic DNA to achieve gene editing;

45      4) according to a phenotypic selectable trait generated by the site-specific editing of the first target gene by the RNP complex, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete organism that contains the editing event is isolated, and optionally, a cell, tissue, organ or complete organism that simultaneously contains the editing events of the first target gene and of the at least one of a second, third or more target genes is isolated.

50      54. The method according to claim 53, wherein the "first target gene" is a gene locus encoding at least one phenotypic selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

55      55. The method according to claim 53 or 54, wherein the "specific site of the first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs, which is capable of conferring the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.

56. The method of claim 55, wherein the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.

57. The method according to claim 55 or 56, wherein the "certain selection pressure" is an environmental pressure or a pressure resulted from an added compound; wherein the environmental pressure is preferably high temperature, low temperature or hypoxia; the pressure resulted from an added compound is preferably a pressure resulted from a salt ion concentration, antibiotic, cytotoxin or herbicide.

5 58. The method according to any one of claims 53-57, wherein the CRISPR/Cas protein is any CRISPR/Cas nuclease capable of performing genome editing.

10 59. The method according to any one of claims 53-58, wherein the feature "to sequentially generate two or more DNA breaks at a specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.

15 60. The method according to any one of claims 53-59, wherein the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.

20 61. The method according to any one of claims 53-59, wherein the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences.

25 62. The method according to any one of claims 53-61, wherein the "second, third or more target genes" refer to other genes that are different in coding from the first target gene.

63. The method according to any one of claims 53-62, wherein the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" shares the same Cas protein with the crRNA or sgRNA targeting the first target gene.

30 64. The method according to any one of claims 53-62, wherein the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" and the crRNA or sgRNA targeting the first target gene use Cas proteins that recognize different PAM sequences.

35 65. The method according to any one of claims 53-64, wherein the method for delivering the RNP complex into a cell is selected from: 1) a PEG-mediated cell transfection method; 2) a liposome-mediated cell transfection method; 3) an electroporation transformation method; 4) a microinjection; or 5) a gene gun bombardment.

66. A method for non-transgenic transient editing of a plant genome, comprising the following steps:

40 1) a combination of at least two crRNA fragments or a combination of at least two sgRNA fragments is designed and synthesized for a specific site of a first target gene of a recipient plant cell or tissue, wherein the crRNA combination combined with tracrRNA or the sgRNA combination alone is capable of guiding a corresponding Cas protein to sequentially generate two or more DNA breaks at the specific site in the first target gene of the recipient cell and to spontaneously repair them respectively, wherein a later DNA break is generated based on a new sequence generated from a previous DNA break repair;

45 2) an appropriate amount of CRISPR/Cas protein or corresponding mRNA thereof is mixed with the combination of crRNA fragments and tracrRNA fragment or with the combination of sgRNA fragments alone as above designed and synthesized in advance that is capable of guiding a site-specific editing of the first target gene to generate endogenous selection markers, optionally, at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes is further added, and incubation is carried out in vitro to form an RNP complex;

50 3) the above RNP complex is delivered into the recipient plant cell or tissue and contacts with the specific site of genomic DNA to achieve gene editing;

55 4) according to a phenotypic selectable trait generated by a site-specific editing of the first target gene by the RNP complex, a corresponding selection pressure is applied to make a selection for the trait, and a cell, tissue, organ or complete plant that contains the editing event is isolated, and optionally, a cell, tissue, organ or complete plant that simultaneously contains the editing events of the first target gene and of the at least one of second, third or more target genes is isolated.

67. The method according to claim 66, wherein the "first target gene" is a gene locus encoding at least one phenotypic

selectable trait, wherein the at least one phenotypic selectable trait is a resistance/tolerance trait or a growth advantage trait.

- 5       **68.** The method according to claim 66 or 67, wherein the "specific site of the first target gene" refers to a site at which a certain type of mutation is generated after sequential cuttings and repairs, which is capable of conferring the recipient cell with a resistance to a certain selection pressure to produce at least one phenotypic selectable resistance/tolerance trait or growth advantage trait.
- 10      **69.** The method according to claim 68, wherein the "certain type of mutation" comprises substitution of single base, substitution of a plurality of bases, or insertion or deletion of an unspecified number of bases.
- 15      **70.** The method according to claim 68 or 69, wherein the "certain selection pressure" is an environmental pressure or a pressure resulted from an added compound; wherein the environmental pressure is preferably high temperature, low temperature or hypoxia; and the pressure resulted from an added compound is preferably a pressure resulted from a salt ion concentration, antibiotic, cytotoxin or herbicide.
- 20      **71.** The method according to any one of claims 66-70, wherein the "recipient plant cell or tissue" is any cell or tissue that can serve as a recipient for transient expression and can be regenerated into a complete plant through tissue culture; wherein the cell is preferably a protoplast cell or a suspension cell; and the tissue is preferably a callus, immature embryo, mature embryo, leaf, shoot tip, young spike or hypocotyl.
- 25      **72.** The method according to any one of claims 66-71, wherein the CRISPR/Cas protein is any CRISPR/Cas nuclease capable of performing genome editing.
- 30      **73.** The method according to any one of claims 66-72, wherein the feature "to sequentially generate two or more DNA breaks at the specific site" refers to that based on a new sequence formed by a previous DNA break repair event generated by a CRISPR/Cas system, a new target RNA is designed to cut the site again.
- 35      **74.** The method according to any one of claims 66-73, wherein the "two or more DNA breaks" are generated when RNP complexes formed by a same CRISPR/Cas nuclease respectively with different gRNAs or sgRNAs sequentially cut corresponding target sequences.
- 40      **75.** The method according to any one of claims 66-73, wherein the "two or more DNA breaks" are generated when RNP complexes respectively formed by each of two or more CRISPR/Cas nucleases that recognize different PAM sequences with respective gRNA or sgRNA, sequentially cut corresponding target sequences.
- 45      **76.** The method according to any one of claims 66-75, wherein the "second, third or more target genes" refer to other genes that are different in coding from the first target gene.
- 50      **77.** The method according to any one of claims 66-76, wherein the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" shares the same Cas protein with the crRNA or sgRNA targeting the first target gene.
- 55      **78.** The method according to any one of claims 66-76, wherein the "at least one of artificially synthesized crRNA and tracrRNA fragments or artificially synthesized sgRNA fragments targeting a second, third or more target genes" and the crRNA or sgRNA targeting the first target gene use Cas proteins that recognize different PAM sequences.
- 60      **79.** The method according to any one of claims 66-78, wherein the method for delivering the RNP complex into the plant cell is selected from: 1) a PEG-mediated protoplast transformation method; 2) a microinjection; 3) a gene gun bombardment; 4) a silicon carbide fiber-mediated method; or 5) a vacuum infiltration method, or any other transient introduction method.
- 65      **80.** The method according to any one of claims 66-79, wherein the "first target gene" is at least one endogenous gene that encodes at least one phenotypic selectable trait selected from herbicide resistance/tolerance, wherein the herbicide resistance/tolerance is selected from the group consisting of resistance/tolerance to EPSPS inhibitor, resistance/tolerance to glutamine synthesis inhibitor, resistance/tolerance to ALS or AHAS inhibitor, resistance/tolerance to ACCase inhibitor, resistance/tolerance to carotenoid biosynthesis inhibitor, resistance/tolerance to cellulose inhibitor, resistance/tolerance to lipid synthesis inhibitor, resistance/tolerance to long-chain fatty acid inhibitor,

resistance/tolerance to microtubule assembly inhibitor, resistance/tolerance to photosystem I electron shunting agent, resistance/tolerance to photosystem II inhibitor or resistance/tolerance to PPO inhibitor, and resistance/tolerance to synthetic growth hormone; preferably, the "first target gene" is selected from PsbA, ALS, EPSPS, ACCase, PPO, HPPD, PDS, GS, DOXPS, TIR1 or AFB5.

- 5           **81.** The method according to claim 80, wherein the "first target gene" is ALS, and the "specific site of gene" refers to site A122, P197, R198, D204, A205, D376, R377, W574, S653 or G654 in an *Arabidopsis* AtALS protein amino acid sequence, and amino acid sites in an ALS protein of another plant which correspond to the above-mentioned amino acid sites by using the AtALS amino acid sequence as reference standard; or  
10          the crRNA or sgRNA targets a target sequence comprising a sequence encoding an AtALS protein amino acid sequence site selected from the group consisting of A122, P197, R198, D204, A205, D376, R377, W574, S653, G654 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an ALS protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the AtALS amino acid sequence as reference standard.  
15          **82.** The method according to claim 80, wherein the "first target gene" is ACCase, and the "specific site of gene" refers to site 11781, E1874, N1878, W1999, W2027, 12041, D2078, C2088 or G2096 in an *Alopecurus myosuroides* AmACCase protein amino acid sequence, and amino acid sites in an ACCase protein of another monocotyledonous plant which correspond to the above-mentioned amino acid sites by using the AmACCase amino acid sequence as  
20          reference standard; or  
25          the crRNA or sgRNA targets a target sequence comprising a sequence encoding an AmACCase amino acid sequence site selected from the group consisting of 11781, E1874, N1878, W1999, W2027, 12041, D2078, C2088, G2096 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an ACCase protein of another monocotyledonous plant which corresponds to the above-mentioned amino acid sites, and any combination thereof, by using the AmACCase amino acid sequence as reference standard.  
30          **83.** The method according to claim 80, wherein the "first target gene" is HPPD, and the "specific site of gene" refers to site H141, L276, P277, N338, G342, R346, D370, P386, K418 or G419 in an *Oryza sativa* OsHPPD protein amino acid sequence, and amino acid sites in an HPPD protein of another plant which correspond to the above-mentioned amino acid sites by using the OsHPPD amino acid sequence as reference standard; or  
35          the crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsHPPD amino acid sequence site selected from the group consisting of H141, L276, P277, N338, G342, R346, D370, P386, K418, G419 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in an HPPD protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the OsHPPD amino acid sequence as reference standard.  
40          **84.** The method according to claim 80, wherein the "first target gene" is PPO, and the "specific site of gene" refers to site S128, V217, S223, V364, K373, L423, Y425 or W470 in an *Oryza sativa* OsPPO1 protein amino acid sequence, and amino acid sites in a PPO protein of another plant which correspond to the above-mentioned amino acid sites by using the amino acid sequence of OsPPO1 as reference standard; or  
45          the crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsPPO1 amino acid sequence site selected from the group consisting of S128, V217, S223, V364, K373, L423, Y425, W470 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in a PPO protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the OsPPO1 amino acid sequence as reference standard.  
50          **85.** The method according to claim 80, wherein the "first target gene" is TIR1, and the "specific site of gene" refers to site F93, F357, C413 or S448 in an *Oryza sativa* OsTIR1 protein amino acid sequence, and amino acid sites in a TIR1 protein of another plant which correspond to the above-mentioned amino acid sites by using the OsTIR1 amino acid sequence as reference standard; or  
55          the crRNA or sgRNA targets a target sequence comprising a sequence encoding an OsTIR1 amino acid sequence site selected from the group consisting of F93, F357, C413, S448 or any combination thereof, and a target sequence comprising a sequence encoding an amino acid site in a TIR1 protein of another plant which corresponds to the above-mentioned amino acid site, and any combination thereof, by using the OsTIR1 amino acid sequence as reference standard.  
**86.** A non-transgenic transient editing system, which uses the method according to any one of claims 53-85.

