



- (51) **International Patent Classification:** Not classified
- (21) **International Application Number:** PCT/IB2013/003222
- (22) **International Filing Date:** 20 December 2013 (20.12.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:** 61/745,003 21 December 2012 (21.12.2012) US
- (71) **Applicant:** CELLECTIS [FR/FR]; 8 Rue De La Croix Jarry, F-75013 Paris (FR).
- (72) **Inventors:** MATHIS, Luc; 30 Rue Gambetta, F-94270 Le Kremlin Bicetre (FR). VOYTAS, Daniel, F.; 2197 Folwell Avenue, Falcon Heights, MN 55108 (US). ZHANG, Feng; 4530 Juneau Lane N., Plymouth, MN 55446 (US). CLASEN, Benjamin; 2255 Jewell Lane, South St. Paul, MN 55075 (US). HAUN, William; 1478 Cumberland Street, St. Paul, MN 55117 (US). STODDARD, Thomas; 420 Ford Rd Apt #335, St. Louis Park, MN 55426 (US).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))



WO 2014/096972 A2

(54) **Title:** POTATOES WITH REDUCED COLD-INDUCED SWEETENING

(57) **Abstract:** Materials and methods are provided for making plants (e.g., Solanum varieties) with decreased accumulation of reducing sugars and acrylamide in cold-stored potatoes, specifically, by making TALE-nuclease-induced mutations in genes encoding vacuolar invertase.

POTATOES WITH REDUCED COLD-INDUCED SWEETENING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority from U.S. Provisional Application Serial No. 61/745,003, filed December 21, 2012.

5

TECHNICAL FIELD

This document provides materials and methods for creating potato varieties with reduced cold-induced sweetening.

10

BACKGROUND

Potato (*Solanum tuberosum*) is an important food crop, with worldwide production estimated at 324 million metric tons in 2011 (Food and Agricultural Organization of the United Nations (FAOSTAT), 2010 Crop Production Data, online at faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor). A large proportion of the total potato crop (61% of the 2010 crop in the United States) is used by processors to produce potato chips, French fries and other processed products. In order to have a year-round supply of high-quality raw potatoes for the processing industry, it is necessary to 'cold-store' the potato tubers until they are needed. Cold storage is variety/processor specific, with temperatures ranging from 3°C to 13°C for up to twelve months, which prevents sprouting, reduces losses due to shrinkage/aging, and minimizes the spread of disease.

20

SUMMARY

This document provides materials and methods for creating potato varieties that have reduced cold-induced sweetening (CIS), which is a phenomenon by which starch is converted to the simple reducing sugars, glucose and fructose, during cold storage. Upon processing at high temperatures, the glucose/fructose can interact with free amino acids in a Maillard reaction, which results in bitter, dark-pigmented products that may have

25

increased levels of acrylamide – a suspected neurotoxin/carcinogen. Potato varieties with reduced CIS also are provided.

The disclosure herein is based at least in part on the discovery that potatoes having reduced CIS can be obtained by using a sequence-specific nuclease to make a targeted mutation or knockout in the vacuolar invertase (VInv) gene. The modified potatoes can have improved storage characteristics and reduced levels of acrylamide upon frying, as compared to the levels of acrylamide in non-modified potatoes upon frying after cold storage. Further, the potatoes do not carry any foreign DNA and therefore may be considered by regulatory agencies as non-GM. This document also is based at least in part on the development of potato cultivars with loss-of-function VInv mutations that are created by sequence-specific nucleases.

In one aspect, this document features a *Solanum* plant, plant part, or plant cell comprising a mutation in at least two VInv alleles endogenous to the plant, plant part, or plant cell, such that the plant, plant part, or plant cell has reduced expression of vacuolar invertase as compared to a control *Solanum* plant, plant part, or plant cell that lacks the mutation. Each mutation can be a deletion of more than one nucleotide base pair. Each mutation can be at a target sequence as set forth in SEQ ID NO:27, or a target sequence having at least 95 percent identity to SEQ ID NO:27; or at a target sequence as set forth in SEQ ID NO:1, or a target sequence having at least 95 percent identity to SEQ ID NO:1. The plant, plant part, or plant cell can have been made using a rare-cutting endonuclease [e.g., a transcription activator-like effector endonuclease (TALE-nuclease)]. The TALE-nuclease can bind to a sequence as set forth in any of SEQ ID NOS:18-23. Each of the at two least VInv alleles can exhibit removal of an endogenous nucleic acid and does not include any exogenous nucleic acid. Every endogenous VInv allele can be mutated. Each VInv allele can exhibit removal of an endogenous nucleic acid, without including any exogenous nucleic acid. The plant, plant part, or plant cell may have no detectable expression of vacuolar invertase. The *Solanum* plant, plant part, or plant cell can be a *S. tuberosum* plant, plant part, or plant cell. The plant, plant part, or plant cell can be subjected to cold storage conditions. The plant, plant part, or plant cell

can have decreased levels of acrylamide as compared to a control plant, plant part, or plant cell that lacks the mutation.

In another aspect, this document features a method for making a *Solanum* plant that has reduced cold-induced sweetening. The method can include (a) contacting a population of *Solanum* plant cells containing a functional VInv allele with a rare-cutting endonuclease targeted to an endogenous VInv sequence, (b) selecting, from the population, a cell in which at least two VInv alleles have been inactivated, and (c) growing the selected plant cell into a *Solanum* plant, wherein the *Solanum* plant has reduced cold-induced sweetening as compared to a control *Solanum* plant in which the VInv alleles have not been inactivated. The *Solanum* plant cells can be protoplasts. The method can include transforming the protoplasts with a nucleic acid encoding the rare-cutting endonuclease. The nucleic acid can be an mRNA. The nucleic acid can be contained within a vector. The method can include introducing into the protoplasts a rare-cutting endonuclease protein. The rare-cutting endonuclease can be a TALE-nuclease. The TALE-nuclease can be targeted to a sequence as set forth in SEQ ID NO:27 or to a sequence having at least 95 percent identity to the sequence set forth in SEQ ID NO:27, or can be targeted to a sequence as set forth in SEQ ID NO:1 or to a sequence having at least 95 percent identity to the sequence set forth in SEQ ID NO:1. The TALE-nuclease can bind to a sequence as set forth in any of SEQ ID NOS:18-23. The method can further include culturing protoplasts to generate plant lines. The method can include isolating genomic DNA containing at least a portion of the VInv locus from the protoplasts. The *Solanum* plant cells can be *S. tuberosum* plant cells.

