



(12) **DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

(13) **A1**

(86) **Date de dépôt PCT/PCT Filing Date:** 2022/09/21
(87) **Date publication PCT/PCT Publication Date:** 2023/03/30
(85) **Entrée phase nationale/National Entry:** 2024/03/19
(86) **N° demande PCT/PCT Application No.:** US 2022/076746
(87) **N° publication PCT/PCT Publication No.:** 2023/049720
(30) **Priorité/Priority:** 2021/09/21 (US63/246,512)

(51) **Cl.Int./Int.Cl.** **C07K 14/415** (2006.01),
C12N 15/82 (2006.01), **C12N 9/22** (2006.01)
(71) **Demandeur/Applicant:**
PAIRWISE PLANTS SERVICES, INC., US
(72) **Inventeurs/Inventors:**
CRAWFORD, BRIAN CHARLES WILDING, US;
GRAHAM, NATHANIEL, US;
MATHEW, LOLITA GEORGE, US
(74) **Agent:** AIRD & MCBURNEY LP

(54) **Titre : PROCÉDES ET COMPOSITIONS POUR REDUIRE L'ÉCLATEMENT DE LA COSSE DANS LE CANOLA**
(54) **Title: METHODS AND COMPOSITIONS FOR REDUCING POD SHATTER IN CANOLA**

(57) **Abrégé/Abstract:**

This invention relates to compositions and methods for modifying SHATTERPROOF MADS-BOX (SHP) genes in canola plants, optionally to reduce pod shattering. The invention further relates to canola plants having reduced pod shatter produced using the methods and compositions of the invention.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

30 March 2023 (30.03.2023)



(10) International Publication Number

WO 2023/049720 A1

(51) International Patent Classification:

C07K 14/415 (2006.01) C12N 15/82 (2006.01)
C12N 9/22 (2006.01)

(21) International Application Number:

PCT/US2022/076746

(22) International Filing Date:

21 September 2022 (21.09.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/246,512 21 September 2021 (21.09.2021) US

(71) Applicant: PAIRWISE PLANTS SERVICES, INC.

[US/US]; 807 East Main Street, Suite 4-100, Durham, North
Carolina 27701 (US).

(72) Inventors: CRAWFORD, Brian Charles Wilding; 1404

Cavalcade Drive, Cary, North Carolina 27519 (US). GRA-
HAM, Nathaniel; 701 Lyon Tree Lane, Durham, North
Carolina 27713 (US). MATHEW, Lolita George; 313
Bolton Grant Drive, Cary, North Carolina 27519 (US).

(74) Agent: BONNEN, Alice M.; Myers Bigel, P.A., P.O. Box

37428, Raleigh, North Carolina 27627 (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, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE,
KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
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, ST, 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:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: METHODS AND COMPOSITIONS FOR REDUCING POD SHATTER IN CANOLA

(57) Abstract: This invention relates to compositions and methods for modifying *SHATTERPROOF MADS-BOX (SHP)* genes in canola plants, optionally to reduce pod shattering. The invention further relates to canola plants having reduced pod shatter produced using the methods and compositions of the invention.

WO 2023/049720 A1

METHODS AND COMPOSITIONS FOR REDUCING POD SHATTER IN CANOLA

STATEMENT OF PRIORITY

5 This application claims the benefit, under 35 U.S.C. § 119 (e), of U.S. Provisional Application No. 63/246,512 filed on September 21, 2021, the entire contents of which is incorporated by reference herein.

STATEMENT REGARDING ELECTRONIC FILING OF A SEQUENCE LISTING

10 A Sequence Listing in XML text format, entitled 1499-73.xml, 556,402 bytes in size, generated on September 7, 2022 and filed herewith, is hereby incorporated by reference into the specification for its disclosures.

FIELD OF THE INVENTION

15 This invention relates to compositions and methods for modifying *SHATTERPROOF MADS-BOX (SHP)* genes in canola plants, optionally to reduce pod shattering. The invention further relates to canola plants having reduced pod shatter produced using the methods and compositions of the invention.

20 BACKGROUND OF THE INVENTION

Canola oil is a vegetable oil derived from a variety of rapeseed that is low in erucic acid, as opposed to colza oil. Both edible and industrial forms of oil are produced from the seed of any of several cultivars of the plant family Brassicaceae, namely cultivars of *Brassica napus L.*, *Brassica rapa subsp. Oleifera* (syn. *B. campestris L.*) and/or *Brassica juncea*, which are also referred to as "canola". Canola seed pods contain the black seeds that are pressed to extract the oil.

25 The canola seedpod is a fruit that encloses and protects the seeds as they are maturing, then dries and opens to disperse the seeds at maturity. There are three major regions in the canola fruit: the valves, the replum, and the valve margins. The valves are the seedpod walls that encircle the developing seeds and connect to the replum, which forms the middle ridge that attaches the fruit to the plant. The valve margins (VM) form at the boundary between the valves and the replum and are specialized for seed dispersal. When the fruit matures and dries, the valves detach from the replum along the margins in a process called dehiscence or pod shatter.

30

Three specialized cell types contribute to the opening of the fruit: the two layers of the margin—the separation layer (or dehiscence zone) and the lignified margin layer—as well as the lignified valve layer (endocarp b). The valves detach through cell-cell separation within the dehiscence zone that occurs following the secretion of hydrolytic enzymes. Lignification of the lignified margin layer and the internal lignified valve layer leads to stress that contributes mechanically to fruit opening. As the fruit dries, differential shrinkage of the remaining thin-walled valve cells relative to the rigid lignified margin and valve layers is thought to create internal tension, causing the shattering that is characteristic of fruit dehiscence.

The timing of dehiscence affects seed yields from canola. The fruit of canola develop throughout the growing season. The earliest fruit are fully developed while later fruit are still developing. The highly differentiated cells in the valve margins weaken the strength of the fruit, leading to seed dispersal at maturity. However, pod shattering is a highly undesirable trait for commercial seed production in canola and can cause significant yield losses of up to 70% in canola. Generally, canola is ‘windrowed’ to reduce seed loss due to shattering but this practice is not completely effective and is labor intensive. Seed losses accelerate further in the presence of high wind velocity and extremely high temperatures during the time of harvesting.

The aim of the present invention is to develop canola varieties having resistance to pod shattering so the standing crop can be directly harvested with combines without significant seed loss by providing new compositions and methods for reducing pod shattering in canola.

SUMMARY OF THE INVENTION

One aspect of the invention provides a canola plant or plant part thereof comprising at least one mutation in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene encoding a Shatterproof MADS-box transcription factor (SHP) polypeptide, optionally wherein the at least one mutation may be a non-natural mutation.

A second aspect of the invention provides a canola plant cell, comprising an editing system the editing system comprising: (a) a CRISPR-Cas effector protein; and (b) a guide nucleic acid comprising a spacer sequence with complementarity to an endogenous target gene encoding a Shatterproof MADS-box transcription factor (SHP) polypeptide in the canola plant cell.

A third aspect of the invention provides a canola plant cell comprising at least one mutation within an endogenous *SHP* gene, wherein the at least one non-natural mutation is a base substitution, a base insertion or a base deletion that is introduced using an editing system

that comprises a nucleic acid binding domain that binds to a target site in the endogenous *SHP* gene, optionally wherein the at least one mutation may be a non-natural mutation.

A fourth aspect of the invention provides a method of producing/breeding a transgene-free edited canola plant, comprising: crossing a canola plant of the invention with a transgene-free plant, thereby introducing the at least one mutation into the canola plant that is transgene-free; and selecting a progeny canola plant that comprises the at least one mutation and is transgene-free, thereby producing a transgene free edited canola plant, optionally wherein the at least one mutation may be a non-natural mutation.