87. Use of the non-transgenic transient editing system according to claim 86 as a selection marker or in disease treatment or biological breeding.

5       88. A genetically modified plant obtained by the method according to any one of claims 66-85, wherein the genome thereof comprises:

- 10       1) an editing event of a first target gene;
- 2) an editing event of the first target gene and an editing event of at least one second target gene; or
- 3) at least one second target gene editing event, wherein the editing event of the first target gene has been removed by genetic separation;

wherein, the genetically modified plant is obtained in a non-transgenic manner.

15       89. A new plant gene mutation obtained by the method according to any one of claims 66-85.

90. A new mutation generated in a plant, which comprises one or a combination of two or more of the following types: substitution of aspartic acid at a site corresponding to *Arabidopsis* ALS376 with any other amino acid, substitution of tryptophan at a site corresponding to *Arabidopsis* ALS574 with any other amino acid, substitution of serine at a site corresponding to *Arabidopsis* ALS653 with any other amino acid, or substitution of serine at a site corresponding to *Arabidopsis* ALS654 with any other amino acid; or substitution of tryptophan at a site corresponding to *Alopecurus myosuroides* ACCase2027 with any other amino acid.

25       91. The mutation according to claim 90, wherein the aspartic acid at a site corresponding to *Arabidopsis* ALS376 is substituted by glutamic acid, the tryptophan at a site corresponding to *Arabidopsis* ALS574 is substituted by leucine or methionine, the serine at a site corresponding to *Arabidopsis* ALS653 is substituted by asparagine or arginine, or the glycine at a site corresponding to *Arabidopsis* ALS654 is substituted by aspartic acid; or, the tryptophan at a site corresponding to *Alopecurus myosuroides* ACCase2027 is substituted by leucine or cysteine.

30       92. The mutation according to claim 90, wherein the aspartic acid at site 350 of *Oryza sativa* ALS is substituted by any other amino acid, the tryptophan at site 548 of *Oryza sativa* ALS is substituted by any other amino acid, or the tryptophan at site 561 of *Solanum tuberosum L.* ALS2 is substituted by any other amino acid; or, the tryptophan at site 2038 of *Oryza sativa* ACCase2 is substituted by any other amino acid.

35       93. The mutation according to any one of claims 90-92, wherein the aspartic acid at site 350 of *Oryza sativa* ALS is substituted by glutamic acid, the tryptophan at site 548 of *Oryza sativa* ALS is substituted by leucine or methionine, or the tryptophan at site 561 of *Solanum tuberosum L.* ALS2 is substituted by leucine or methionine; or, the tryptophan at site 2038 of *Oryza sativa* ACCase2 is substituted by leucine or cysteine.

40       94. A protein or biologically active fragment thereof, comprising the new mutation according to any one of claims 89-93.

95. A nucleic acid, comprising a nucleic acid sequence or complementary sequence thereof that encodes the protein or biologically active fragment thereof according to claim 94.

45       96. A recombinant expression vector, comprising the nucleic acid according to claim 95 and a promoter operably linked thereto.

97. An expression cassette, comprising the nucleic acid according to claim 95.

98. A plant cell, comprising the expression cassette according to claim 97.

50       99. A plant regenerated by using the plant cell according to claim 98.

100. A method for producing a plant with improved resistance or tolerance to herbicides, which comprises transforming or transfecting a plant cell with the recombinant expression vector according to claim 96 or the expression cassette according to claim 97, and regenerating the transformed or transfected plant cell into a plant.

55       101. A method for controlling weeds in a plant cultivation site, wherein the plant includes the plant according to claim 88 or 99 or the plant produced by the method according to claim 100, the method comprises applying to the cultivation

site one or more herbicides in an effective amount to control the weeds.

**102.** Use of the new mutation according to any one of claims 89-93, the protein or biologically active fragment thereof according to claim 94, the nucleic acid according to claim 95, the recombinant expression vector according to claim 96, or the expression cassette according to claim 97 in improving resistance or tolerance of a plant cell, a plant tissue, a plant part or a plant to herbicides.