In another aspect, this document features a method for producing a food product. The method can include (a) providing a *Solanum* plant or plant part that (i) contains a mutation in at least two VInv alleles endogenous to the plant or plant part, such that the plant, plant part, or plant cell has reduced expression of vacuolar invertase as compared to a control *Solanum* plant or plant part that lacks the mutation, and (ii) has been subjected to cold storage; and (b) producing a food product from the plant or plant part. The method can further include (c) cooking the plant or plant part to obtain a food product having reduced levels of acrylamide as compared to a food product produced

from a control cooked plant or plant part that lacks the mutation and that was subjected to cold-induced storage prior to being cooked. The cooked plant or plant part can have about the same level of acrylamide as a cooked *Solanum* plant or plant part that was not subjected to cold storage prior to cooking. Each mutation can be at a target sequence as set forth in SEQ ID NO:27 or a target sequence having at least 95 percent identity to SEQ ID NO:27, or at a target sequence as set forth in SEQ ID NO:1 or a target sequence having at least 95 percent identity to SEQ ID NO:1. Each mutation can have been made using a rare-cutting endonuclease (e.g., a TALE-nuclease). The TALE-nuclease can bind to a sequence as set forth in any of SEQ ID NOS:18-23. The *Solanum* plant or plant part can be a *S. tuberosum* plant or plant part. The *Solanum* plant or plant part may have no detectable expression of vacuolar invertase.

In still another aspect, this document features a food product produced from a *Solanum* plant or plant part that (i) contains a mutation in each VInv allele endogenous to the plant or plant part, such that the plant, plant part, or plant cell has no functional VInv allele, and (ii) has been subjected to cold storage. Each mutation can be at a target sequence as set forth in SEQ ID NO:27 or a target sequence having at least 95 percent identity to SEQ ID NO:27, or at a target sequence as set forth in SEQ ID NO:1 or a target sequence having at least 95 percent identity to SEQ ID NO:1. Each mutation can have been made using a rare-cutting endonuclease (e.g., a TALE-nuclease). The TALE-nuclease can bind to a sequence as set forth in any of SEQ ID NOS:18-23. The food product can have been cooked. The food product can have decreased levels of acrylamide as compared to a cooked food product made from a control plant or plant part that lacks the mutation and that was subjected to cold storage prior to being cooked. The cooked food product can have about the same level of acrylamide as a *Solanum* plant or plant part that has not been subjected to cold storage. The *Solanum* plant or plant part can be a *S. tuberosum* plant or plant part (e.g., from a variety selected from the group consisting of Ranger Russet, Atlantic, and Burbank). The food product can be a potato chip or a French fry.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this

invention pertains. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 shows target sites for VInv TALE-nucleases. A DNA sequence from the VInv gene is shown (SEQ ID NO:1). The underlined sequences represent target sites (SEQ ID NOS:18-23) for TALE-nucleases that recognize the VInv gene.

FIG. 2 shows examples of TALE-nuclease-induced mutations in the VInv gene. The top line of each panel shows the DNA sequence of the recognition site for the VInv TALE-nucleases (underlined). The other sequences show representative mutations that were induced by imprecise non-homologous end-joining (NHEJ). Deletion sizes are given on the right.

FIG. 3 shows exemplary alleles of the VInv 2 locus for the variety “Ranger Russet,” which was the germplasm used for development of a mutant plant. Diagnostic single nucleotide polymorphisms (SNPs) that differentiate allele types are underlined and in bold type.

FIG. 4 shows exemplary deletion profiles for a regenerated mutant “Ranger Russet” plant. TALE nuclease recognition sites are underlined, and SNP sites are shaded.

DETAILED DESCRIPTION

The potato genome contains a small family of enzymes called invertases, which play an important role in regulating the carbon partitioning between source tissues

(leaves) and sink tissues (tubers, fruits, seeds). The enzymes irreversibly catalyze the starch → sucrose → glucose/fructose reaction. Plants have three classes of invertase enzymes, but the vacuolar invertase (VInv) is thought to play an important role in CIS.

This document provides potato plant varieties, particularly of the species *S. tuberosum*, that have reduced or even lack VInv activity. Methods for generating such plant varieties, methods for using such plant varieties to produce food products, and food products produced from such plant varieties also are provided.

As used herein, the terms “plant” and “plant part” refer to cells, tissues, organs, seeds, and severed parts (e.g., roots, leaves, and flowers) that retain the distinguishing characteristics of the parent plant. “Seed” refers to any plant structure that is formed by continued differentiation of the ovule of the plant, following its normal maturation point at flower opening, irrespective of whether it is formed in the presence or absence of fertilization and irrespective of whether or not the seed structure is fertile or infertile.

The term “allele(s)” means any of one or more alternative forms of a gene at a particular locus. In a diploid (or amphidiploid) cell of an organism, alleles of a given gene are located at a specific location or locus on a chromosome, with one allele being present on each chromosome of the pair of homologous chromosomes. Similarly, in a tetraploid cell of an organism, one allele is present on each chromosome of the group of four homologous chromosomes. “Heterozygous” alleles are different alleles residing at a specific locus, positioned individually on corresponding homologous chromosomes. “Homozygous” alleles are identical alleles residing at a specific locus, positioned individually on corresponding homologous chromosomes in the cell.

“Wild type” as used herein refers to a typical form of a plant or a gene as it most commonly occurs in nature. A “wild type VInv allele” is a naturally occurring VInv allele (e.g., as found within naturally occurring *S. tuberosum* plants) that encodes a functional VInv protein, while a “non-functional mutant VInv allele” is a VInv allele that does not encode a functional VInv protein. Such a “non-functional mutant VInv allele” can include one or more mutations in its nucleic acid sequence, where the mutation(s) result in no detectable amount of functional VInv protein in the plant or plant cell *in vivo*.

The potato genome usually contains only one VInv gene, but because cultivated potato is a tetraploid, multiple alleles of VInv are present in each variety. The methods provided herein can be used to inactivate at least one (e.g., at least two, at least three, or all four) functional alleles of VInv, thereby removing at least some full-length RNA transcripts and functional VInv protein from potato cells, and in some cases completely removing all full-length RNA transcripts and functional VInv protein.