A fifth aspect of the invention provides a method of providing a plurality of canola plants having one or more improved yield traits, the method comprising planting two or more canola plants of the invention in a growing area, thereby providing the plurality of canola plants having a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a plurality of control canola plants devoid of the at least one mutation, optionally wherein the at least one mutation may be a non-natural mutation.

In a sixth aspect, a method of creating a mutation in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene in a canola plant is provided, the method comprising: (a) targeting a gene editing system to a portion of the endogenous *SHP* gene that (i) comprises a sequence having at least 90% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; and/or (ii) encodes a sequence having at least 80% identity to any one **NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, and (b) selecting a canola plant that comprises a modification located in a region of the endogenous *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally located in a region of the endogenous *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**.

In a seventh aspect, a method is provided for generating variation in a Shatterproof MADS-box transcription factor (*SHP*) polypeptide in a canola plant cell, comprising: introducing an editing system into a canola plant cell, wherein the editing system is targeted to a region of *SHATTERPROOF MADS-BOX (SHP)* gene; and contacting the region of the *SHP* gene with the editing system, thereby introducing a mutation into the *SHP* gene and generating variation in the *SHP* polypeptide in the canola plant cell.

A eighth aspect provides a method for editing a specific site in the genome of a canola plant cell, the method comprising: cleaving, in a site-specific manner, a target site within an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene in the canola plant cell, the endogenous *SHP* gene: (a) comprising a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, (b) comprising a region having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, (c) encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, (d) encoding a region having at least 80% sequence identity to an amino acid sequence of any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby generating an edit in the endogenous *SHP* gene of the canola plant cell and producing a canola plant cell comprising the edit in the endogenous *SHP* gene.

A ninth aspect provides a method for making a canola plant, the method comprising (a) contacting a population of canola plant cells comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene with a nuclease linked to a nucleic acid binding domain (e.g., editing system) that binds to a sequence (i) having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, (ii) comprising a region having at least 80% identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, and/or (iv) encoding a region having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**; (b) selecting a canola plant cell from the population of canola plant cells in which an endogenous *SHP* gene has been mutated, thereby producing a canola plant cell comprising a mutation in the endogenous *SHP* gene; and (c) growing the selected canola plant cell into a canola plant comprising a mutation in the endogenous *SHP* gene.

A tenth aspect provides a method for reducing pod shattering and/or reducing lignification (reduced lignin content) in the pod valve margin in a canola plant, comprising (a) contacting a canola plant cell comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene with a nuclease targeting the endogenous *SHP* gene, wherein the nuclease is linked to a nucleic acid binding domain (e.g., editing system) that binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene: (i) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178,**

206, 207, 240 or 241; (ii) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (iv) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291** to produce a canola plant cell comprising a mutation in the endogenous *SHP* gene; and (b) growing the canola plant cell comprising a mutation in the endogenous *SHP* gene into a canola plant comprising the mutation in the endogenous *SHP* gene thereby producing a canola plant having a mutated endogenous *SHP* gene and reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

A eleventh aspect provides a method of producing a canola plant or part thereof comprising at least one cell having a mutated endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing the canola plant or part thereof comprising at least one cell having a mutation in the endogenous *SHP* gene.

A twelfth aspect of the invention provides a method for producing a canola plant or part thereof comprising a mutated endogenous *SHATTERPROOF MADS-BOX (SHP)* gene and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to

a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing the canola plant or part thereof comprising an endogenous *SHP* gene having a mutation and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

An thirteenth aspect provides a guide nucleic acid that binds to a target site in a *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the target site is in a region of the *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally a region of the *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**.

In a fourteenth aspect, a system is provided that comprises a guide nucleic acid of the invention and a CRISPR-Cas effector protein that associates with the guide nucleic acid.

A fifteenth aspect provides a gene editing system comprising a CRISPR-Cas effector protein in association with a guide nucleic acid, wherein the guide nucleic acid comprises a spacer sequence that binds to an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene.

In a sixteenth aspect, a complex comprising a guide nucleic acid and a CRISPR-Cas effector protein comprising a cleavage domain is provided, wherein the guide nucleic acid binds to a target site in endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, and the cleavage domain cleaves a target strand in the *SHP* gene.

In a seventeenth aspect, an expression cassette is provided, the expression cassette comprising (a) a polynucleotide encoding CRISPR-Cas effector protein comprising a cleavage

domain and (b) a guide nucleic acid that binds to a target site in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the guide nucleic acid comprises a spacer sequence that is complementary to and binds to (i) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (ii) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) a portion of a nucleic acid encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (iv) a portion of a nucleic acid encoding an amino acid sequence having at least 80% identity to any one **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

In another aspect, plants are provided that comprise in their genome one or more mutated *SHATTERPROOF MADS-BOX (SHP)* genes produced by the methods of the invention, optionally wherein the at least one mutation may be a non-natural mutation, optionally wherein the mutated *SHP* gene comprises a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323** and/or encodes a SHP polypeptide having at least 90% sequence identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317**.

A further aspect of the invention provides a canola plant or plant part thereof comprising at least one mutation in at least one endogenous *SHATTERPROOF MADS-BOX (SHP)* gene having a gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*), optionally wherein the at least one mutation may be a non-natural mutation.

In a further aspect, a guide nucleic acid is provided that binds to target nucleic acid in a *SHATTERPROOF MADS-BOX (SHP)* gene having a gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*). Further provided are polypeptides, polynucleotides, nucleic acid constructs, expression cassettes and vectors for making a canola plant of this invention.

These and other aspects of the invention are set forth in more detail in the description of the invention below.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NOs:1-17 are exemplary Cas12a amino acid sequences useful with this invention.

SEQ ID NOs:18-20 are exemplary Cas12a nucleotide sequences useful with this invention.

SEQ ID NO:21-22 are exemplary regulatory sequences encoding a promoter and intron.

SEQ ID NOs:23-29 are exemplary cytosine deaminase sequences useful with this invention.

SEQ ID NOs:30-40 are exemplary adenine deaminase amino acid sequences useful with this invention.

SEQ ID NO:41 is an exemplary uracil-DNA glycosylase inhibitor (UGI) sequences useful with this invention.

SEQ ID NOs:42-44 provide example peptide tags and affinity polypeptides useful with this invention.

SEQ ID NOs:45-55 provide example RNA recruiting motifs and corresponding affinity polypeptides useful with this invention.

SEQ ID NOs:56-57 are exemplary Cas9 polypeptide sequences useful with this invention.

SEQ ID NOs:58-68 are exemplary Cas9 polynucleotide sequences useful with this invention.

SEQ ID NO:69 (BnaA04g01810D) is an example *SHP3* (*SHP3A*) genomic sequence from canola.

SEQ ID NO:70 is an example *SHP3* (*SHP3A*) coding sequence from canola.

SEQ ID NO:71 is an example *SHP3*(*SHP3A*) polypeptide sequence from canola.

SEQ ID NOs:72-96 are example portions or regions of *SHP3* genomic and coding sequences from canola (BnaA04g01810D, *SHP3A*).

SEQ ID NOs:97-99 are an example portions or regions of an *SHP3* polypeptide from canola (*SHP3A*).

SEQ ID NO:100 (BnaA07g18050D) is an example *SHP2* (*SHP2A*) genomic sequence from canola.

SEQ ID NO:101 is an example *SHP2* (*SHP2A*) coding sequence from canola.

SEQ ID NO:102 is an example *SHP2* (*SHP2A*) polypeptide sequence from canola.