10

15

20

25

30

35

40

45

50

55

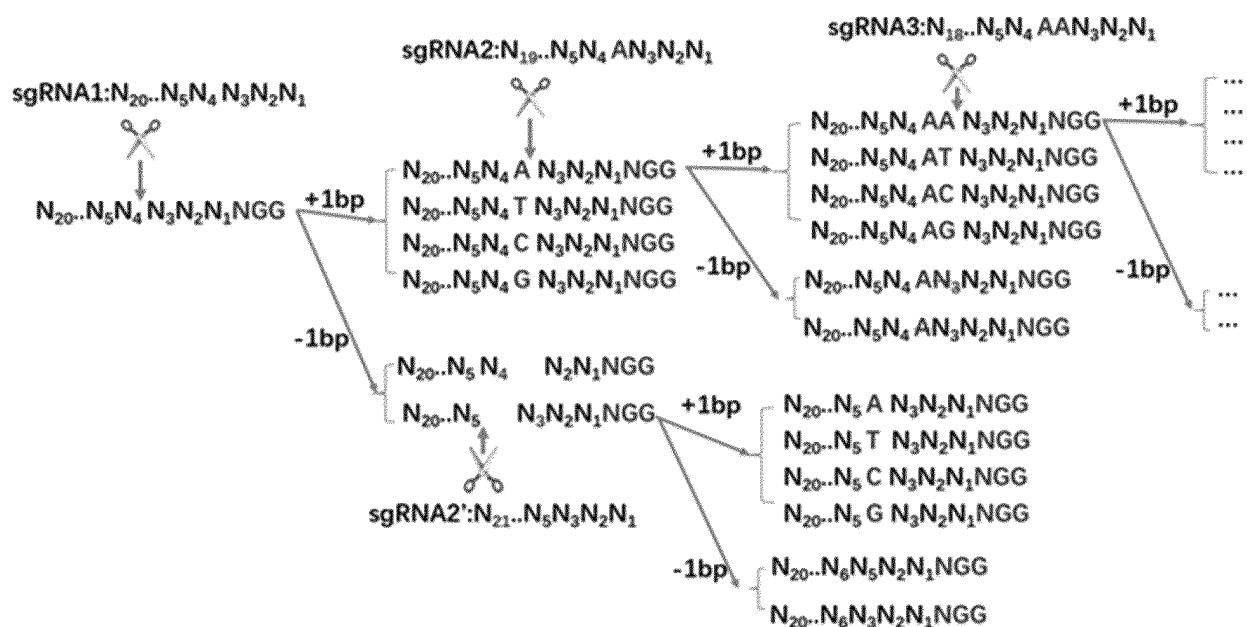


Figure 1

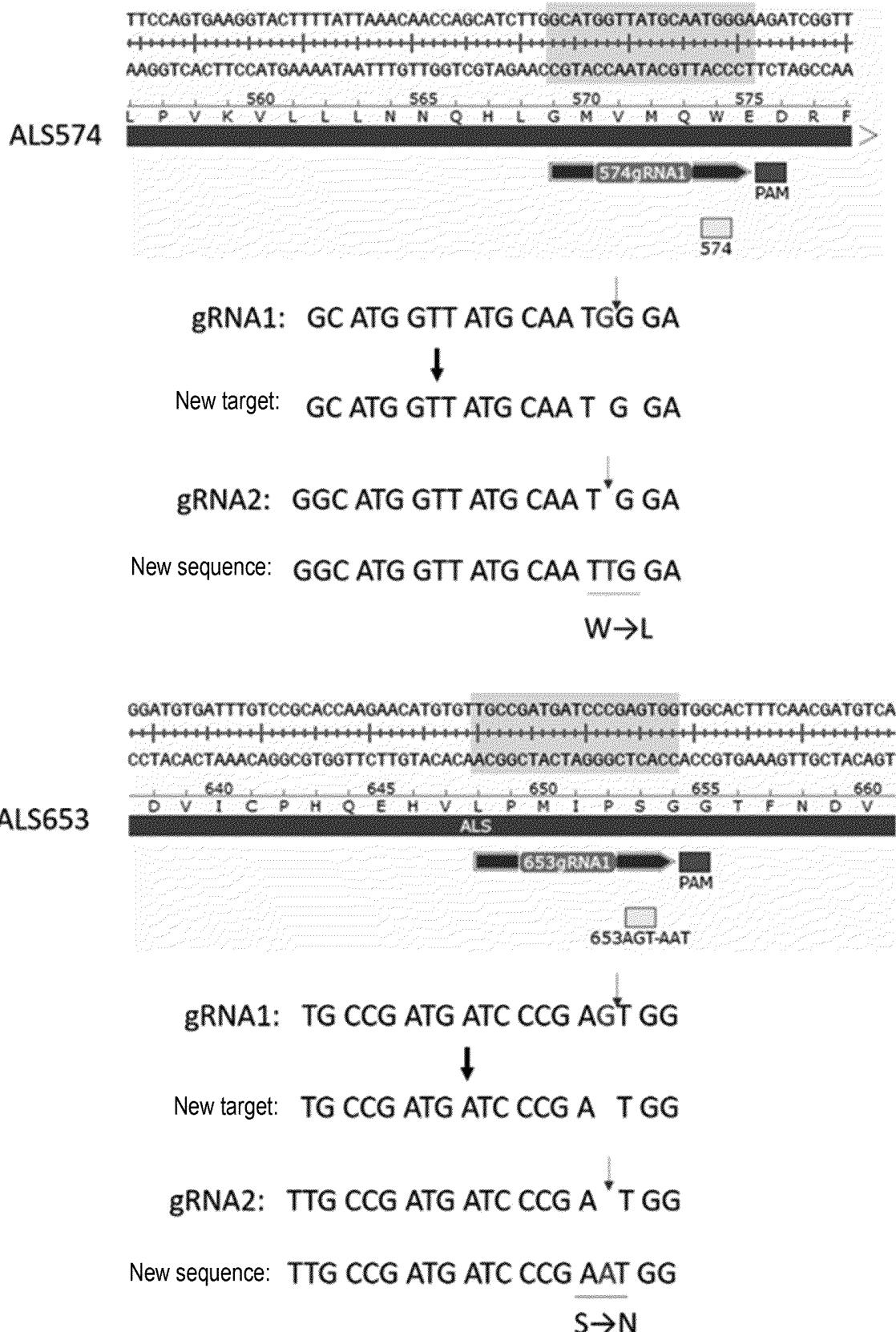


Figure 2

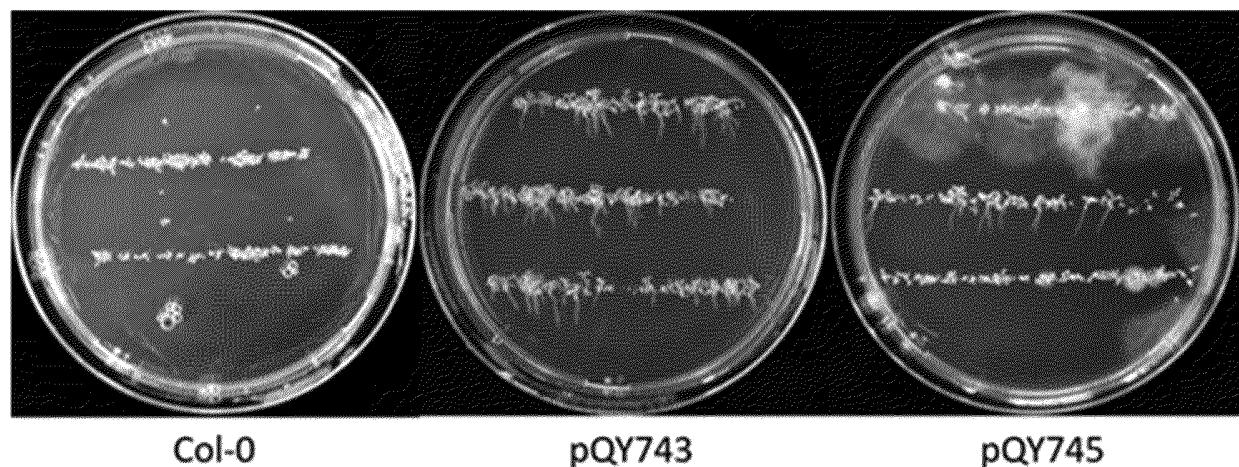


Figure 3

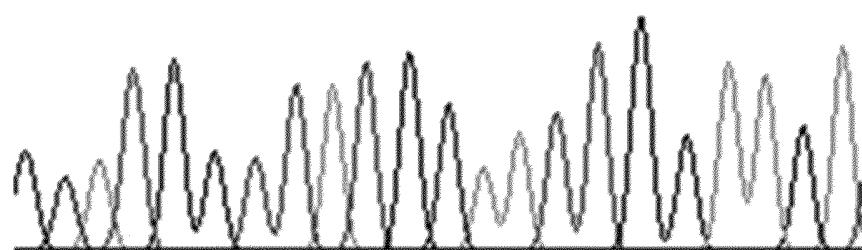


Figure 4

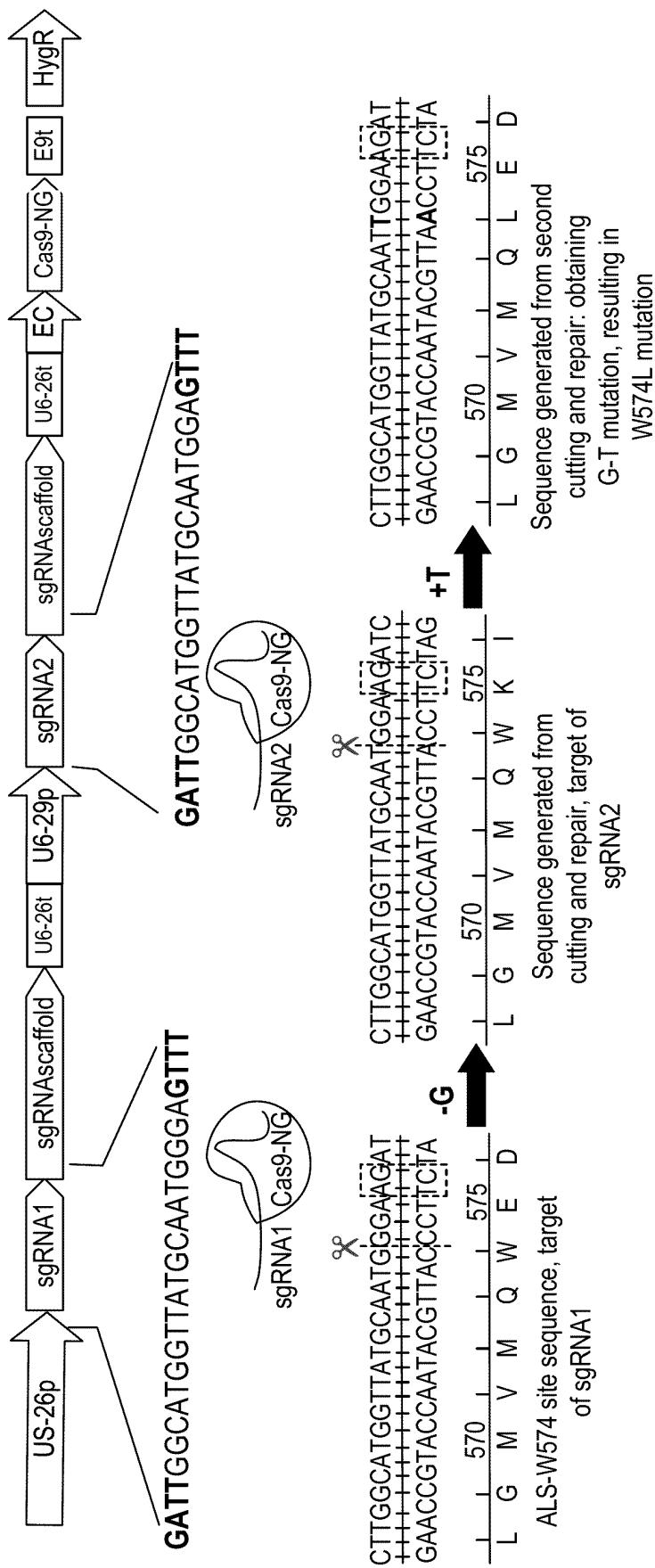


Figure 5

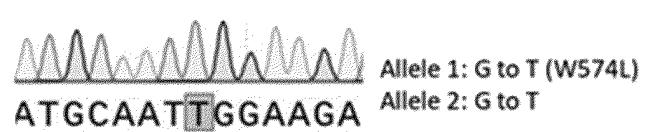
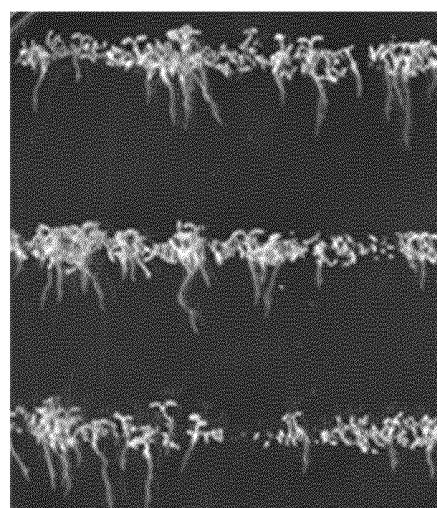