A representative example of a naturally occurring *S. tuberosum* VInv nucleotide sequence is shown in Table 4 herein. The *S. tuberosum* plants, cells, plant parts, seeds, and progeny thereof that are provided herein have a mutation in each endogenous VInv allele, such that expression of the gene is reduced or completely inhibited. Thus, in some cases, the plants, cells, plant parts, seeds, and progeny do not exhibit detectable levels of vacuolar invertase expressed from the VInv gene.

The plants, plant cells, plant parts, seeds, and progeny provided herein can be generated using a TALE-nuclease system to make a targeted knockout in each allele of the VInv gene. Thus, this document provides materials and methods for using rare-cutting endonucleases (e.g., TALE-nucleases) to generate potato plants and related products (e.g., seeds and plant parts) that are particularly suitable for cold storage before use in making food products for human and animal consumption, due to targeted knockouts in the VInv gene. Other sequence-specific nucleases also may be used to generate the desired plant material, including engineered homing endonucleases or zinc finger nucleases.

The term "rare-cutting endonucleases" herein refer to natural or engineered proteins having endonuclease activity directed to nucleic acid sequences having a recognition sequence (target sequence) about 12-40 bp in length (e.g., 14-40, 15-36, or 16-32 bp in length). Typical rare-cutting endonucleases cause cleavage inside their recognition site, leaving 4 nt staggered cuts with 3'OH or 5'OH overhangs. These rare-cutting endonucleases may be meganucleases, such as wild type or variant proteins of homing endonucleases, more particularly belonging to the dodecapeptide family (LAGLIDADG (SEQ ID NO:28); see, WO 2004/067736) or may result from fusion proteins that associate a DNA binding domain and a catalytic domain with cleavage

activity. TAL-effector endonucleases (TALE-nucleases) and zinc-finger-nucleases (ZFN) are examples of fusions of DNA binding domains with the catalytic domain of the endonuclease *FokI*. Customized TALE-nucleases are commercially available under the trade name TALENTM (Cellestis, Paris, France). For a review of rare-cutting
5 endonucleases, see Baker, *Nature Methods* 9:23-26, 2012).

“Mutagenesis” as used herein refers to processes in which mutations are introduced into a selected DNA sequence. Mutations induced by endonucleases generally are obtained by a double strand break, which results in insertion/deletion mutations (“indels”) that can be detected by deep-sequencing analysis. Such mutations typically are
10 deletions of several base pairs, and have the effect of inactivating the mutated allele. In the methods described herein, for example, mutagenesis occurs via double stranded DNA breaks made by TALE-nucleases targeted to selected DNA sequences in a plant cell. Such mutagenesis results in “TALE-nuclease-induced mutations” (e.g., TALE-nuclease-induced knockouts) and reduced expression of the targeted gene. Following mutagenesis,
15 plants can be regenerated from the treated cells using known techniques (e.g., planting seeds in accordance with conventional growing procedures, followed by self-pollination).

The term “expression” as used herein refers to the transcription of a particular nucleic acid sequence to produce sense or antisense RNA or mRNA, and/or the translation of an mRNA molecule to produce a polypeptide (e.g., a therapeutic protein),
20 with or without subsequent post-translational events.

“Reducing the expression” of a gene or polypeptide in a plant or a plant cell includes inhibiting, interrupting, knocking-out, or knocking-down the gene or polypeptide, such that transcription of the gene and/or translation of the encoded polypeptide is reduced as compared to a corresponding control plant or plant cell in
25 which expression of the gene or polypeptide is not inhibited, interrupted, knocked-out, or knocked-down. Expression levels can be measured using methods such as, for example, reverse transcription-polymerase chain reaction (RT-PCR), Northern blotting, dot-blot hybridization, in situ hybridization, nuclear run-on and/or nuclear run-off, RNase protection, or immunological and enzymatic methods such as ELISA, radioimmunoassay,
30 and western blotting.

In general, a *Solanum* plant, plant part, or plant cell can have its expression of vacuolar invertase reduced by more than 60 percent (e.g., by more than 70 percent, more than 80 percent, or more than 90 percent) as compared to a control *Solanum* plant that lacks the mutation(s). The control *Solanum* plant can be, for example, the wild-type of the *Solanum* plant of which the invertase gene has been mutated.

In some cases, a nucleic acid can have a nucleotide sequence with at least about 75 percent sequence identity to a representative VInv nucleotide sequence. For example, a nucleotide sequence can have at least 75 percent, at least 80 percent, at least 85 percent, at least 90 percent, at least 91 percent, at least 92 percent, at least 93 percent, at least 94 percent, at least 95 percent, at least 96 percent, at least 97 percent, at least 98 percent, or at least 99 percent sequence identity to a representative, naturally occurring VInv nucleotide sequence.

In some cases, a mutation can be at a target sequence as set forth in a VInv sequence as set forth here (e.g., SEQ ID NO:1 or SEQ ID NO:27), or at a target sequence that is at least 95 percent (e.g., at least 96 percent, at least 97 percent, at least 98 percent, or at least 99 percent) identical to the sequence set forth in a VInv sequence as set forth here (e.g., SEQ ID NO:1 or SEQ ID NO:27).

The percent sequence identity between a particular nucleic acid or amino acid sequence and a sequence referenced by a particular sequence identification number is determined as follows. First, a nucleic acid or amino acid sequence is compared to the sequence set forth in a particular sequence identification number using the BLAST 2 Sequences (Bl2seq) program from the stand-alone version of BLASTZ containing BLASTN version 2.0.14 and BLASTP version 2.0.14. This stand-alone version of BLASTZ can be obtained online at fr.com/blast or at ncbi.nlm.nih.gov. Instructions explaining how to use the Bl2seq program can be found in the readme file accompanying BLASTZ. Bl2seq performs a comparison between two sequences using either the BLASTN or BLASTP algorithm. BLASTN is used to compare nucleic acid sequences, while BLASTP is used to compare amino acid sequences. To compare two nucleic acid sequences, the options are set as follows: -i is set to a file containing the first nucleic acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second

nucleic acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastn; -o is set to any desired file name (e.g., C:\output.txt); -q is set to -1; -r is set to 2; and all other options are left at their default setting. For example, the following command can be used to generate an output file containing a comparison between two sequences: C:\BI2seq -i
5 c:\seq1.txt -j c:\seq2.txt -p blastn -o c:\output.txt -q -1 -r 2. To compare two amino acid sequences, the options of BI2seq are set as follows: -i is set to a file containing the first amino acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second amino acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastp; -o is set to any desired file name (e.g., C:\output.txt); and all other options are left at their default
10 setting. For example, the following command can be used to generate an output file containing a comparison between two amino acid sequences: C:\BI2seq -i c:\seq1.txt -j c:\seq2.txt -p blastp -o c:\output.txt. If the two compared sequences share homology, then the designated output file will present those regions of homology as aligned sequences. If the two compared sequences do not share homology, then the designated
15 output file will not present aligned sequences.