SEQ ID NOs:103-144 are example portions or regions of *SHP2* genomic and coding sequences from canola (BnaA07g18050D, *SHP2A*).

SEQ ID NOs:145-147 are an example portions or regions of an *SHP2* polypeptide from canola (*SHP2A*).

5 **SEQ ID NO:148** (BnaA05g02990D) is an example *SHP4* (*SHP4A*) genomic sequence from canola.

SEQ ID NO:149 is an example *SHP4*(*SHP4A*) coding sequence from canola.

SEQ ID NO:150 is an example *SHP4* (*SHP4A*) polypeptide sequence from canola.

10 **SEQ ID NOs:151-173** and **324-338** are example portions or regions of *SHP4* genomic and coding sequences from canola (BnaA05g02990D, *SHP4A*).

SEQ ID NOs:174-176 are an example portions or regions of an *SHP4* polypeptide from canola (*SHP4A*).

SEQ ID NO:177 (BnaA09g55330D) is an example *SHP1* (*SHP1A*) genomic sequence from canola.

15 **SEQ ID NO:178** is an example *SHP1* (*SHP1A*) coding sequence from canola.

SEQ ID NO:179 is an example *SHP1* polypeptide sequence from canola (*SHP1A*).

SEQ ID NOs:180-202 are example portions or regions of a canola *SHP1* genomic and coding sequences (BnaA09g55330D, *SHP1A*).

20 **SEQ ID NOs:203-205** are an example portions or regions of an *SHP1* polypeptide from canola (*SHP1A*).

SEQ ID NO:206 (BnaC04g23360D) is an example *SHP3* (*SHP3C*) genomic sequence from canola.

SEQ ID NO:207 is an example *SHP3* (*SHP3C*) coding sequence from canola.

SEQ ID NO:208 is an example *SHP3* (*SHP3C*) polypeptide sequence from canola.

25 **SEQ ID NOs:209-236** are example portions or regions of *SHP3* genomic and coding sequences from canola (BnaC04g23360D, *SHP3C*).

SEQ ID NOs:237-239 are an example portions or regions of an *SHP3* polypeptide from canola (*SHP3C*).

30 **SEQ ID NO:240** (BnaC06g16910D) is an example *SHP2* (*SHP2C*) genomic sequence from canola.

SEQ ID NO:241 is an example *SHP2* (*SHP2C*) coding sequence from canola.

SEQ ID NO:242 is an example *SHP2* (*SHP2C*) polypeptide sequence from canola.

SEQ ID NOs:243-288 are example portions or regions of *SHP2* genomic and coding sequences from canola (BnaC06g16910D, *SHP2C*).

SEQ ID NOs:289-291 are an example portions or regions of an SHP2 polypeptide from canola (SHP2C).

5 **SEQ ID NOs:292-297** and **342-346** are example spacer sequences for nucleic acid guides useful with this invention.

SEQ ID NOs:298, 300, 302, 306, 312, 314, and 321 are example mutated *SHP2* (*SHP2A*) genomic sequences (BnaA07g18050D, **SEQ ID NO:100**) produced using the methods of the invention.

10 **SEQ ID NOs:299, 301, 303, 307, 313 and 315** are example mutated SHP2 (SHP2A) polypeptide sequences encoded by the mutated *SHP2* genomic sequences of **SEQ ID NOs:298, 300, 302, 309, 312 and 314**, respectively.

SEQ ID NOs:304, 308, 310, 316, 318 and 319 are example mutated *SHP2* (*SHP2C*) genomic sequences (BnaC06g16910D, **SEQ ID NO:240**) produced using the methods of the invention.

SEQ ID NOs:305, 309, 311, and 317 are example mutated SHP2 (SHP2C) polypeptide sequences encoded by the mutated *SHP2* genomic sequences of **SEQ ID NOs:304, 308, 310 and 316**, respectively.

20 **SEQ ID NO:320** is an example deleted portion from an SHP2C gene ((BnaC06g16910D, **SEQ ID NO:240**).

SEQ ID NO:322 and **SEQ ID NO:323** are example mutated *SHP4* (SHP4A) genomic sequences (BnaA05g02990D, **SEQ ID NO:148**) produced using the methods of the invention.

DETAILED DESCRIPTION

25 The present invention now will be described hereinafter with reference to the accompanying drawings and examples, in which embodiments of the invention are shown. This description is not intended to be a detailed catalog of all the different ways in which the invention may be implemented, or all the features that may be added to the instant invention. For example, features illustrated with respect to one embodiment may be incorporated into other
30 embodiments, and features illustrated with respect to a particular embodiment may be deleted from that embodiment. Thus, the invention contemplates that in some embodiments of the invention, any feature or combination of features set forth herein can be excluded or omitted. In addition, numerous variations and additions to the various embodiments suggested herein will be

apparent to those skilled in the art in light of the instant disclosure, which do not depart from the instant invention. Hence, the following descriptions are intended to illustrate some particular embodiments of the invention, and not to exhaustively specify all permutations, combinations and variations thereof.

5 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 belongs. The terminology used in the description of the invention herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention.

10 All publications, patent applications, patents and other references cited herein are incorporated by reference in their entireties for the teachings relevant to the sentence and/or paragraph in which the reference is presented.

Unless the context indicates otherwise, it is specifically intended that the various features of the invention described herein can be used in any combination. Moreover, the present invention also contemplates that in some embodiments of the invention, any feature or
15 combination of features set forth herein can be excluded or omitted. To illustrate, if the specification states that a composition comprises components A, B and C, it is specifically intended that any of A, B or C, or a combination thereof, can be omitted and disclaimed singularly or in any combination.

As used in the description of the invention and the appended claims, the singular forms
20 "a," "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise.

Also as used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative ("or").

25 The term "about," as used herein when referring to a measurable value such as an amount or concentration and the like, is meant to encompass variations of $\pm 10\%$, $\pm 5\%$, $\pm 1\%$, $\pm 0.5\%$, or even $\pm 0.1\%$ of the specified value as well as the specified value. For example, "about X" where X is the measurable value, is meant to include X as well as variations of $\pm 10\%$, $\pm 5\%$, $\pm 1\%$, $\pm 0.5\%$, or even $\pm 0.1\%$ of X. A range provided herein for a measurable value may include any
30 other range and/or individual value therein.

As used herein, phrases such as "between X and Y" and "between about X and Y" should be interpreted to include X and Y. As used herein, phrases such as "between about X and Y"

mean "between about X and about Y" and phrases such as "from about X to Y" mean "from about X to about Y."

Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. For example, if the range 10 to 15 is disclosed, then 11, 12, 13, and 14 are also disclosed.

The term "comprise," "comprises" and "comprising" as used herein, specify the presence of the stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

As used herein, the transitional phrase "consisting essentially of" means that the scope of a claim is to be interpreted to encompass the specified materials or steps recited in the claim and those that do not materially affect the basic and novel characteristic(s) of the claimed invention. Thus, the term "consisting essentially of" when used in a claim of this invention is not intended to be interpreted to be equivalent to "comprising."

As used herein, the terms "increase," "increasing," "increased," "enhance," "enhanced," "enhancing," and "enhancement" (and grammatical variations thereof) describe an elevation of at least about 5%, 10%, 15%, 20%, 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400%, 500% or more as compared to a control. For example, a canola plant comprising a mutation in a *SHATTERPROOF MADS-BOX (SHP)* gene as described herein can exhibit an increase in harvestable seed, wherein the increase in harvestable seed is an increase of at least 10% over that produced by a control plant (e.g., an increase of at least about 10% to about 70% in harvestable seed, e.g., about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70%, or any range or value therein).