Figure 6

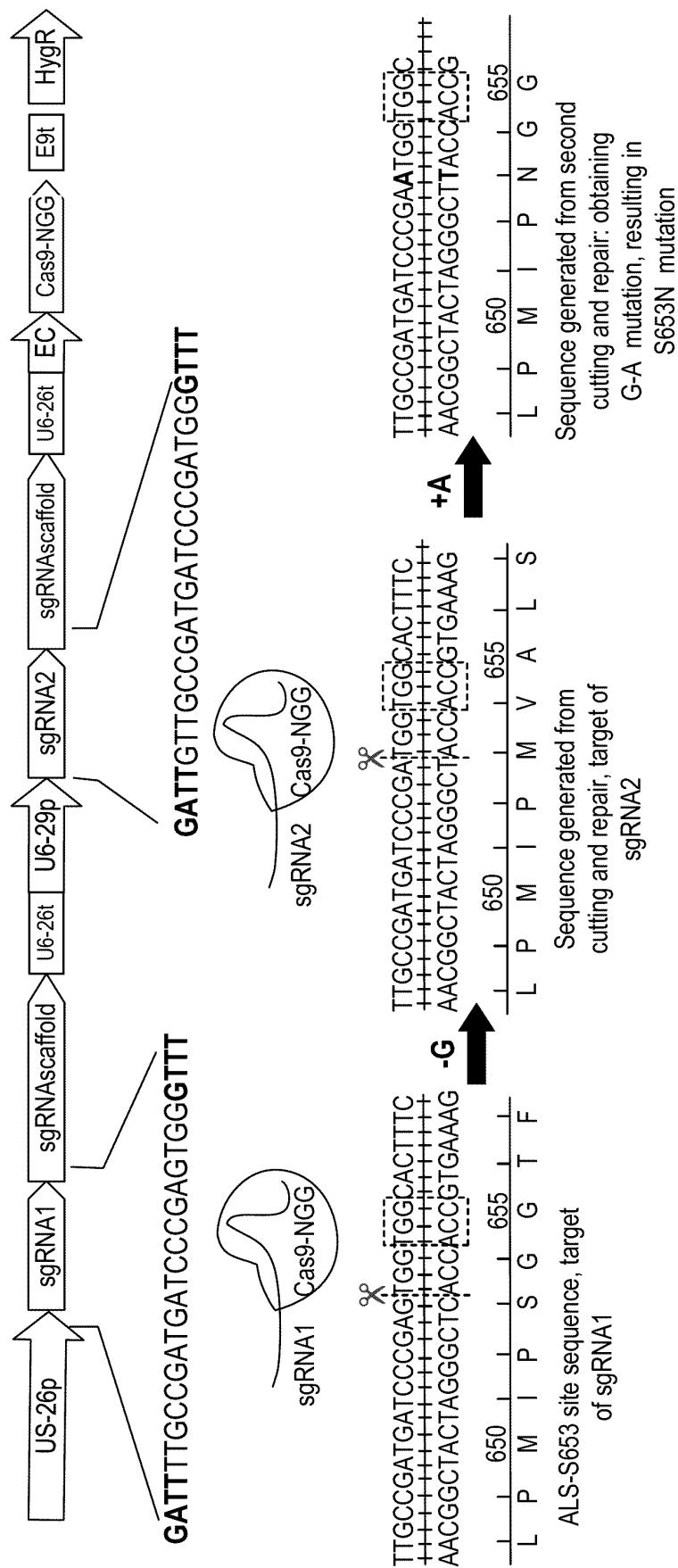


Figure 7

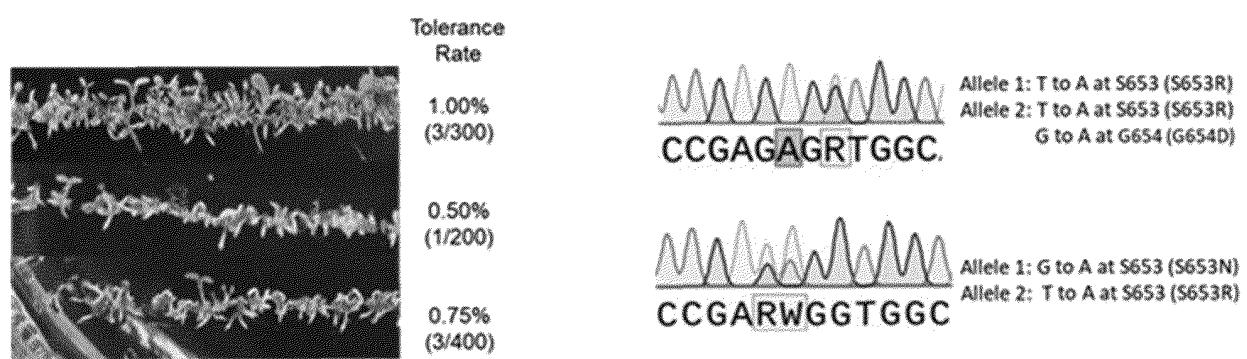


Figure 8

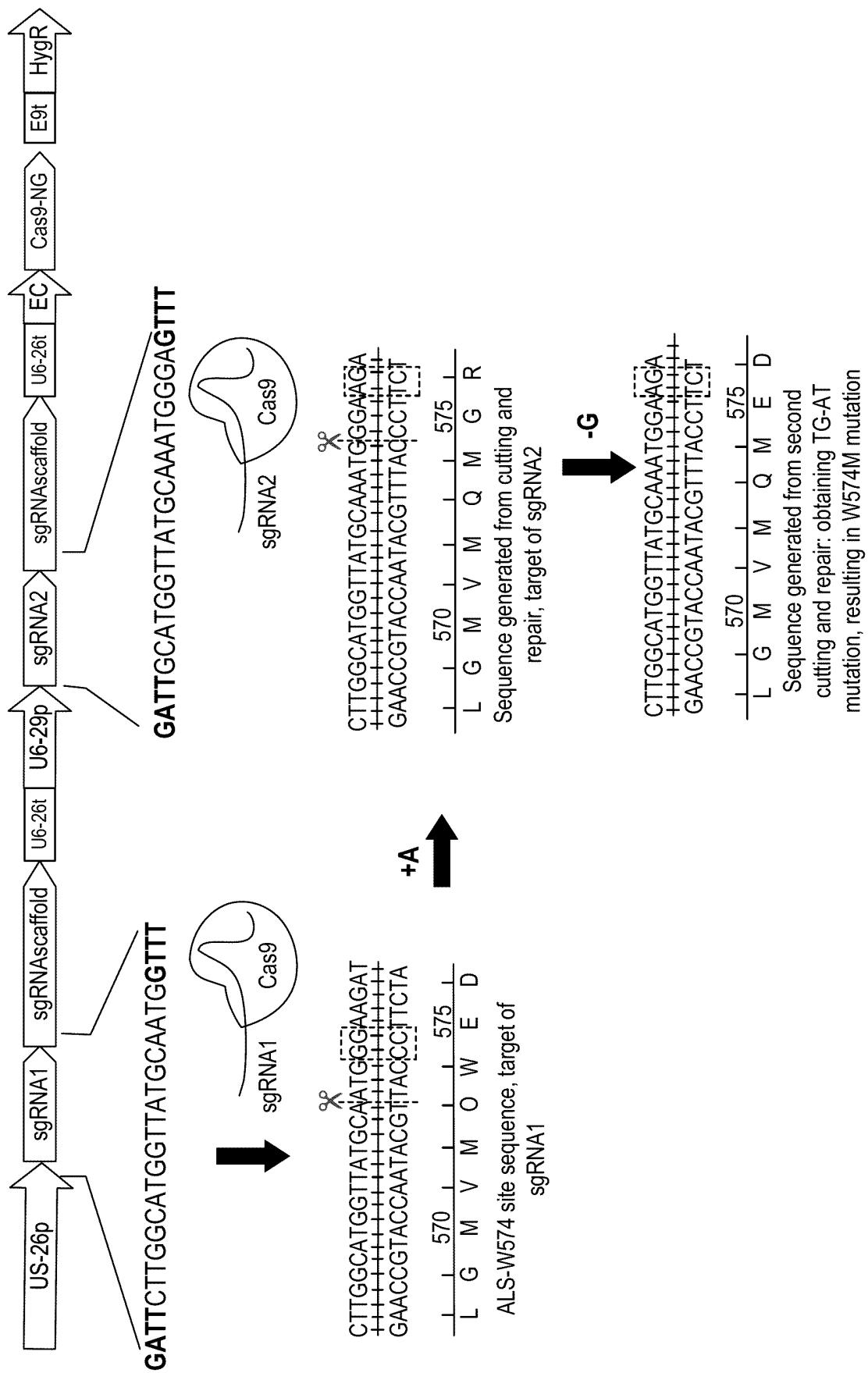


Figure 9

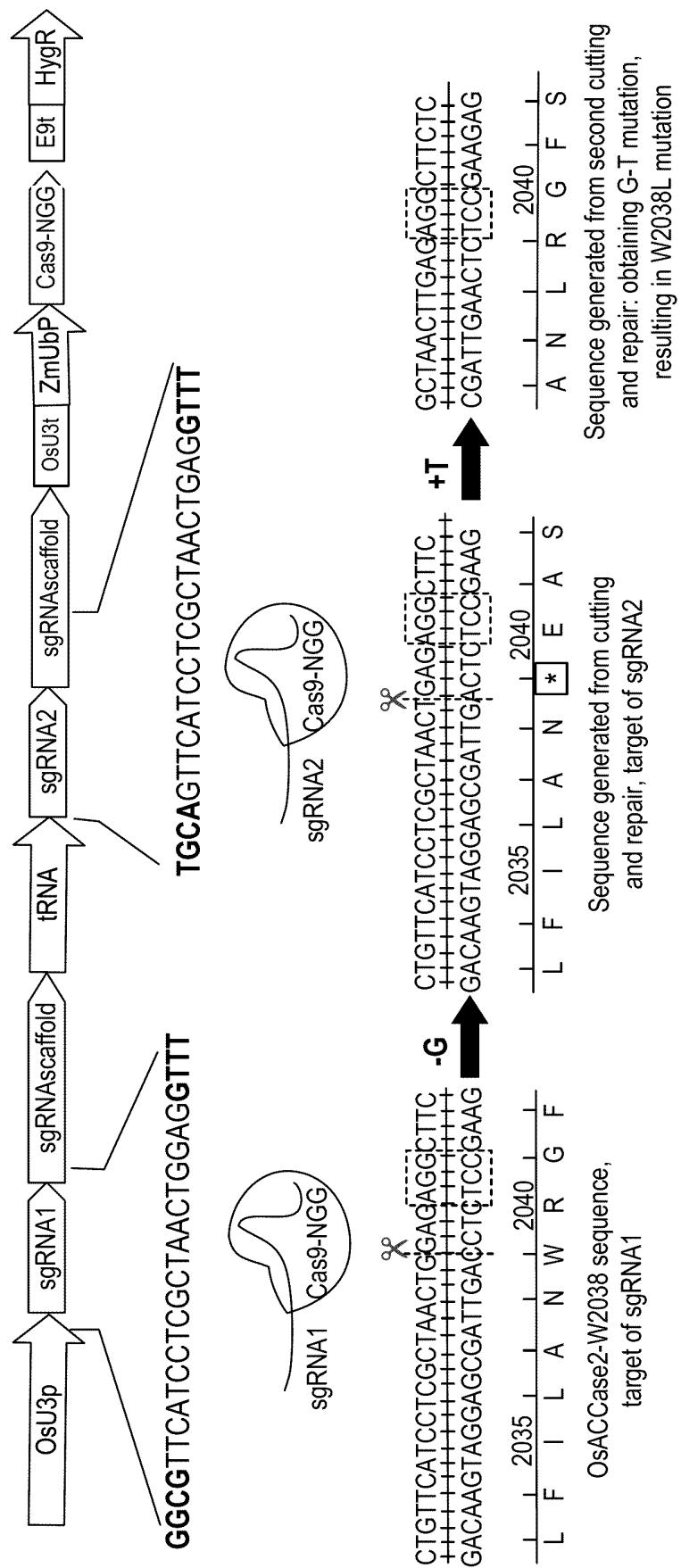