Once aligned, the number of matches is determined by counting the number of positions where an identical nucleotide or amino acid residue is presented in both sequences. The percent sequence identity is determined by dividing the number of matches either by the length of the sequence set forth in the identified sequence (e.g.,
20 SEQ ID NO:1), or by an articulated length (e.g., 100 consecutive nucleotides or amino acid residues from a sequence set forth in an identified sequence), followed by multiplying the resulting value by 100. For example, a nucleic acid sequence that has 120 matches when aligned with the sequence set forth in SEQ ID NO:1 is 86.3 percent identical to the sequence set forth in SEQ ID NO:1 (i.e., $120 \div 139 \times 100 = 86.3$). It is
25 noted that the percent sequence identity value is rounded to the nearest tenth. For example, 75.11, 75.12, 75.13, and 75.14 is rounded down to 75.1, while 75.15, 75.16, 75.17, 75.18, and 75.19 is rounded up to 75.2. It also is noted that the length value will always be an integer.

Methods for selecting endogenous target sequences and generating TALE-
30 nucleases targeted to such sequences can be performed as described elsewhere. *See*, for

example, PCT Publication No. WO 2011/072246, which is incorporated herein by reference in its entirety. In some embodiments, software that specifically identifies TALE-nuclease recognition sites, such as TALE-NT 2.0 (Doyle et al., *Nucleic Acids Res* 40:W117-122, 2012) can be used.

5 Transcription activator-like (TAL) effectors are found in plant pathogenic bacteria in the genus *Xanthomonas*. These proteins play important roles in disease, or trigger defense, by binding host DNA and activating effector-specific host genes (*see, e.g.,* Gu et al., *Nature* 435:1122-1125, 2005; Yang et al., *Proc. Natl. Acad. Sci. USA* 103:10503-10508, 2006; Kay et al. *Science* 318:648-651, 2007; Sugio et al., *Proc. Natl. Acad. Sci. USA* 104:10720-10725, 2007; and Römer et al. *Science* 318:645-648, 2007). Specificity depends on an effector-variable number of imperfect, typically 34 amino acid repeats (Schornack et al., *J. Plant Physiol.* 163:256-272, 2006; and WO 2011/072246). Polymorphisms are present primarily at repeat positions 12 and 13, which are referred to herein as the repeat variable-diresidue (RVD).

15 The RVDs of TAL effectors correspond to the nucleotides in their target sites in a direct, linear fashion, one RVD to one nucleotide, with some degeneracy and no apparent context dependence. This mechanism for protein-DNA recognition enables target site prediction for new target specific TAL effectors, as well as target site selection and engineering of new TAL effectors with binding specificity for the selected sites.

20 TAL effector DNA binding domains can be fused to other sequences, such as endonuclease sequences, resulting in chimeric endonucleases targeted to specific, selected DNA sequences, and leading to subsequent cutting of the DNA at or near the targeted sequences. Such cuts (*i.e.,* double-stranded breaks) in DNA can induce mutations into the wild type DNA sequence via NHEJ or homologous recombination, for example. In some cases, TALE-nucleases can be used to facilitate site directed mutagenesis in complex genomes, knocking out or otherwise altering gene function with great precision and high efficiency. As described in the Examples below, TALE-nucleases targeted to the *S. tuberosum* VInv gene can be used to mutagenize the endogenous gene, resulting in plants without detectable expression of VInv. The fact that
25 some endonucleases (*e.g., FokI*) function as dimers can be used to enhance the target
30

specificity of the TALE-nuclease. For example, in some cases a pair of TALE-nuclease monomers targeted to different DNA sequences (e.g., the underlined target sequences shown in FIG. 1) can be used. When the two TALE-nuclease recognition sites are in close proximity, as depicted in FIG. 1, the inactive monomers can come together to create a functional enzyme that cleaves the DNA. By requiring DNA binding to activate the nuclease, a highly site-specific restriction enzyme can be created.

Methods for using TALE-nucleases to generate potato plants, plant cells, or plant parts having mutations in endogenous genes include, for example, those described in the Examples herein. For example, one or more nucleic acids encoding TALE-nucleases targeted to selected VInv sequences (e.g., the VInv sequences shown in FIG. 1) can be transformed into plant cells (e.g., protoplasts), where they can be expressed. In some cases, one or more TALE-nuclease proteins can be introduced into plant cells (e.g., protoplasts). The cells, or a plant cell line or plant part generated from the cells, can subsequently be analyzed to determine whether mutations have been introduced at the target site(s), through nucleic acid-based assays or protein-based assays to detect expression levels as described above, for example, or using nucleic acid-based assays (e.g., PCR and DNA sequencing, or PCR followed by a T7E1 assay; Mussolino et al., *Nucleic Acids Res.* 39:9283-9293, 2011) to detect mutations at the genomic loci. In a T7E1 assay, genomic DNA can be isolated from pooled calli, and sequences flanking TALE-nuclease recognition sites for VInv can be PCR-amplified. Amplification products then can be denatured and re-annealed. If the re-annealed fragments form a heteroduplex, T7 endonuclease I cuts at the site of mismatch. The digested products can be visualized by gel electrophoresis to quantify mutagenesis activity of the TALE-nuclease.

More recently, a new genome engineering tool has been developed based on the RNA-guided Cas9 nuclease from the type II prokaryotic CRISPR (Clustered Regularly Interspaced Short palindromic Repeats) adaptive immune system (*see, e.g.,* Belahj et al., *Plant Methods* 9:39, 2013). This system allows for cleaving DNA sequences that are flanked by a short sequence motif, referred as proto-spacer adjacent motif (PAM). Cleavage is achieved by engineering a specific crRNA that is complementary to the

target sequence, which associates into the living cell with the endonuclease Cas9 from *S. pyogenes* that is heterologously expressed. In the crRNA/Cas9 complex, a dual tracrRNA:crRNA structure acts as guide RNA that directs the endonuclease Cas9 to the cognate target sequence. Since there are several PAM motifs present in the nucleotide sequence of the *Vinv* gene, crRNA specific to *Vinv* gene may be designed to introduce mutations or to inactivate all or part of the *Vinv* gene alleles within *Solanum* plant cells in which the Cas9 endonuclease and the crRNA are transfected and expressed. This approach can be used as an alternative to TALE-nucleases in some instances, to obtain the plants as described herein.