As used herein, the terms "reduce," "reduced," "reducing," "reduction," "diminish," and "decrease" (and grammatical variations thereof), describe, for example, a decrease of at least about 5%, 10%, 15%, 20%, 25%, 35%, 50%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% as compared to a control. In particular embodiments, the reduction can result in no or essentially no (*i.e.*, an insignificant amount, *e.g.*, less than about 10% or even 5%) detectable activity or amount. As an example, a canola plant comprising a mutation in a *SHATTERPROOF MADS-BOX (SHP)* gene as described herein can

exhibit a reduction in pod shattering of at least 10% when compared to a control canola plant devoid of the at least one mutation (e.g., a reduction in pod shattering of at least about 10% to about 100%, e.g., about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%, or any range or value therein).

A control canola plant is typically the same plant as the edited plant, but the control plant has not been similarly edited and therefore is devoid of the mutation. A control plant maybe an isogenic plant and/or a wild type plant. Thus, a control plant can be the same breeding line, variety, or cultivar as the subject plant into which a mutation as described herein is introgressed, but the control breeding line, variety, or cultivar is free of the mutation. In some embodiments, a comparison between a canola plant of the invention and a control canola plant is made under the same growth conditions, e.g., the same environmental conditions (soil, hydration, light, heat, nutrients, and the like).

As used herein, the terms "express," "expresses," "expressed" or "expression," and the like, with respect to a nucleic acid molecule and/or a nucleotide sequence (e.g., RNA or DNA) indicates that the nucleic acid molecule and/or a nucleotide sequence is transcribed and, optionally, translated. Thus, a nucleic acid molecule and/or a nucleotide sequence may express a polypeptide of interest or, for example, a functional untranslated RNA.

A "heterologous" or a "recombinant" nucleotide sequence is a nucleotide sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of a naturally occurring nucleotide sequence. A "heterologous" nucleotide/polypeptide may originate from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention.

A "native" or "wild type" nucleic acid, nucleotide sequence, polypeptide or amino acid sequence refers to a naturally occurring or endogenous nucleic acid, nucleotide sequence, polypeptide or amino acid sequence. In some contexts, a "wild type" nucleic acid is a nucleic acid that is not edited as described herein and can differ from an "endogenous" gene that may be edited as described herein (e.g., a mutated endogenous gene). In some contexts, a "wild type" nucleic acid (e.g., unedited) may be heterologous to the organism in which the wild type nucleic acid is found (e.g., a transgenic organism). As an example, a "wild type endogenous

SHATTERPROOF MADS-BOX (SHP) gene" is a *SHP* gene that is naturally occurring in or endogenous to the reference organism, e.g., a canola plant, and may be subject to modification as described herein, after which, such a modified endogenous gene is no longer wild type. In some embodiments, an endogenous *SHP* gene may be an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally wherein the endogenous *SHP* gene has a gene identification number (gene ID) of BnaA04g01810D (*SHP3A*), BnaA07g18050D (*SHP2A*), BnaA05g02990D (*SHP4A*), BnaA09g55330D (*SHP1A*), BnaC04g23360D (*SHP3C*), and/or BnaC06g16910D (*SHP2C*) (BrassicaEDB – a Gene Expression Database for *Brassica* Crops (brassica.biodb.org/analysis) or plants.ensembl.org).

As used herein, the term "heterozygous" refers to a genetic status wherein different alleles reside at corresponding loci on homologous chromosomes.

As used herein, the term "homozygous" refers to a genetic status wherein identical alleles reside at corresponding loci on homologous chromosomes.

As used herein, the term "allele" refers to one of two or more different nucleotides or nucleotide sequences that occur at a specific locus.

A "null allele" is a nonfunctional allele caused by a genetic mutation that results in a complete lack of production of the corresponding protein or produces a protein that is non-functional.

A "recessive mutation" is a mutation in a gene that produces a phenotype when homozygous but the phenotype is not observable when the locus is heterozygous.

A "dominant mutation" is a mutation in a gene that produces a mutant phenotype in the presence of a non-mutated copy of the gene. A dominant mutation may be a loss or a gain of function mutation, a hypomorphic mutation, a hypermorphic mutation or a weak loss of function or a weak gain of function.

A "dominant negative mutation" is a mutation that produces an altered gene product (e.g., having an aberrant function relative to wild type), which gene product adversely affects the function of the wild-type allele or gene product. For example, a "dominant negative mutation" may block a function of the wild type gene product. A dominant negative mutation may also be referred to as an "antimorphic mutation."

A "semi-dominant mutation" refers to a mutation in which the penetrance of the phenotype in a heterozygous organism is less than that observed for a homozygous organism.

A "weak loss-of-function mutation" is a mutation that results in a gene product having partial function or reduced function (partially inactivated) as compared to the wildtype gene product.

A "hypomorphic mutation" is a mutation that results in a partial loss of gene function, which may occur through reduced expression (e.g., reduced protein and/or reduced RNA) or reduced functional performance (e.g., reduced activity), but not a complete loss of function/activity. A "hypomorphic" allele is a semi-functional allele caused by a genetic mutation that results in production of the corresponding protein that functions at anywhere between 1% and 99% of normal efficiency.

A "hypermorphic mutation" is a mutation that results in increased expression of the gene product and/or increased activity of the gene product.

A "locus" is a position on a chromosome where a gene or marker or allele is located. In some embodiments, a locus may encompass one or more nucleotides.

As used herein, the terms "desired allele," "target allele" and/or "allele of interest" are used interchangeably to refer to an allele associated with a desired trait. In some embodiments, a desired allele may be associated with either an increase or a decrease (relative to a control) of or in a given trait, depending on the nature of the desired phenotype.

A marker is "associated with" a trait when said trait is linked to it and when the presence of the marker is an indicator of whether and/or to what extent the desired trait or trait form will occur in a plant/germplasm comprising the marker. Similarly, a marker is "associated with" an allele or chromosome interval when it is linked to it and when the presence of the marker is an indicator of whether the allele or chromosome interval is present in a plant/germplasm comprising the marker.

As used herein, the terms "backcross" and "backcrossing" refer to the process whereby a progeny plant is crossed back to one of its parents one or more times (e.g., 1, 2, 3, 4, 5, 6, 7, 8, etc.). In a backcrossing scheme, the "donor" parent refers to the parental plant with the desired gene or locus to be introgressed. The "recipient" parent (used one or more times) or "recurrent" parent (used two or more times) refers to the parental plant into which the gene or locus is being introgressed. For example, see Ragot, M. et al. *Marker-assisted Backcrossing: A Practical Example*, in TECHNIQUES ET UTILISATIONS DES MARQUEURS MOLECULAIRES LES COLLOQUES, Vol. 72, pp. 45-56 (1995); and Openshaw et al., *Marker-assisted Selection in Backcross Breeding*, in PROCEEDINGS OF THE SYMPOSIUM "ANALYSIS OF MOLECULAR MARKER DATA," pp. 41-43 (1994). The initial cross gives rise to the F1 generation. The term "BC1" refers to the

second use of the recurrent parent, "BC2" refers to the third use of the recurrent parent, and so on.

As used herein, the terms "cross" or "crossed" refer to the fusion of gametes via pollination to produce progeny (e.g., cells, seeds or plants). The term encompasses both sexual crosses (the pollination of one plant by another) and selfing (self-pollination, e.g., when the pollen and ovule are from the same plant). The term "crossing" refers to the act of fusing gametes via pollination to produce progeny.