Figure 10

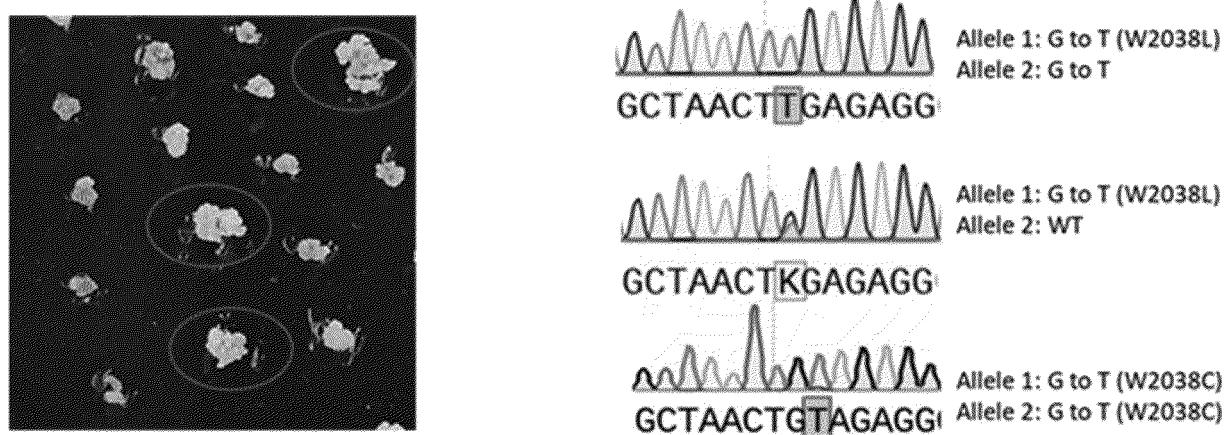


Figure 11

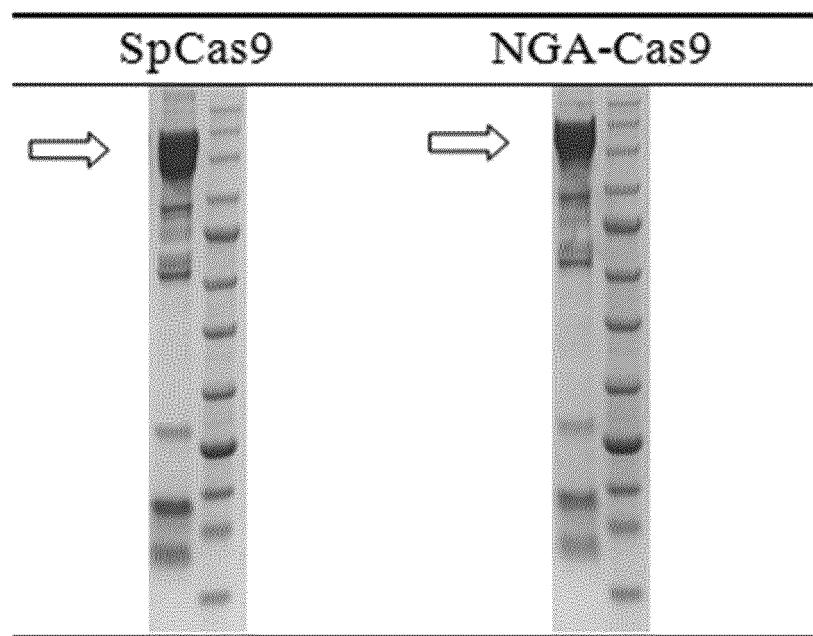


Figure 12

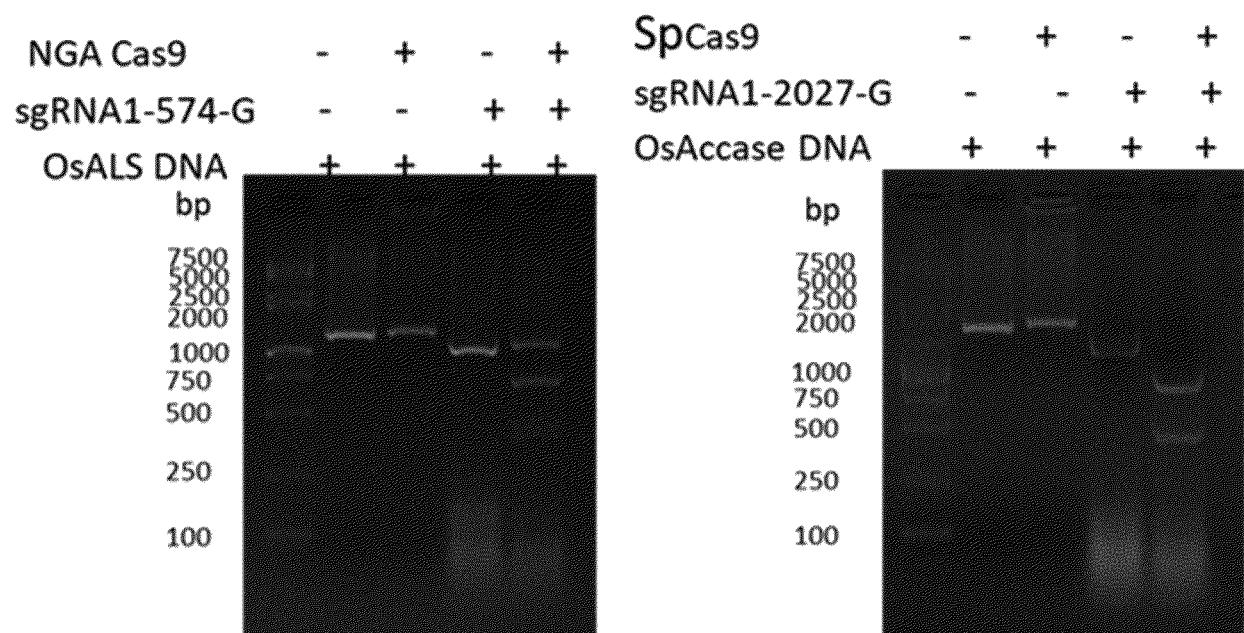


Figure 13

5'TGCAATGGGAGGGATAGGTTT3'