This document also encompasses further mutations that could be introduced in other *Solanum* genes so as to, for example:

- provide further acrylamide reduction by modifying the expression of genes involved in asparagine synthesis;
- prevent black spot bruise by reducing polyphenol oxidase-5 expression;
- prevent Potato Virus Y by reducing eIF4E gene expression;
- prevent late blight; or
- improve nematode, herbicide, or insect resistance.

Thus, the methods provided herein can be used to obtain gene stacking in a *Solanum* trait.

This disclosure also provides methods for producing food products using potato plant varieties with reduced CIS, as well as food products made by such methods. The methods provided herein can include, for example, providing or making *S. tuberosum* plants or plant parts that contain a TALE-nuclease-induced mutation in two or more endogenous *VInv* alleles and that have been subjected to cold storage, and using standard cooking and/or manufacturing methods to produce a food product (including, without limitation, potato chips, French fries, potato flakes, and mashed potatoes) from the plants or plant parts. In some embodiments, the reduced CIS can be observed as a reduction in bitterness and/or dark-pigmentation as compared to the bitterness and/or pigmentation observed in food products made from control plants or plant parts that do not contain the mutated *VInv* alleles and that have been subjected to cold storage. In some

embodiments, the food products (e.g., food products made using methods that include cooking the plants or plant parts) can have reduced acrylamide levels as compared to the levels of acrylamide in food products made from *S. tuberosum* plants or plant parts that do not have mutations in the endogenous VInv alleles and that have been subjected to cold storage (e.g., prior to cooking). In some cases, the food products can have levels of acrylamide that are comparable to the levels of acrylamide in food products made from *S. tuberosum* plants or plant parts that were not subjected to cold storage.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

Example 1 – Engineering sequence-specific nucleases to mutagenize the VInv gene

To completely inactivate or knock-out the alleles of the VInv gene in *S. tuberosum*, sequence-specific nucleases were designed that target the protein coding region in the vicinity of the start codon. Three TALE-nuclease pairs were designed to target the VInv gene family within the first 200 bp of the coding sequence using software that specifically identifies TALE-nuclease recognition sites. The TALE-nuclease recognition sites for the VInv genes are underlined in FIG. 1 and are listed in Table 1. TALE-nucleases were synthesized using methods similar to those described elsewhere (Cermak et al., *Nucleic Acids Res.* 39:e82, 2011; Reyon et al., *Nat. Biotechnol.* 30:460-465, 2012; and Zhang et al., *Nat. Biotechnol.* 29:149-153, 2011).

Example 2 – VInv TALE-nuclease activity in yeast

To assess the activity of the TALE-nucleases targeting the VInv gene, activity assays were performed in yeast by methods similar to those described elsewhere (Christian et al., *Genetics* 186:757-761, 2010). For these assays, a target plasmid was constructed with the TALE-nuclease recognition site cloned in a non-functional β -galactosidase reporter gene. The target site was flanked by a direct repeat of β -galactosidase coding sequence such that if the reporter gene was cleaved by the TALE-nuclease, recombination would occur between the direct repeats and function would be

restored to the β -galactosidase gene. β -galactosidase activity, therefore, served as a measure of TALE-nuclease cleavage activity.

In the yeast assay, all of the VInv TALE-nuclease pairs (VInv_T01, VInv_T02 and VInv_T03) exhibited high cleavage activity under two distinct temperature
5 conditions (i.e., 37°C and 30°C). Cleavage activities were normalized to the benchmark nuclease, I-SceI. Results are summarized in Table 2.

Example 3 – Activity of VInv TALE-nucleases at their endogenous target sites in *S. tuberosum*

10 TALE-nuclease activity at endogenous target sites in *S. tuberosum* was measured by expressing the TALE-nucleases in protoplasts and surveying the TALE-nuclease target sites for mutations introduced by NHEJ. Methods for protoplast preparation were performed as described elsewhere (Shepard, in: Genetic Improvement of Crops/Emergent Techniques (pp.185-219), Rubenstein, Gengenbach, Philips, and Green (Eds.), Univ. of
15 Minnesota Press, Minneapolis, MN, 1980; and Shepard and Totten, *Plant Physiol.* 60:313-316, 1977). Briefly, *S. tuberosum* mini tubers were planted in moistened vermiculite and grown under low light conditions for 3-5 weeks. Young, fully expanded leaves were collected and surface sterilized, and protoplasts were isolated.

TALE-nuclease-encoding plasmids, together with a YFP-encoding plasmid, were
20 introduced into *S. tuberosum* protoplasts by PEG-mediated transformation as described elsewhere (Yoo et al., *Nature Protocols* 2:1565-1572, 2007). Twenty-four hours after treatment, transformation efficiency was measured by evaluating an aliquot of the transformed protoplasts using a fluorescent microscope to monitor YFP fluorescence. The remainder of the transformed protoplasts was harvested, and genomic DNA was
25 prepared using a CTAB-based method. Using genomic DNA prepared from the protoplasts as a template, a 272-bp fragment encompassing the TALE-nuclease recognition site was amplified by PCR. Allele types were analyzed by individual clonal direct sequencing and 454 pyro-sequencing. Sequencing reads with indel mutations in the spacer region were considered as having been derived from imprecise repair of a
30 cleaved TALE-nuclease recognition site by NHEJ. Mutagenesis frequency was

calculated as the number of sequencing reads with NHEJ mutations out of the total sequencing reads. The values were then normalized by the transformation efficiency.

The activity of the VInv TALE-nuclease pairs, VInv_T01, VInv_T02 and VInv_T03, against their target gene is summarized in Table 3. The TALE-nucleases induced NHEJ mutations in VInvT1, VInvT2, and VInvT3, ranging from 3.6% to 9.5%. Examples of TALE-nuclease-induced mutations in VInvT1, VInvT2, and VInvT3 are shown in FIG. 2.

Example 4 – Regeneration of *S. tuberosum* lines with TALE-nuclease-induced mutations

in VInv

S. tuberosum lines were created with mutations in one or more alleles of the VInv gene. Protoplasts were isolated from surface sterilized leaves, and transformed with plasmids encoding one of the following: (i) TALE-nuclease VInv_T01 (ii) TALE-nuclease VInv_T02; (iii) TALE-nuclease VInv_T03; or (iv) YFP. Transformation efficiencies were monitored by the delivery of the YFP plasmid, which is visualized using a fluorescent microscope or by flow cytometry.