As used herein, the terms "introgression," "introgressing" and "introgressed" refer to both the natural and artificial transmission of a desired allele or combination of desired alleles of a genetic locus or genetic loci from one genetic background to another. For example, a desired allele at a specified locus can be transmitted to at least one (e.g., one or more) progeny via a sexual cross between two parents of the same species, where at least one of the parents has the desired allele in its genome. Alternatively, for example, transmission of an allele can occur by recombination between two donor genomes, e.g., in a fused protoplast, where at least one of the donor protoplasts has the desired allele in its genome. The desired allele may be a selected allele of a marker, a QTL, a transgene, or the like. Offspring comprising the desired allele can be backcrossed one or more times (e.g., 1, 2, 3, 4, or more times) to a line having a desired genetic background, selecting for the desired allele, with the result being that the desired allele becomes fixed in the desired genetic background. For example, a marker associated with increased yield under non-water stress conditions may be introgressed from a donor into a recurrent parent that does not comprise the marker and does not exhibit increased yield under non-water stress conditions. The resulting offspring could then be backcrossed one or more times and selected until the progeny possess the genetic marker(s) associated with increased yield under non-water stress conditions in the recurrent parent background.

A "genetic map" is a description of genetic linkage relationships among loci on one or more chromosomes within a given species, generally depicted in a diagrammatic or tabular form. For each genetic map, distances between loci are measured by the recombination frequencies between them. Recombination between loci can be detected using a variety of markers. A genetic map is a product of the mapping population, types of markers used, and the polymorphic potential of each marker between different populations. The order and genetic distances between loci can differ from one genetic map to another.

As used herein, the term "genotype" refers to the genetic constitution of an individual (or group of individuals) at one or more genetic loci, as contrasted with the observable and/or

detectable and/or manifested trait (the phenotype). Genotype is defined by the allele(s) of one or more known loci that the individual has inherited from its parents. The term genotype can be used to refer to an individual's genetic constitution at a single locus, at multiple loci, or more generally, the term genotype can be used to refer to an individual's genetic make-up for all the genes in its genome. Genotypes can be indirectly characterized, e.g., using markers and/or directly characterized by nucleic acid sequencing.

As used herein, the term "germplasm" refers to genetic material of or from an individual (e.g., a plant), a group of individuals (e.g., a plant line, variety or family), or a clone derived from a line, variety, species, or culture. The germplasm can be part of an organism or cell or can be separate from the organism or cell. In general, germplasm provides genetic material with a specific genetic makeup that provides a foundation for some or all of the hereditary qualities of an organism or cell culture. As used herein, germplasm includes cells, seed or tissues from which new plants may be grown, as well as plant parts that can be cultured into a whole plant (e.g., leaves, stems, buds, roots, pollen, cells, etc.).

As used herein, the terms "cultivar" and "variety" refer to a group of similar plants that by structural or genetic features and/or performance can be distinguished from other varieties within the same species.

As used herein, the terms "exotic," "exotic line" and "exotic germplasm" refer to any plant, line or germplasm that is not elite. In general, exotic plants/germplasms are not derived from any known elite plant or germplasm, but rather are selected to introduce one or more desired genetic elements into a breeding program (e.g., to introduce novel alleles into a breeding program).

As used herein, the term "hybrid" in the context of plant breeding refers to a plant that is the offspring of genetically dissimilar parents produced by crossing plants of different lines or breeds or species, including but not limited to the cross between two inbred lines.

As used herein, the term "inbred" refers to a substantially homozygous plant or variety. The term may refer to a plant or plant variety that is substantially homozygous throughout the entire genome or that is substantially homozygous with respect to a portion of the genome that is of particular interest.

A "haplotype" is the genotype of an individual at a plurality of genetic loci, i.e., a combination of alleles. Typically, the genetic loci that define a haplotype are physically and genetically linked, i.e., on the same chromosome segment. The term "haplotype" can refer to

polymorphisms at a particular locus, such as a single marker locus, or polymorphisms at multiple loci along a chromosomal segment.

A canola plant in which at least one (e.g., one or more, e.g., 1, 2, 3, or 4, or more) endogenous *SHP* gene (e.g., an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene) is modified as described herein (e.g., comprises a modification as described herein) may have reduced pod shattering as compared to a canola plant that does not comprise (is devoid of) the modification in the at least one endogenous *SHP* gene. In some embodiments, a canola plant in which at least one endogenous *SHP* gene is modified as described herein may exhibit reduced lignification (reduced lignin content) in the pod valve margin of the canola plant comprising the at least one endogenous *SHP* gene modified as described herein. In some embodiments, a canola plant in which at least one endogenous *SHP* gene is modified as described herein may exhibit an increase in harvestable seed as compared to a canola plant that does not comprise (is devoid of) the modification in the at least one endogenous *SHP* gene.

As used herein, "reduced pod shattering" means a reduction in pod shattering of at least 10% when compared to a control canola plant devoid of the at least one mutation (e.g., a reduction in pod shattering of at least about 10% to about 100%, e.g., a reduction of about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%, or any range or value therein). Reduced pod shattering may result in an increase in harvestable seed.

As used herein, "reduced lignification in the pod valve margin" means a reduction in detectable lignin content by at least 10% at the pod valve margin when compared to pod valve margins in a control canola plant devoid of the at least one mutation (e.g., a reduction in lignification in the pod valve margin of at least about 10% to about 100%, e.g., a reduction of about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%, or any range or value therein), wherein a reduction of 100% means no detectable lignin (e.g., no lignin staining in the pod valve margin). Reduced lignification in the pod valve margin can result in reduced pod shatter.

As used herein, an "increase in harvestable seed" means an increase in harvestable seed of at least 10% when compared to a control canola plant devoid of the at least one mutation (e.g., an increase of at least about 10% to about 70% in harvestable seed, e.g., about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70%, or any range or value therein).

As used herein a "control plant" means a canola plant that does not contain an edited *SHP* gene or gene as described herein that imparts an altered phenotype of reduced pod shattering and/or increased harvestable seed and/or reduced lignification (reduced lignin content) in the pod valve margin. A control canola plant is used to identify and select a canola plant edited as described herein and that has an enhanced trait or altered phenotype as compared to the control canola plant. A suitable control plant can be a plant of the parental line used to generate a plant comprising a mutated *SHP* gene(s), for example, a wild type plant devoid of an edit in an endogenous *SHP* gene as described herein. A suitable control plant can also be a plant that contains recombinant nucleic acids that impart other traits, for example, a transgenic plant having enhanced herbicide tolerance. A suitable canola control plant can in some cases be a progeny of a heterozygous or hemizygous transgenic canola plant line that is devoid of the mutated *SHP* gene as described herein, known as a negative segregant, or a negative isogenic line.

An enhanced trait (e.g., improved yield trait) may include, for example, decreased days from planting to maturity, increased stalk size, increased number of leaves, increased plant height growth rate in vegetative stage, increased ear size, increased ear dry weight per plant, increased number of kernels per ear, increased weight per kernel, increased number of kernels per plant, decreased ear void, extended grain fill period, reduced plant height, increased number of root branches, increased total root length, increased yield (e.g., increase in harvestable seed), increased nitrogen use efficiency, and/or increased water use efficiency as compared to a control plant. An altered phenotype may be, for example, plant height, biomass, canopy area, anthocyanin content, chlorophyll content, water applied, water content, and water use efficiency.