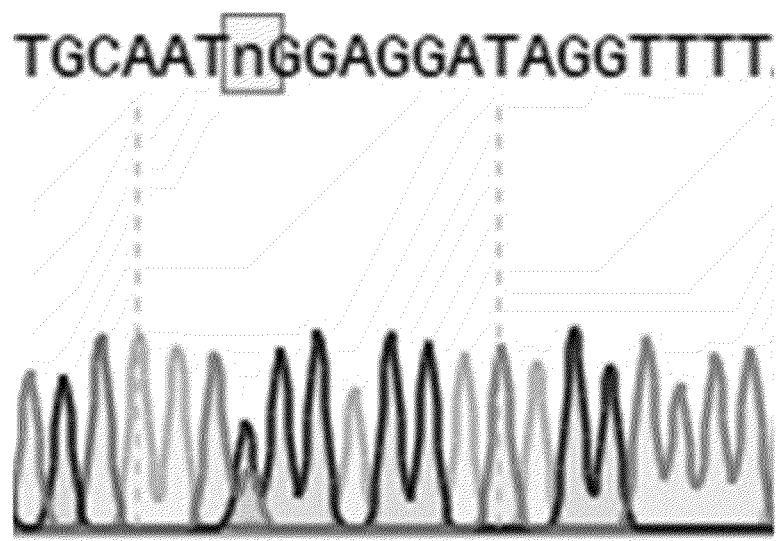


Figure 14

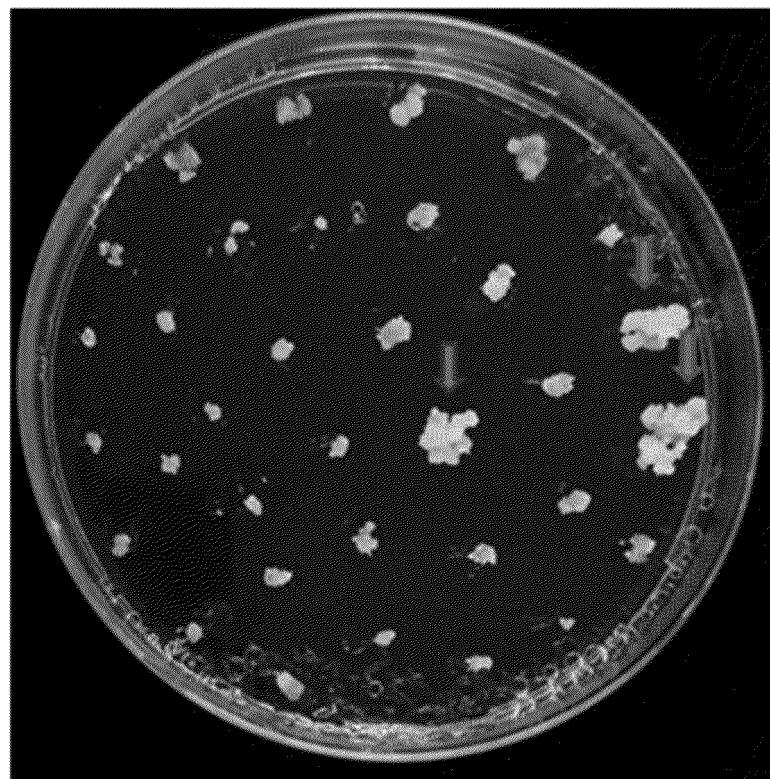


Figure 15

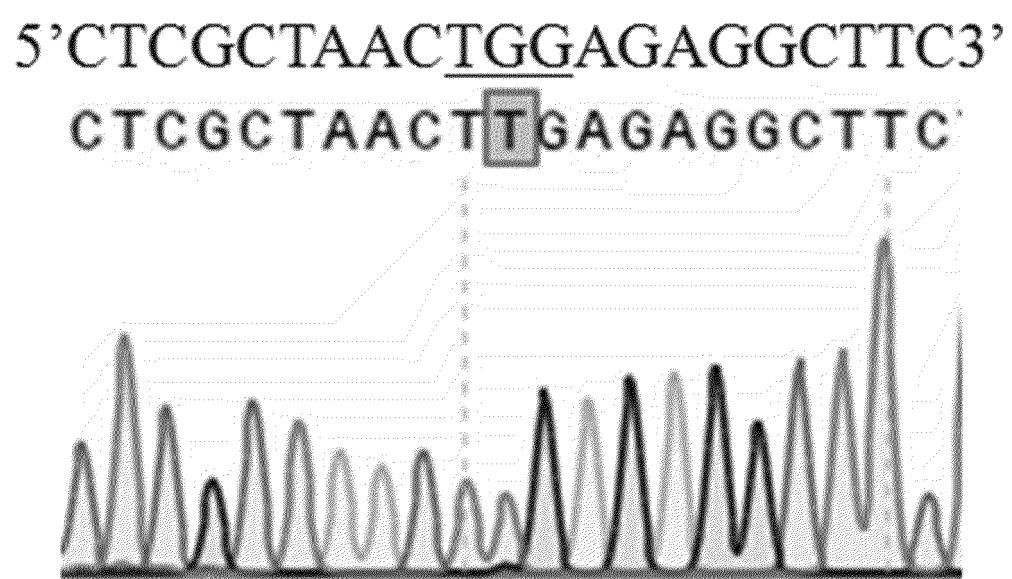


Figure 16

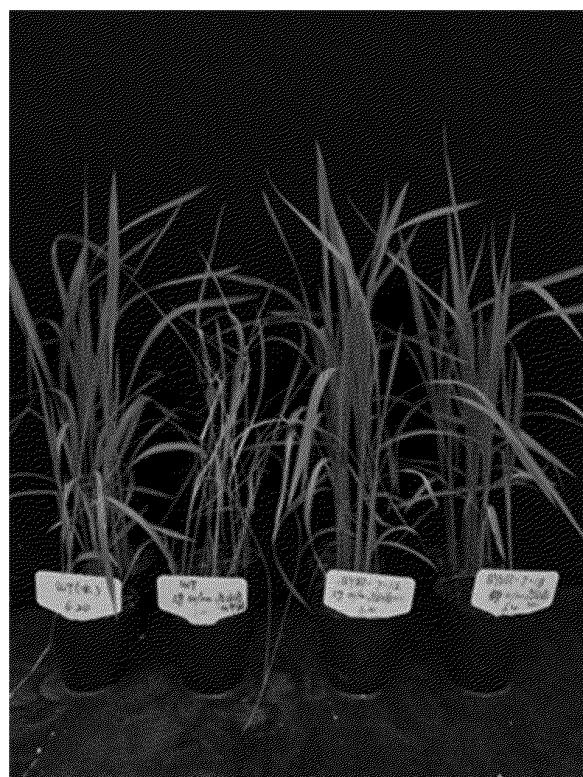


Figure 17

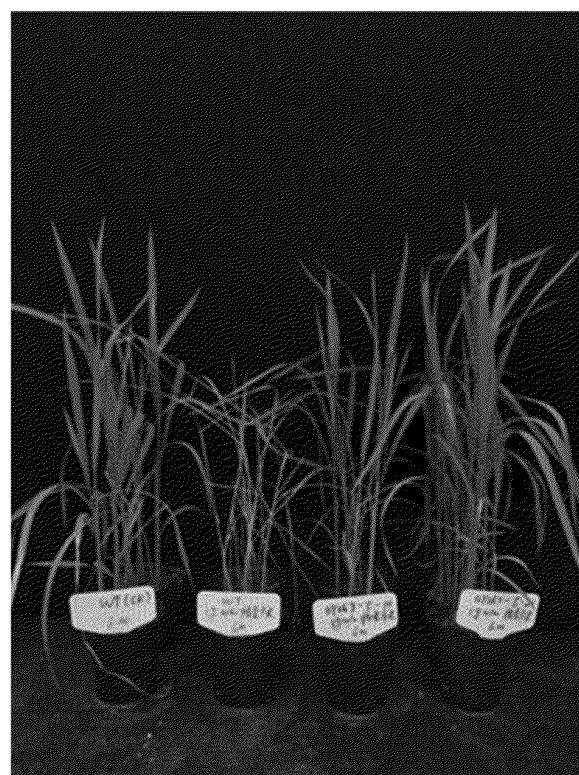


Figure 18

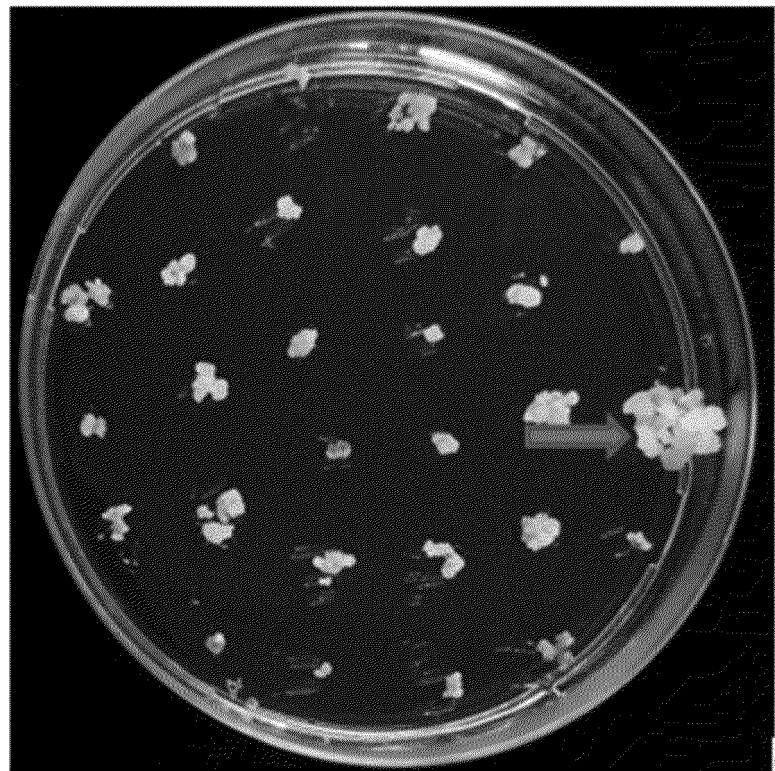


Figure 19

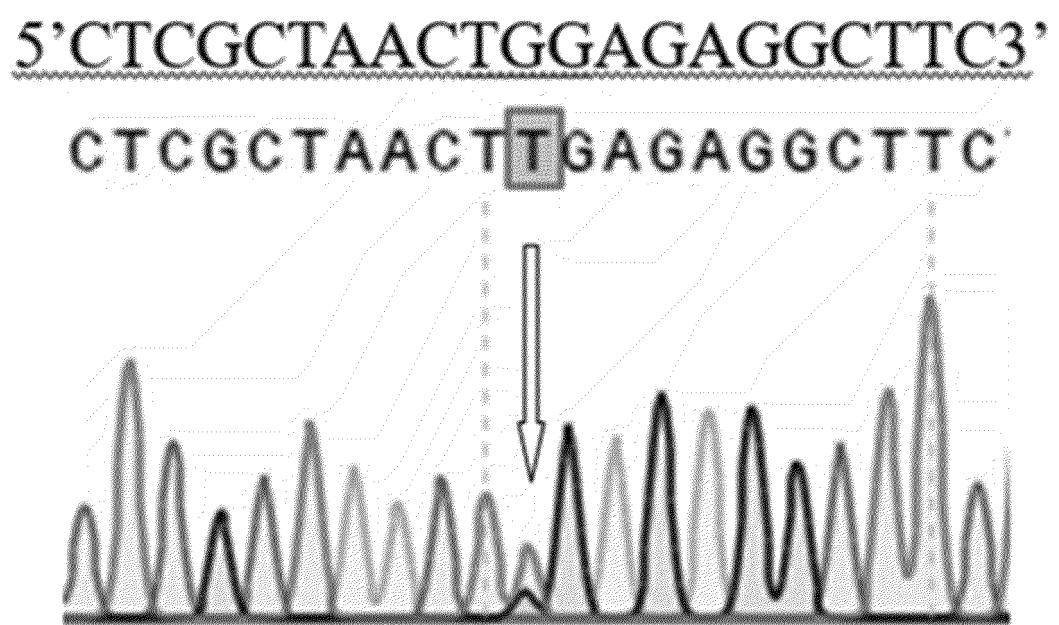


Figure 20

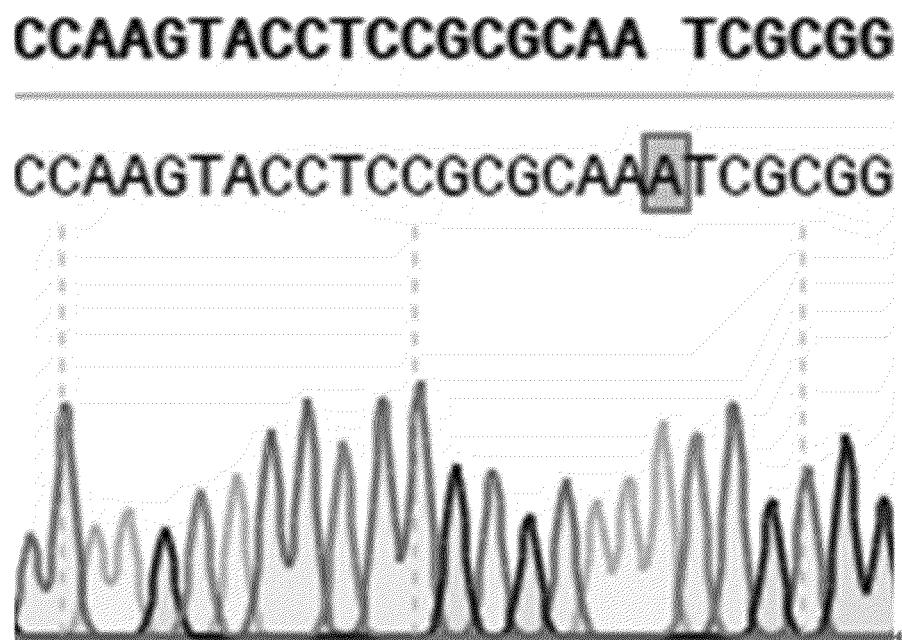


Figure 21

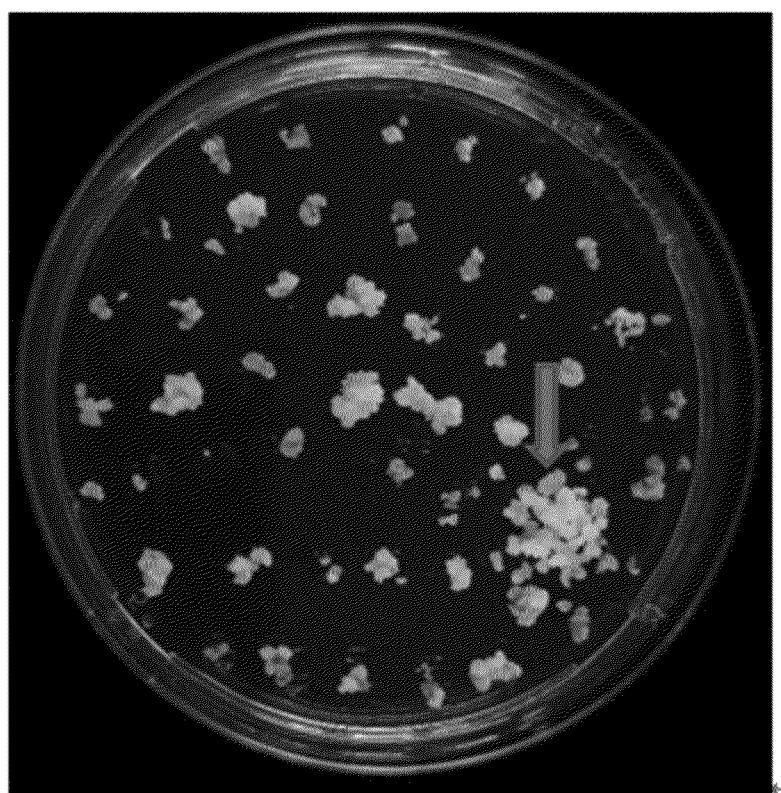


Figure 22

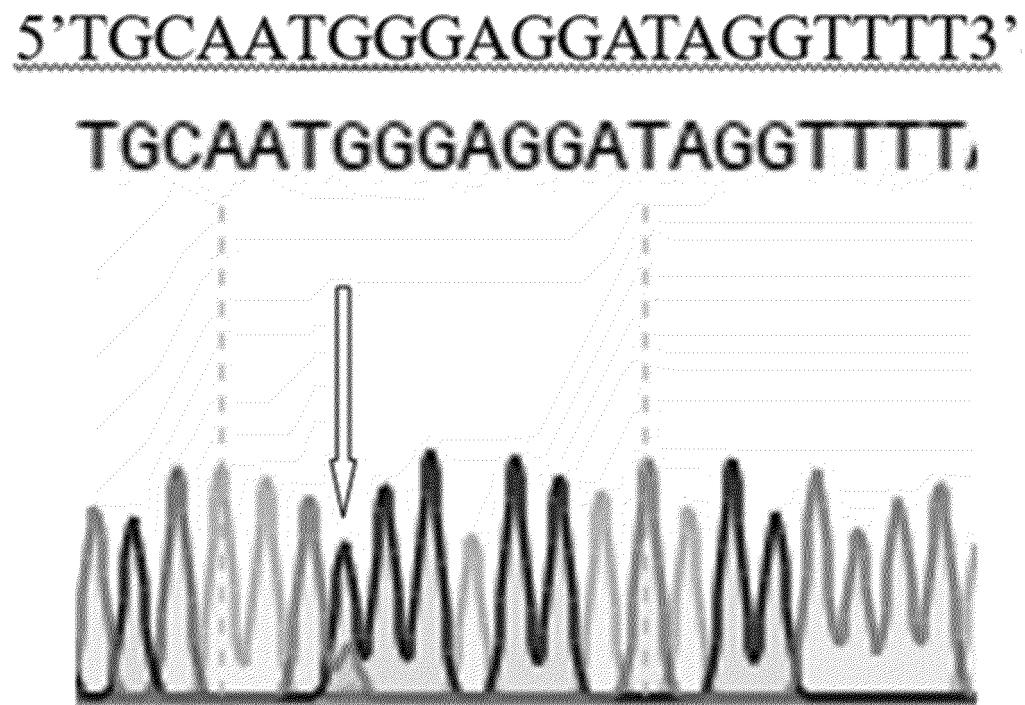


Figure 23

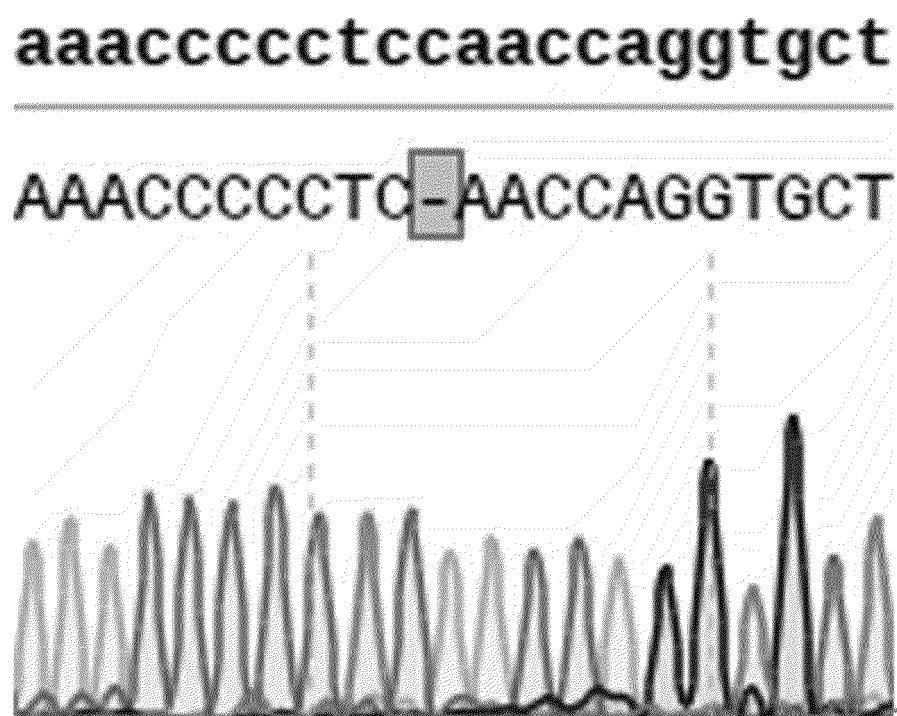


Figure 24

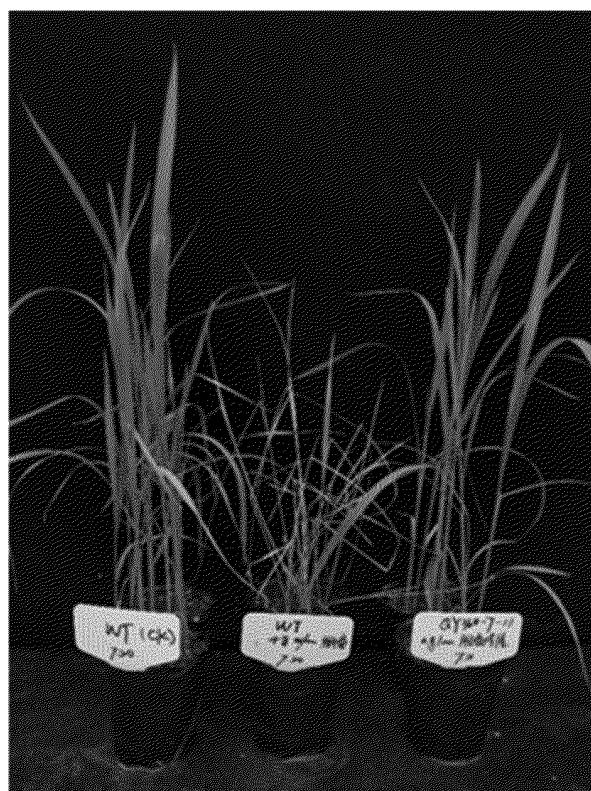


Figure 25

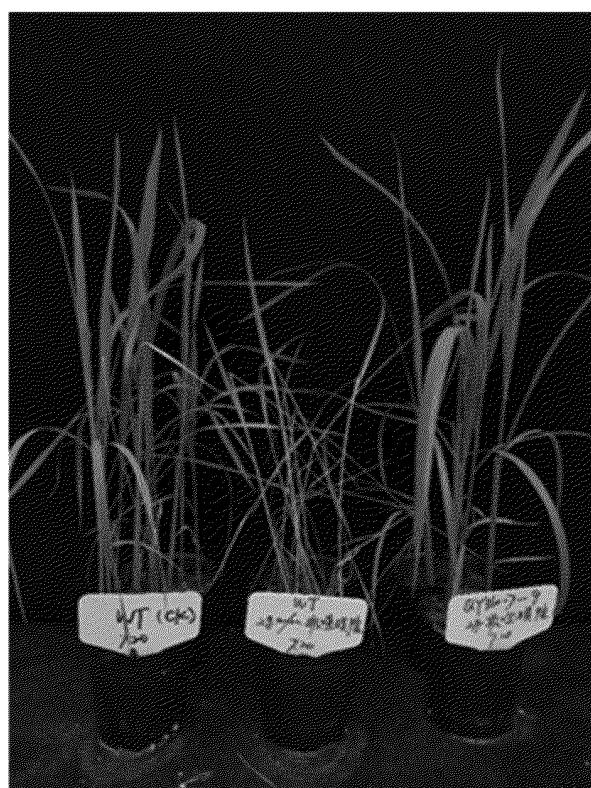


Figure 26

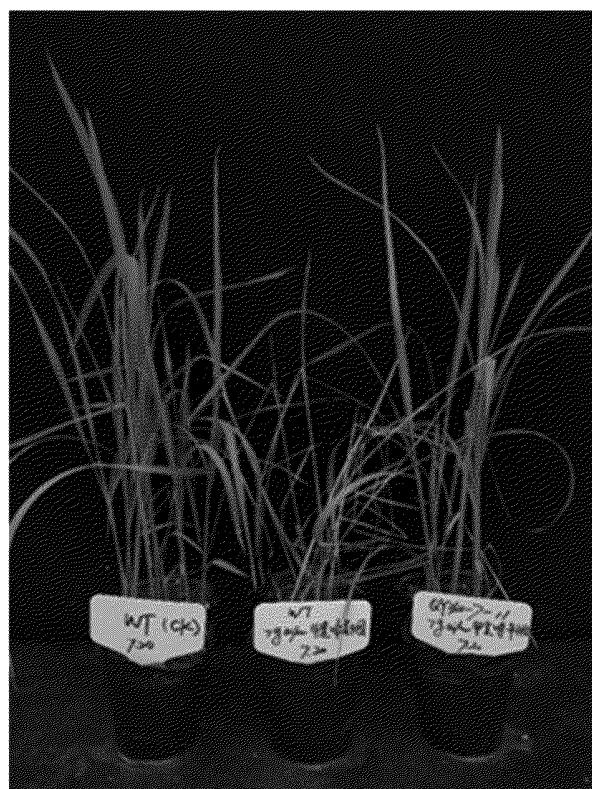


Figure 27

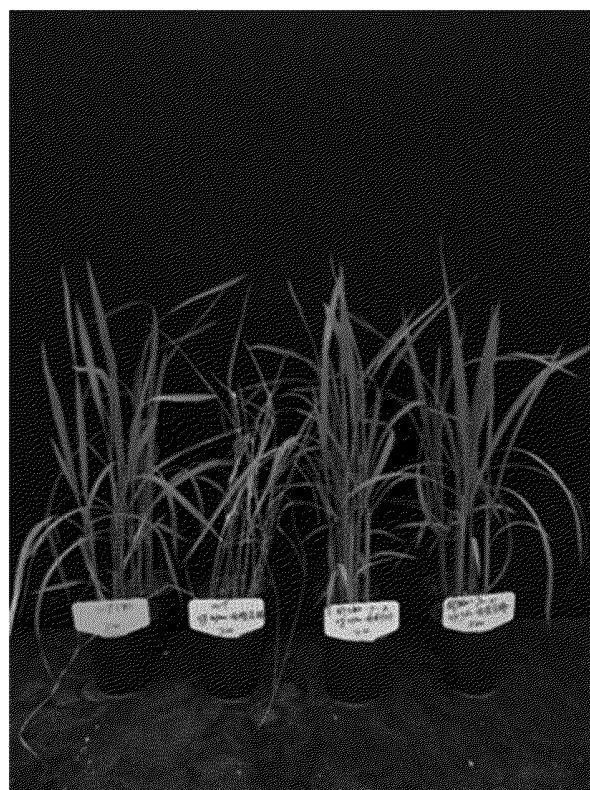


Figure 28

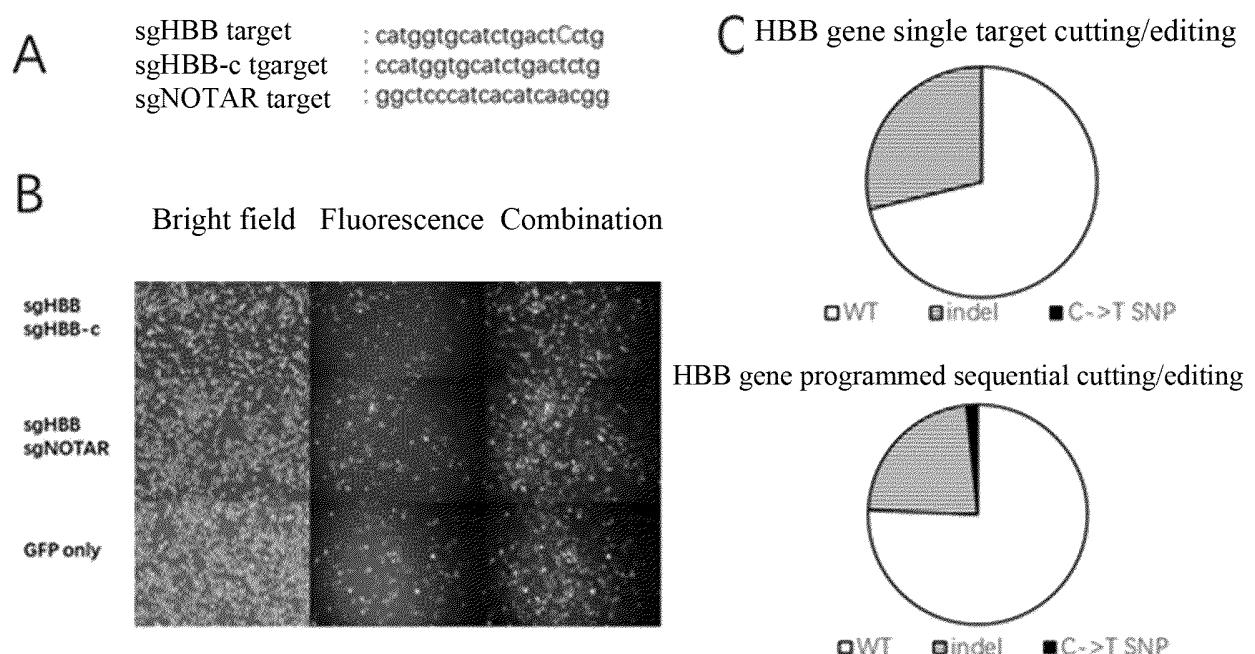


Figure 29

INTERNATIONAL SEARCH REPORT		International application No. PCT/CN2020/120633
5	<b>A. CLASSIFICATION OF SUBJECT MATTER</b> C12N 15/63(2006.01)i; C12N 15/11(2006.01)i  According to International Patent Classification (IPC) or to both national classification and IPC	
10	<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) C12N  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched	
15	Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) database: CNABS, VEN(DWPI+SIP0ABS), CNTXT, EPTXT, USTXT, WOTXT, CNKI, 百度学术, ISI-WEB OF SCIENCE, PubMed: crRNA, sgRNA, gRNA, cas, 基因编辑, gene edit+, 两, 二, two, 多, multiple, several	
20	<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>	
20	Category*	Citation of document, with indication, where appropriate, of the relevant passages
25	X	CN 110373430 A (QINGDAO QINGYUAN COMPOUND CO., LTD.) 25 October 2019 (2019-10-25) see description paragraphs [0004]-[0031], embodiments 1-5
30	A	WO 2017095967 A2 (DUKE UNIVERSITY et al.) 08 June 2017 (2017-06-08) see entire document
35	A	孙毓等(SUN, Yu et al.). "基于CRISPR/Cas9技术构建多重sgRNA实现在人原代T细胞中敲除ADRB2基因(ADRB2 Gene Knockout in Human Primary T Cells by Multiple sgRNAs Constructed using CRISPR/Cas9 Technology)" 中国实验血液学杂志( <i>Journal of Experimental Hematology</i> ), Vol. 27, No. 5, 18 October 2019 (2019-10-18), ISSN: 1009-2137, see pp. 1682-1690
40	A	Wang, ZP et al. "Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwifruit" <i>Plant Biotechnology Journal</i> , Vol. 16, No. 8, 06 February 2018 (2018-02-06), ISSN: 1467-7644, see pp. 1424-1432
45	<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.	
50	* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed  "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family	
55	Date of the actual completion of the international search <b>06 January 2021</b>	Date of mailing of the international search report <b>18 January 2021</b>
55	Name and mailing address of the ISA/CN <b>China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China</b>	Authorized officer
55	Facsimile No. <b>(86-10)62019451</b>	Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/120633

## Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

- 5