After PEG-mediated transformation, protoplasts were cultured using methods and media described elsewhere (Gamborg et al., in: Plant Tissue Culture Methods and Applications in Agriculture (pp.115-153), Thorpe (Ed.), Academic Press, Inc., New York, NY, 1981), with slight modifications. Protoplasts were re-suspended in liquid plating medium at a cell density of 1×10^5 /ml in a small petri dish, and stored at 25°C in the dark. At day 14 after transformation, when the majority of the protoplasts had divided at least once, the protoplast culture was diluted two-fold in a suspension of P.-medium. At day 28 after transformation, the protoplast cultures were plated on a solid reservoir (10 ml) of CUL medium (Haberlach et al., *Plant Science* 39:67-74, 1985). At this point, protoplast-derived calli were visible to the eye.

At day 65 after transformation, protoplast-derived calli identified as mutants (e.g., using methods as described in Example 5) were transferred to a solid reservoir of DIF medium (Haberlach et al., *supra*). Calli were transferred to fresh DIF medium at biweekly intervals. As shoots formed, they were excised and placed into a solid reservoir

assay is employed using previously validated methods (Bethke and Busse, *Am. J. Potato Res.* 85:414-421, 2008).

Table 1
TALE-nuclease target sequences

Gene	Target Sequence Left	SEQ ID NO:	Target Sequence Right	SEQ ID NO:
Vinv_T1	TTCCTCCCGGATCAACC	18	GAAGTCCCTTAAAATCA	19
Vinv_T2	TTCCTCTCCTCTTTCCT	20	CTTCTTCCGATCCTCA	21
Vinv_T3	TAGCCTTCTTCCCGATC	22	CCGGACTTGCAGAGTAA	23

Table 2
VInv TALE-nuclease activity in yeast

TALE-nuclease Pair Name	TALE-nuclease Target Sequence	SEQ ID NO:	Activity in Yeast*
VInv_T01	TTCCTCCCGGATCAACCCGATTCCGGCCACCGAAGTCCCTTAAAATCA	24	37°C 0.94 30°C 0.95
VInv_T02	TTCCTCTCCTCTTTCCTTTTGCTTCTGTAGCCTTCTTCCGATCCTCA	25	0.92 0.89
VInv_T03	TAGCCTTCTTCCCGATCCTCAACAACCCGACTTGCAGAGTAA	26	0.96 0.82

*Normalized to I-SceI (max = 1.0)

Table 3
454 Pyro-Sequencing Data for VInv TALE-nuclease

TALE-nuclease name	Location of target site	NHEJ mutagenesis freq. with TALE-nuclease*
VInv_T01	VInvT1	3.6% (4614)
VInv_T02	VInvT2	9.5% (4957)
VInv_T03	VInvT3	9.9% (3350)

* The total number of 454 sequencing reads used for this analysis was indicated in parentheses.

Table 4*S. tuberosum* VInv complete CDS; GenBank JN661860; SEQ ID NO:27)

5 ATGGCCACGCAGTACCATTCCAGTTATGACCCGGAAAACCTCCGCCTCCCATTACACATT
 CCTCCCGGATCAACCCGATTCCGGCCACCGGAAGTCCCTTAAAATCATCTCCGGCATTTC
 TCCTCTCCTCTTTCCTTTTGCTTCTGTAGCCTTCTTCCGATCCTCAACAACCAGTCA
 CCGGACTTGCAGAGTAACTCCCGTTCGCCGGCGCCCGTCAAGAGGTGTTTCTCAGGG
 AGTCTCCGATAAGACTTTTCGAGATGTCGTCAATGCTAGTCACGTTTCTTATGCGTGGT
 CCAATGCTATGCTTAGCTGGCAAAGAACTGCTTACCATTTTCAACCTCAAAAAAATTGG
 10 ATGAACGATCCTAATGGTCCATTGTACCACAAGGGATGGTATCATCTTTTTTATCAATA
 CAATCCAGATTCAGCTATTTGGGGAAAATATCACATGGGGCCATGCCGTATCCAAGGACT
 TGATCCACTGGCTCTACTTGCCTTTTGCCATGGTTCCTGATCAATGGTACGATATTAAC
 GGTGTCTGGACTGGGTCCGCTACCATCCTACCCGATGGTCAGATCATGATGCTTTATAC
 CGGTGACACTGATGATTATGTGCAAGTGCAAAATCTTGCGTACCCCACTTATCTG
 15 ATCCTCTCCTTCTAGACTGGGTCAAGTACAAAGGCAACCCGGTTCTGGTTCCCTCCACCC
 GGCATTGGTGTCAAGGACTTTAGAGACCCGACCACTGCTTGGACCCGACCCCAAATGG
 GCAATGGCTCTTAACAATCGGGTCTAAGATTGGTAAAACGGGTATTGCACTTGTTTATG
 AACTTCCAACCTCACAAGCTTTAAGCTATTGGATGAAGTGCTGCATGCGGTTCCGGGT
 ACGGGTATGTGGGAGTGTGTGGACTTTTACCCGGTATCGACTGAAAAACAAACGGGTT
 20 GGACACATCATATAACGGCCCGGGTGTAAAGCATGTGTTAAAAGCAAGTTTAGATGACA
 ATAAGCAAGATCACTATGCTATTGGGACGTATGACTTGACAAAGAACAATGGACACCC
 GATAAGCCGGAATTGGATTGTGGAATTGGGTTGAAGCTGGATTATGGGAAATATTATGC
 ATCAAAGACATTTTATGACCCGAAGAAACAACGAAGAGTACTGTGGGGATGGATTGGGG
 AACTGATAGTGAATCTGCTGACCTGCAGAAGGGATGGGCATCTGTACAGAGTATTCCA
 25 AGGACAGTGCTTTACGACAAGAAGACAGGGACACATCTACTTCAGTGGCCAGTTGAAGA
 AATTGAAAGCTTAAGAGCGGGTGATCCTATTGTTAAGCAAGTCAATCTTCAACCAGGTT
 CAATTGAGCTACTCCATGTTGACTCAGCTGCAGAGTTGGATATAGAAGCCTCATTTGAA
 GTGGACAAAGTTCGCGCTCCAGGGAATAATTGAAGCAGATCATGTAGGTTTTCAGCTGCTC
 TACTAGTGGAGGTGCTGCTAGCAGAGGCATTTTGGGACCATTTGGTGTGCTTGTAAATTG
 30 CTGATCAAACGCTATCTGAGCTAACGCCAGTTTACTTCTTCATTTCTAAAGGAGCTGAT
 GGTGAGCTGAGACTCACTTCTGTGCTGATCAAACCTAGATCCTCAGAGGCTCCGGGAGT
 TGCTAAACGAGTTTATGGTAGTTCAGTACCCGTGTTGGACGGTGAAAAACATTCGATGA
 GATTATTGGTGGACCACTCAATTGTGGAGAGCTTTGCTCAAGGAGGAAGAACAGTCATA
 ACATCGCGAATTTACCCAACAAGGCAGTGAATGGAGCAGCACGACTCTTCGTTTTCAA
 35 TAATGCCACAGGGGCTAGCGTGACTGCCTCCGTCAAGATTTGGTCACTTGAGTCGGCTA
 ATATTCGATCCTTCCCCTTGCAAGACTTGTA