In some embodiments, a plant of this invention may comprise one or more improved yield traits including, but not limited to, In some embodiments, the one or more improved yield traits includes higher yield (bu/acre), increased biomass, increased plant height, increased stem diameter, increased leaf area, increased number of flowers, increased kernel row number, optionally wherein ear length is not substantially reduced, increased kernel number, increased

kernel size, increased ear length, decreased tiller number, decreased tassel branch number, increased number of pods, including an increased number of pods per node and/or an increased number of pods per plant, increased number of seeds per pod, increased number of seeds, increased seed size, and/or increased seed weight (e.g., increase in 100-seed weight) as compared to a control plant devoid of the at least one mutation. In some embodiments, a plant of this invention may comprise one or more improved yield traits including, but not limited to, optionally an increase in yield (bu/acre), seed size (including kernel size), seed weight (including kernel weight), increased kernel row number (optionally wherein ear length is not substantially reduced), increased number of pods, increased number of seeds per pod and an increase in ear length as compared to a control plant or part thereof.

As used herein a "trait" is a physiological, morphological, biochemical, or physical characteristic of a plant or particular plant material or cell. In some instances, this characteristic is visible to the human eye and can be measured mechanically, such as seed or plant size, weight, shape, form, length, height, growth rate and development stage, or can be measured by biochemical techniques, such as detecting the protein, starch, certain metabolites, or oil content of seed or leaves, or by observation of a metabolic or physiological process, for example, by measuring tolerance to water deprivation or particular salt or sugar concentrations, or by the measurement of the expression level of a gene or genes, for example, by employing Northern analysis, RT-PCR, microarray gene expression assays, or reporter gene expression systems, or by agricultural observations such as hyperosmotic stress tolerance or yield. However, any technique can be used to measure the amount of, the comparative level of, or the difference in any selected chemical compound or macromolecule in the transgenic plants.

As used herein an "enhanced trait" means a characteristic of a canola plant resulting from mutations in a *SHP* gene(s) as described herein. Such traits include, but are not limited to, an enhanced agronomic trait characterized by enhanced plant morphology, physiology, growth and development, yield, nutritional enhancement, disease or pest resistance, or environmental or chemical tolerance. In some embodiments, an enhanced trait/altered phenotype may be, for example, decreased days from planting to maturity, increased stalk size, increased number of leaves, increased plant height growth rate in vegetative stage, increased ear size, increased ear dry weight per plant, increased number of kernels per ear, increased weight per kernel, increased number of kernels per plant, decreased ear void, extended grain fill period, reduced plant height, increased number of root branches, increased total root length, drought tolerance, increased water use efficiency, cold tolerance, increased nitrogen use efficiency, and/or increased yield. In

some embodiments, a trait is increased yield under nonstress conditions or increased yield under environmental stress conditions. Stress conditions can include both biotic and abiotic stress, for example, drought, shade, fungal disease, viral disease, bacterial disease, insect infestation, nematode infestation, cold temperature exposure, heat exposure, osmotic stress, reduced nitrogen nutrient availability, reduced phosphorus nutrient availability and high plant density. "Yield" can be affected by many properties including without limitation, plant height, plant biomass, pod number, pod position on the plant, number of internodes, incidence of pod shatter, grain size, ear size, ear tip filling, kernel abortion, efficiency of nodulation and nitrogen fixation, efficiency of nutrient assimilation, resistance to biotic and abiotic stress, carbon assimilation, plant architecture, resistance to lodging, percent seed germination, seedling vigor, and juvenile traits. Yield can also be affected by efficiency of germination (including germination in stressed conditions), growth rate (including growth rate in stressed conditions), flowering time and duration, ear number, ear size, ear weight, seed number per ear or pod, seed size, composition of seed (starch, oil, protein) and characteristics of seed fill.

Also used herein, the term "trait modification" encompasses altering the naturally occurring trait by producing a detectable difference in a characteristic in a canola plant comprising a mutation in an endogenous *SHP* gene as described herein relative to a canola plant not comprising the mutation, such as a wild-type plant, or a negative segregant. In some cases, the trait modification can be evaluated quantitatively. For example, the trait modification can entail an increase or decrease in an observed trait characteristic or phenotype as compared to a control plant. It is known that there can be natural variations in a modified trait. Therefore, the trait modification observed can entail a change of the normal distribution and magnitude of the trait characteristics or phenotype in the plants as compared to a control plant.

The present disclosure relates to a canola plant with improved economically relevant characteristics, more specifically reduced pod shattering, reduced lignification (reduced lignin content) in the pod valve margin, and/or increased harvestable seed yield. More specifically the present disclosure relates to a canola plant comprising a mutation(s) in an *SHP* gene(s) as described herein, wherein the canola plant has reduced pod shattering, reduced lignification (reduced lignin content) in the pod valve margin, and/or increased harvestable seed yield as compared to a control plant devoid of said mutation(s). In some embodiments, a canola plant of the present disclosure exhibits an improved trait that is related to yield, including but not limited to increased nitrogen use efficiency, increased nitrogen stress tolerance, increased water use efficiency and/or increased drought tolerance, as defined and discussed *infra*.

Yield can be defined as the measurable produce of economic value from a crop. Yield can be defined in the scope of quantity and/or quality. Yield can be directly dependent on several factors, for example, the number and size of organs (e.g., number of flowers), plant architecture (such as the number of branches, plant biomass, e.g., increased root biomass, steeper root angle and/or longer roots, and the like), flowering time and duration, grain fill period. Root architecture and development, photosynthetic efficiency, nutrient uptake, stress tolerance, early vigor, delayed senescence and functional stay green phenotypes may be factors in determining yield. Optimizing the above-mentioned factors can therefore contribute to increasing crop yield.

Reference herein to an increase/improvement in yield-related traits can also be taken to mean an increase in biomass (weight) of one or more parts of a plant, which can include above ground and/or below ground (harvestable) plant parts. In particular, such harvestable parts are seeds, and performance of the methods of the disclosure results in plants with increased yield and in particular increased seed yield relative to the seed yield of suitable control plants. In some embodiments, performance of the methods of the disclosure results in canola plants having increased harvestable seed (and thus increased seed yield), optionally reduced lignification (decreased lignin content) in the pod valve margin, relative to suitable control canola plants due to reduced seed loss resulting from reduced pod shattering. The term "yield" of a plant can relate to vegetative biomass (root and/or shoot biomass), to reproductive organs, and/or to propagules (such as seeds) of that plant.

Increased yield of a plant of the present disclosure can be measured in a number of ways, including test weight, seed number per plant, seed weight, seed number per unit area (for example, seeds, or weight of seeds, per acre), bushels per acre, tons per acre, or kilo per hectare. Increased yield can result from improved utilization of key biochemical compounds, such as nitrogen, phosphorous and carbohydrate, or from improved responses to environmental stresses, such as cold, heat, drought, salt, shade, high plant density, and attack by pests or pathogens.

"Increased yield" can manifest as one or more of the following: (i) increased plant biomass (weight) of one or more parts of a plant, particularly aboveground (harvestable) parts, of a plant, increased root biomass (increased number of roots, increased root thickness, increased root length) or increased biomass of any other harvestable part; or (ii) increased early vigor, defined herein as an improved seedling aboveground area approximately three weeks post-germination.

"Early vigor" refers to active healthy plant growth especially during early stages of plant growth, and can result from increased plant fitness due to, for example, the plants being better adapted to their environment (for example, optimizing the use of energy resources, uptake of nutrients and partitioning carbon allocation between shoot and root). Early vigor, for example, can be a combination of the ability of seeds to germinate and emerge after planting and the ability of the young plants to grow and develop after emergence. Plants having early vigor also show increased seedling survival and better establishment of the crop, which often results in highly uniform fields with the majority of the plants reaching the various stages of development at substantially the same time, which often results in increased yield. Therefore, early vigor can be determined by measuring various factors, such as kernel weight, percentage germination, percentage emergence, seedling growth, seedling height, root length, root and shoot biomass, canopy size and color and others.