1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
    - a.  forming part of the international application as filed:
      - in the form of an Annex C/ST.25 text file.
      - on paper or in the form of an image file.
    - b.  furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
    - c.  furnished subsequent to the international filing date for the purposes of international search only:
      - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
      - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
  - 10 2.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
  - 15 3. Additional comments:

20

25

30

35

40

45

50

55

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/120633

5

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

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

- 10 1.  Claims Nos.: **1-13,25-30,38-65,87**  
because they relate to subject matter not required to be searched by this Authority, namely:  
  - [1] Claims 1-13, 25-30, 38-65, and 87 relate to methods for treatment of the human or animal body by means of surgery or therapy.
- 15 2.  Claims Nos.: **14-24,31-37,86,88-89,94-102** (部分)  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  - [1] The compositional structure of the products to which claims 14-24, 31-37, 86, 88-89, and 94-102 (in part) relate is unclear, and thus a meaningful search can not be carried out.
- 20 3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

- 25 [1] Invention 1: claims 66-85, relating to a method for performing non-transgenic instantaneous editing on a plant genome. Said method relates to sequentially producing two or more DNA breakages at specific positions of a target gene, and separately and spontaneously repairing same, later DNA breakages being produced on the basis of new sequences produced after repair of earlier DNA breakages.
- 30 [2] Invention 2: claims 90-102 (in part), relating to a mutation occurring in a plant and corresponding to the aspartic acid at the position of ALS376 in *Arabidopsis thaliana* being substituted by any other amino acid.
- [3] Invention 3: claims 90-102 (in part), relating to a mutation occurring in a plant and corresponding to the aspartic acid at the position of ALS574 in *Arabidopsis thaliana* being substituted by any other amino acid.
- 35 [4] Invention 4: claims 90-102 (in part), relating to a mutation occurring in a plant and corresponding to the aspartic acid at the position of ALS653 in *Arabidopsis thaliana* being substituted by any other amino acid.
- [5] Invention 5: claims 90-102 (in part), relating to a mutation occurring in a plant and corresponding to the aspartic acid at the position of ALS654 in *Arabidopsis thaliana* being substituted by any other amino acid.
- [6] Invention 6: claims 90-102 (in part), relating to a mutation occurring in a plant and corresponding to the aspartic acid at the position of ACCase2027 in *Alopecurus myosuroides* being substituted by any other amino acid.