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended
5 claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A *Solanum* plant, plant part, or plant cell comprising a mutation in at least two VInv alleles endogenous to said plant, plant part, or plant cell, such that said plant, plant part, or plant cell has reduced expression of vacuolar invertase as compared to a control *Solanum* plant, plant part, or plant cell that lacks said mutation.
2. The plant, plant part, or plant cell of claim 1, wherein each said mutation is a deletion of more than one nucleotide base pair.
3. The plant, plant part, or plant cell of claim 1, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:27, or at a target sequence having at least 95 percent identity to SEQ ID NO:27.
4. The plant, plant part, or plant cell of claim 1, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:1, or at a target sequence having at least 95 percent identity to SEQ ID NO:1.
5. The plant, plant part, or plant cell of claim 1, wherein said plant, plant part, or plant cell was made using a rare-cutting endonuclease.
6. The plant, plant part, or plant cell of claim 5, wherein said rare-cutting endonuclease is a transcription activator-like effector endonuclease (TALE-nuclease).
7. The plant, plant part, or plant cell of claim 6, wherein said TALE-nuclease binds to a sequence as set forth in any of SEQ ID NOS:18-23.
8. The plant, plant part, or plant cell of claim 1, wherein each of said at least two VInv alleles exhibits removal of an endogenous nucleic acid and does not include any exogenous nucleic acid.
9. The plant, plant part, or plant cell of claim 1, wherein every endogenous VInv allele is mutated.

10. The plant, plant part, or plant cell of claim 9, wherein each said VInv allele exhibits removal of an endogenous nucleic acid and does not include any exogenous nucleic acid.
11. The plant, plant part, or plant cell of claim 9, wherein said plant, plant part, or plant cell has no detectable expression of vacuolar invertase.
12. The plant, plant part, or plant cell of claim 1, wherein said *Solanum* plant, plant part, or plant cell is a *S. tuberosum* plant, plant part, or plant cell.
13. The plant, plant part, or plant cell of claim 1, wherein said plant, plant part, or plant cell is subjected to cold storage conditions.
14. The plant, plant part, or plant cell of claim 13, wherein said plant, plant part, or plant cell has decreased levels of acrylamide as compared to a control plant, plant part, or plant cell that lacks said mutation.
15. A method for making a *Solanum* plant that has reduced cold-induced sweetening, wherein said method comprises:
 - (a) contacting a population of *Solanum* plant cells comprising a functional VInv allele with a rare-cutting endonuclease targeted to an endogenous VInv sequence,
 - (b) selecting, from said population, a cell in which at least two VInv alleles have been inactivated, and
 - (c) growing said selected plant cell into a *Solanum* plant, wherein said *Solanum* plant has reduced cold-induced sweetening as compared to a control *Solanum* plant in which said VInv alleles have not been inactivated.
16. The method of claim 15, wherein said *Solanum* plant cells are protoplasts.
17. The method of claim 16, comprising transforming said protoplasts with a nucleic acid encoding said rare-cutting endonuclease.
18. The method of claim 17, wherein said nucleic acid is an mRNA.

19. The method of claim 17, wherein said nucleic acid is contained within a vector.
20. The method of claim 15, comprising introducing into said protoplasts a rare-cutting endonuclease protein.
21. The method of any one of claims 15 to 20, wherein said rare-cutting endonuclease is a TALE-nuclease.
22. The method of claim 21, wherein said TALE-nuclease is targeted to a sequence as set forth in SEQ ID NO:27, or to a sequence having at least 95 percent identity to the sequence set forth in SEQ ID NO:27.
23. The method of claim 21, wherein said TALE-nuclease is targeted to a sequence as set forth in SEQ ID NO:1, or to a sequence having at least 95 percent identity to the sequence set forth in SEQ ID NO:1.
24. The method of claim 21, wherein said TALE-nuclease binds to a sequence as set forth in any of SEQ ID NOS:18-23.
25. The method of claim 16, further comprising culturing said protoplasts to generate plant lines.
26. The method of claim 16, comprising isolating genomic DNA comprising at least a portion of the VInv locus from said protoplasts.
27. The method of claim 15, wherein said *Solanum* plant cells are *S. tuberosum* plant cells.
28. A method for producing a food product, comprising:
 - (a) providing a *Solanum* plant or plant part that (i) comprises a mutation in at least two VInv alleles endogenous to said plant or plant part, such that said plant, plant part, or plant cell has reduced expression of vacuolar invertase as compared to a control *Solanum* plant or plant part that lacks said mutation, and (ii) has been subjected to cold storage; and

(b) producing a food product from said plant or plant part.

29. The method of claim 28, further comprising:

(c) cooking said plant or plant part to obtain a food product having reduced levels of acrylamide as compared to a food product produced from a control cooked plant or plant part that lacks said mutation and that was subjected to cold-induced storage prior to being cooked.

30. The method of claim 29, wherein said cooked plant or plant part has about the same level of acrylamide as a cooked *Solanum* plant or plant part that was not subjected to cold storage prior to cooking.

31. The method of claim 28, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:27, or at a target sequence having at least 95 percent identity to SEQ ID NO:27.

32. The method of claim 28, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:1, or at a target sequence having at least 95 percent identity to SEQ ID NO:1.