Further, increased yield can also manifest as increased total seed yield, which may result from one or more of an increase in seed biomass (seed weight) due to an increase in the seed weight on a per plant and/or on an individual seed basis an increased number of, for example, flowers/panicles per plant; an increased number of pods; an increased number of nodes; an increased number of flowers ("florets") per panicle/plant; increased seed fill rate; an increased number of filled seeds; increased seed size (length, width, area, perimeter, and/or weight), which can also influence the composition of seeds; and/or increased seed volume, which can also influence the composition of seeds. In one embodiment, increased yield can be increased seed yield, for example, increased seed weight; increased number of filled seeds; and/or increased harvest index.

Increased yield can also result in modified architecture, or can occur because of modified plant architecture.

Increased yield can also manifest as increased harvest index, which is expressed as a ratio of the yield of harvestable parts, such as seeds, over the total biomass

The disclosure also extends to harvestable parts of a plant such as, but not limited to, seeds, leaves, fruits, flowers, bolls, pods, siliques, nuts, stems, rhizomes, tubers and bulbs. The disclosure furthermore relates to products derived from a harvestable part of such a plant, such as dry pellets, powders, oil, fat and fatty acids, starch or proteins.

The present disclosure provides a method for increasing "yield" of a plant or "broad acre yield" of a plant or plant part defined as the harvestable plant parts per unit area, for example

seeds, or weight of seeds, per acre, pounds per acre, bushels per acre, tones per acre, tons per acre, kilo per hectare.

As used herein "nitrogen use efficiency" refers to the processes which lead to an increase in the plant's yield, biomass, vigor, and growth rate per nitrogen unit applied. The processes can include the uptake, assimilation, accumulation, signaling, sensing, retranslocation (within the plant) and use of nitrogen by the plant.

As used herein "increased nitrogen use efficiency" refers to the ability of plants to grow, develop, or yield faster or better than normal when subjected to the same amount of available/applied nitrogen as under normal or standard conditions; ability of plants to grow, develop, or yield normally, or grow, develop, or yield faster or better when subjected to less than optimal amounts of available/applied nitrogen, or under nitrogen limiting conditions.

As used herein "nitrogen limiting conditions" refers to growth conditions or environments that provide less than optimal amounts of nitrogen needed for adequate or successful plant metabolism, growth, reproductive success and/or viability.

As used herein the "increased nitrogen stress tolerance" refers to the ability of plants to grow, develop, or yield normally, or grow, develop, or yield faster or better when subjected to less than optimal amounts of available/applied nitrogen, or under nitrogen limiting conditions.

Increased plant nitrogen use efficiency can be translated in the field into either harvesting similar quantities of yield, while supplying less nitrogen, or increased yield gained by supplying optimal/sufficient amounts of nitrogen. The increased nitrogen use efficiency can improve plant nitrogen stress tolerance and can also improve crop quality and biochemical constituents of the seed such as protein yield and oil yield. The terms "increased nitrogen use efficiency", "enhanced nitrogen use efficiency", and "nitrogen stress tolerance" are used inter-changeably in the present disclosure to refer to plants with improved productivity under nitrogen limiting conditions.

As used herein "water use efficiency" refers to the amount of carbon dioxide assimilated by leaves per unit of water vapor transpired. It constitutes one of the most important traits controlling plant productivity in dry environments. "Drought tolerance" refers to the degree to which a plant is adapted to arid or drought conditions. The physiological responses of plants to a deficit of water include leaf wilting, a reduction in leaf area, leaf abscission, and the stimulation of root growth by directing nutrients to the underground parts of the plants. Typically, plants are more susceptible to drought during flowering and seed development (the reproductive stages), as plant's resources are deviated to support root growth. In addition, abscisic acid (ABA), a plant

stress hormone, induces the closure of leaf stomata (microscopic pores involved in gas exchange), thereby reducing water loss through transpiration, and decreasing the rate of photosynthesis. These responses improve the water-use efficiency of the plant on the short term. The terms "increased water use efficiency", "enhanced water use efficiency", and "increased drought tolerance" are used inter-changeably in the present disclosure to refer to plants with improved productivity under water-limiting conditions.

As used herein "increased water use efficiency" refers to the ability of plants to grow, develop, or yield faster or better than normal when subjected to the same amount of available/applied water as under normal or standard conditions; ability of plants to grow, develop, or yield normally, or grow, develop, or yield faster or better when subjected to reduced amounts of available/applied water (water input) or under conditions of water stress or water deficit stress.

As used herein "increased drought tolerance" refers to the ability of plants to grow, develop, or yield normally, or grow, develop, or yield faster or better than normal when subjected to reduced amounts of available/applied water and/or under conditions of acute or chronic drought; ability of plants to grow, develop, or yield normally when subjected to reduced amounts of available/applied water (water input) or under conditions of water deficit stress or under conditions of acute or chronic drought.

As used herein, "drought stress" refers to a period of dryness (acute or chronic/prolonged) that results in water deficit and subjects plants to stress and/or damage to plant tissues and/or negatively affects grain/crop yield; a period of dryness (acute or chronic/prolonged) that results in water deficit and/or higher temperatures and subjects plants to stress and/or damage to plant tissues and/or negatively affects grain/crop yield.

As used herein, "water deficit" refers to the conditions or environments that provide less than optimal amounts of water needed for adequate/successful growth and development of plants.

As used herein, "water stress" refers to the conditions or environments that provide improper (either less/insufficient or more/excessive) amounts of water than that needed for adequate/successful growth and development of plants/crops thereby subjecting the plants to stress and/or damage to plant tissues and/or negatively affecting grain/crop yield.

As used herein "water deficit stress" refers to the conditions or environments that provide less/insufficient amounts of water than that needed for adequate/successful growth and

development of plants/crops thereby subjecting the plants to stress and/or damage to plant tissues and/or negatively affecting grain yield.

As used herein, the terms "nucleic acid," "nucleic acid molecule," "nucleotide sequence" and "polynucleotide" refer to RNA or DNA that is linear or branched, single or double stranded, or a hybrid thereof. The term also encompasses RNA/DNA hybrids. When dsRNA is produced synthetically, less common bases, such as inosine, 5-methylcytosine, 6-methyladenine, hypoxanthine and others can also be used for antisense, dsRNA, and ribozyme pairing. For example, polynucleotides that contain C-5 propyne analogues of uridine and cytidine have been shown to bind RNA with high affinity and to be potent antisense inhibitors of gene expression. Other modifications, such as modification to the phosphodiester backbone, or the 2'-hydroxy in the ribose sugar group of the RNA can also be made.

As used herein, the term "nucleotide sequence" refers to a heteropolymer of nucleotides or the sequence of these nucleotides from the 5' to 3' end of a nucleic acid molecule and includes DNA or RNA molecules, including cDNA, a DNA fragment or portion, genomic DNA, synthetic (e.g., chemically synthesized) DNA, plasmid DNA, mRNA, and anti-sense RNA, any of which can be single stranded or double stranded. The terms "nucleotide sequence" "nucleic acid," "nucleic acid molecule," "nucleic acid construct," "oligonucleotide" and "polynucleotide" are also used interchangeably herein to refer to a heteropolymer of nucleotides. Nucleic acid molecules and/or nucleotide sequences provided herein are presented herein in the 5' to 3' direction, from left to right and are represented using the standard code for representing the nucleotide characters as set forth in the U.S. sequence rules, 37 CFR §§1.821 - 1.825 and the World Intellectual Property Organization (WIPO) Standard ST.25. A "5' region" as used herein can mean the region of a polynucleotide that is nearest the 5' end of the polynucleotide. Thus, for example, an element in the 5' region of a polynucleotide can be located anywhere from the first nucleotide located at the 5' end of the polynucleotide to the nucleotide located halfway through the polynucleotide. A "3' region" as used herein can mean the region of a polynucleotide that is nearest the 3' end of the polynucleotide. Thus, for example, an element in the 3' region of a polynucleotide can be located anywhere from the first nucleotide located at the 3' end of the polynucleotide to the nucleotide located halfway through the polynucleotide.