- 40 [7] Because invention 1 primarily discusses the steps of a gene editing method, and not a mutation at a specific point, while inventions 2-6 relate to mutations at different points, the inventions described above therefore do not share a same or corresponding special technical feature, do not share a technical relationship, do not form a single general inventive concept, and thus do not comply with PCT Rules 13.1, 13.2, and 13.3.

45

50

55

Form PCT/ISA/210 (continuation of first sheet) (January 2015)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/120633

5

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

- 10            1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
- 15            2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
- 20            3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
- 25            4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **66-85**

- 30
- 35
- 40
- 45
- 50
- 55
- Remark on Protest**
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
  - The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
  - No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2020/120633**

5	Patent document cited in search report		Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)	
10	CN	110373430	A	25 October 2019	WO	2019196717	A1
15	WO	2017095967	A2	08 June 2017	AU	2016362282	A1
20					KR	20180081618	A
25					EP	3384055	A4
30					US	2018353615	A1
35					CN	108779466	A
40					MX	2018005377	A
45					EA	201891317	A2
50					EA	201891317	A3
55					EP	3384055	A2

Form PCT/ISA/210 (patent family annex) (January 2015)

## REFERENCES CITED IN THE DESCRIPTION

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

## Patent documents cited in the description

- CN 201911081617X [0001]
- CN 202010821877 [0001]
- CN 202010974151 [0001]

## Non-patent literature cited in the description

- **SHEN et al.** Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*, 2018 [0005]
- **JINEK et al.** A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012, vol. 337, 816-821 [0017]
- **FINEGOLD ; MARTIN.** Diagnostic Microbiology. V Mosby, 1982, 13-15 [0151]
- **SAMBROOK, J. ; RUSSELL, DAVID W.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, 2001 [0182]
- **EDWIN F. GEORGE ; MICHAEL A. HALL ; GEERT-JAN DE KLERK.** Plant Propagation by Tissue Culture. Springer, 2008 [0182]
- **XING HL ; DONG L ; WANG ZP ; ZHANG HY ; HAN CY ; LIU B ; WANG XC ; CHEN QJ.** A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.*, 29 November 2014, vol. 14 (1), 327 [0184] [0190] [0210]
- **WANG ZP ; XING HL ; DONG L ; ZHANG HY ; HAN CY ; WANG XC ; CHEN QJ.** *Genome Biol.*, 21 July 2015, vol. 16, 144, <https://www.addgene.org/71287> [0184]
- **NISHIMASU et al.** Engineered CRISPR-Cas9 nucleic acid with expanded Targeting space. *Science*, 2018, vol. 361 (6408), 1259-1262 [0184]
- **NISHIMASU et al.** Engineered CRISPR-Cas9 nucleic acid with expanded targeting space. *Science*, 2018, vol. 361 (6408), 1259-1262 [0210]
- **XING HL ; DONG L ; WANG ZP ; ZHANG HY ; HAN CY ; LIU B ; WANG XC ; CHEN QJ.** *BMC Plant Biol.*, 2014 [0211]
- **Sci China Life Sci.**, January 2019, vol. 62 (1), 1-7 [0270]