33. The method of claim 28, wherein each said mutation was made using a rare-cutting endonuclease.

34. The method of claim 33, wherein said rare-cutting endonuclease is a TALE-nuclease.

35. The method of claim 34, wherein said TALE-nuclease binds to a sequence as set forth in any of SEQ ID NOS:18-23.

36. The method of claim 28, wherein said *Solanum* plant or plant part is a *S. tuberosum* plant or plant part.

37. The method of claim 28, wherein said *Solanum* plant or plant part has no detectable expression of vacuolar invertase.

38. A food product produced from a *Solanum* plant or plant part that (i) comprises a mutation in each VInv allele endogenous to said plant or plant part, such that said plant, plant part, or plant cell has no functional VInv allele, and (ii) has been subjected to cold storage.
39. The food product of claim 38, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:27, or at a target sequence having at least 95 percent identity to SEQ ID NO:27.
40. The food product of claim 38, wherein each said mutation is at a target sequence as set forth in SEQ ID NO:1, or at a target sequence having at least 95 percent identity to SEQ ID NO:1.
41. The food product of claim 38, wherein each said mutation was made using a rare-cutting endonuclease.
42. The food product of claim 41, wherein said rare-cutting endonuclease is a TALE-nuclease.
43. The food product of claim 42, wherein said TALE-nuclease binds to a sequence as set forth in any of SEQ ID NOS:18-23.
44. The food product of claim 38, wherein said food product has been cooked.
45. The food product of claim 44, wherein said food product has decreased levels of acrylamide as compared to a cooked food product made from a control plant or plant part that lacks said mutation and that was subjected to cold storage prior to being cooked.
46. The food product of claim 44, wherein said cooked food product has about the same level of acrylamide as a *Solanum* plant or plant part that has not been subjected to cold storage.
47. The food product of claim 38, wherein said *Solanum* plant or plant part is a *S. tuberosum* plant or plant part.

48. The food product of claim 38, wherein said food product is a potato chip or a French fry.

Figure 1

VInv_T1

AttcctcccggaatcaaacCgattccggccaccggaagtcccttaaaatcatctccggcatcttcctctctcttctct
 tttgctttctgtagccttcttccgatacctcaacaaccaGtcaccggacttgacagagtaac (SEQ ID NO:1)

VInv_T2

AttcctcccggaatcaaacCgattccggccaccggaagtcccttaaaatcatctccggcatcttcctctctcttctct
 tttgctttctgtagccttcttccgatacctcaacaaccaGtcaccggacttgacagagtaac (SEQ ID NO:1)

VInv_T3

AttcctcccggaatcaaacCgattccggccaccggaagtcccttaaaatcatctccggcatcttcctctctcttctct
 tttgctttctgtagccttcttccgatacctcaacaaccaGtcaccggacttgacagagtaac (SEQ ID NO:1)

Figure 2

VInv_T1			
Sequence			
<u>TTCTCCCGGATCAACCCGATTC</u> <u>CGGCCACCGGAAGTCCCTTAAAAATCA</u>			
<u>TTCTCCCGGATCAACTCCCTT</u> ----- <u>AAATCA</u>	0	2	
TTCTCCCGGATCAGCACCGGAAATC----- <u>CTTAAAACTCA</u>	21	3	
TTCTCCCGGATCAACCCGATT----- <u>GTCCCTTAAAAATCA</u>	10	4	
TTCTCCCGGATCAACCCGAT----- <u>TTCCCTTAAAAATCA</u>	13	5	
TTCTCCCGGATCAACCCGAT----- <u>TTCCCTTAAAAATCA</u>	38	6	
TTCTCCCGGATCAACCCGATTCGGC----- <u>TTCCCTTAAAAATCA</u>	30	7	
VInv_T2			
Sequence			
<u>TTCTCTCCTCTTTCCCTTTTGCTTTCTGTAGCCTTCTTTCCGATCCTCA</u>			
<u>TTCTCTCCTCTTTCCCTTTTGCT</u> ----- <u>AGCCTTCTTTCCGATCCTCA</u>	0	8	
TTCTCTCCTCTCTCCTTTTGCTTT--- <u>TAGTCTTCTTTCCGATCCTCA</u>	6	9	
TTCTCTCCTCTTTCCCTTTTGCTGT--- <u>AGCCTTCTTTCCGATCCTCA</u>	3	10	
TTCTCTCCTCTTTCCCTTTTGCTGT--- <u>AGCCTTCTTTCCGATCCTCA</u>	4	11	
TTCTCTCCTCTTTCCCTTTTGCTGT--- <u>AGCCTTCTTTCCGATCCTCA</u>	6	9	
TTCTCTCCTCTTTCCCTTTTGCGT----- <u>AGCCTTCTTTCCGATCCTCA</u>	5	12	
VInv_T3			
Sequence			
<u>TAGCCTTCTTTCCGATCCTCAACAAC</u> <u>CAGTCAACCGGACTTGCAGAGTAA</u>			
<u>TAGCCTTCTTTCCGATCCTCAAC</u> ----- <u>ACACCGGACTTGCAGAGTAA</u>	0	13	
<u>TAGCCTTCTTTCCGATCCTCAAA</u> ----- <u>GTACCGGACTTGCAGAGTAA</u>	6	14	
<u>TAGCCTTCTTTCCGATCCTC</u> ----- <u>TCACCGGACTTGCAGAGTAA</u>	5	15	
<u>TAGCCTTCTTTCCGATCCTCAG</u> ----- <u>TCACCGGACTTGCAGAGTAA</u>	9	16	
<u>TAGCCTTCTTTCCGATCCTCAG</u> ----- <u>TCACCGGACTTGCAGAGTAA</u>	7	17	
<u>TAGCCTTCTTTCCGATCCTCAG</u> ----- <u>TCACCGGACTTGCAGAGTAA</u>	7	17	

Figure 4

Vinv_T2 Site	Number of Deletions	SEQ ID NO:
TTCCCTCCTCCTCTCTCCCTTTTGCTTTTCTTTAGCTTCTTTCCGATCCCTCA	0	31
TTCCCTCCTCCTCTCTCCCTTTTGCTTTTCTTTAGCTTCTTTCCGATCCCTCA	4	32
TTCCCTCCTCCTCTCTCCCTTTTGCTTTTCTTTAGCTTCTTTCCGATCCCTCA	4	33
TTCCCTCCTCCTCTCTCCCTTTTGCTTTTCTTTAGCTTCTTTCCGATCCCTCA	17	34