As used herein with respect to nucleic acids, the term "fragment" or "portion" refers to a nucleic acid that is reduced in length relative (e.g., reduced by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 400,

450, 500, 550, 600, 650, 700, 750, 800, 850, or 900 or more nucleotides or any range or value therein) to a reference nucleic acid and that comprises, consists essentially of and/or consists of a nucleotide sequence of contiguous nucleotides identical or almost identical (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical) to a
 5 corresponding portion of the reference nucleic acid. Such a nucleic acid fragment may be, where appropriate, included in a larger polynucleotide of which it is a constituent. As an example, a repeat sequence of guide nucleic acid of this invention may comprise a "portion" of a wild type CRISPR-Cas repeat sequence (e.g., a wild type CRISPR-Cas repeat; e.g., a repeat
 10 from the CRISPR Cas system of, for example, a Cas9, Cas12a (Cpf1), Cas12b, Cas12c (C2c3), Cas12d (CasY), Cas12e (CasX), Cas12g, Cas12h, Cas12i, C2c4, C2c5, C2c8, C2c9, C2c10, Cas14a, Cas14b, and/or a Cas14c, and the like).

In some embodiments, a nucleic acid fragment may comprise, consist essentially of or consist of about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27,
 15 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 320, 330, 340, 350, 360, 370, 380, 390, 395, 400, 410, 415, 420, 425, 430, 435, 440, 445, 450, 500,
 20 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1900, 2000, 3000, 4000 or 5000 or more consecutive nucleotides, or any range or value therein, of a nucleic acid encoding a SHP polypeptide, optionally a fragment of a *SHP* gene may be about 10, 11, 12, 13, 14, 15, 16, 17,
 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43,
 25 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150 consecutive nucleotides to about 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 340, 345, 350,
 30 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740 or 750 or more consecutive nucleotides in length, or any range or value therein (e.g., a fragment or portion of any one **SEQ ID NOs:69, 70, 100, 101, 148, 149,**

177, 178, 206, 207, 240 or 241 (e.g., **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**).

In some embodiments, a "sequence-specific nucleic acid binding domain" may bind to one or more fragments or portions of nucleotide sequences (e.g., DNA, RNA) encoding, for example, a Shatterproof MADS-box transcription factor (SHP) polypeptide as described herein.

As used herein with respect to polypeptides, the term "fragment" or "portion" may refer to a polypeptide that is reduced in length relative to a reference polypeptide and that comprises, consists essentially of and/or consists of an amino acid sequence of contiguous amino acids identical or almost identical (e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical) to a corresponding portion of the reference polypeptide. Such a polypeptide fragment may be, where appropriate, included in a larger polypeptide of which it is a constituent. In some embodiments, a polypeptide fragment may comprise, consist essentially of, or consist of at least about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 260, 270, 280, or 290 or more consecutive amino acids of a reference polypeptide. In some embodiments, a polypeptide fragment may comprise, consist essentially of or consist of about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, or 150, or more consecutive amino acid residues, or any range or value therein, of a SHP (e.g., a fragment or a portion of any one of the polypeptides of **SEQ ID NOs:71, 102, 150, 179, 208, or 242** (e.g., **SEQ ID NOs:100-102, 148-150, 177-179, 206-208, 240-242 or 292-294**)). In some embodiments, an SHP polypeptide fragment may comprise, consist essentially of or consist of about 13, 59, 63, 64, 68, 69, 71, 72, 76, or 77 consecutive amino acid residues (see, e.g., **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**).

In some embodiments, a fragment of a SHP polypeptide can be a truncated SHP polypeptide resulting from a mutation of the SHP genomic sequence encoding the SHP polypeptide as described herein. For example, a fragment of a SHP polypeptide can be the N-terminus of an SHP polypeptide or a portion thereof (see, e.g., about the first 170-190 consecutive amino acid residues (e.g., the first 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190 consecutive amino acid residues, e.g., the first 172 to 185 (e.g., 172, 173, 174, 175, 176, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, or 189) consecutive amino acid residues, and any range or value therein) (e.g., **SEQ ID**

NOs:71, 102, 150, 179, 208, or 242). In some embodiments, a fragment of an SHP polypeptide may be the result of a mutation generated in at least one endogenous gene encoding an SHP polypeptide (e.g., an endogenous *SHPI* gene, an endogenous *SHPI2* gene, an endogenous *SHPI3* gene, and/or an endogenous *SHPI4* gene) as described herein (e.g., a deletion, insertion and the like, in one or more of the endogenous *SHPI* genes in a canola plant). In some embodiments, SHP polypeptides edited as described herein are N-terminal truncated polypeptides missing the first about 170-190 consecutive amino acids, see, e.g., **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, or 317.** In some embodiments, a truncated SHP polypeptide (N-terminal fragment) may comprise at least one amino acid substitution at the C-terminal end of the truncated polypeptide that are not present in the polypeptide encoded by the endogenous *SHPI* gene (e.g., a substitution of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid residues at the C-terminus, optionally 1 amino acid residue to about 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues) (see, e.g., **SEQ ID NO:313 or SEQ ID NO:315).**

In some embodiments, a fragment of a SHP polypeptide may be from the C-terminus of the SHP polypeptide. For example, a fragment of a SHP polypeptide can be about the last 76 or 77 consecutive amino acid residues of an SHP polypeptide or a portion thereof (see, e.g., **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291).** In some embodiments, a fragment of a SHP polypeptide can comprise the portion of the SHP polypeptide encoded by the second to last exon and/or the last exon of an *SHPI* gene. In some embodiments, a fragment of a SHP polypeptide can comprise the portion of the SHP polypeptide encoded by all but the second to last exon and/or the last exon of an *SHPI* gene.

In some embodiments, such a deletion when comprised in a canola plant can result in the canola plant exhibiting reduced pod shatter and/or reduced lignification (reduced lignin content) in the pod valve margin, as compared to a canola plant not comprising (devoid of) said deletion. An *SHPI* gene may be edited in one or more than one location (and using one or more different editing tools), thereby providing a *SHPI* gene comprising one or more than one mutation. In some embodiments, an SHP polypeptide mutated as described herein may comprise one or more than one edit that may result in a polypeptide having a deletion of one or more amino acid residues (e.g., a deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, or more consecutive amino acid residue, and any range or value therein (e.g., a truncated polypeptide), optionally a deletion of about 100 to about 600 consecutive amino acid

residues (e.g., about 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, or 600, and any range or value therein).

5 In some embodiments, a "portion" or "region" in reference to a nucleic acid means at least 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 110, 115, 120, 10 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 285, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 395, 400, 405, 410, 415, 420, 425, 430, 435, 440, 445, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2500, 3000, 3500, 4000, 4500, or 5000 or more consecutive nucleotides from a gene (e.g., consecutive nucleotides from an *SHP* gene), 15 optionally a "portion" or "region" of a *SHP* gene may be about 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 consecutive nucleotides to about 155, 160, 165, 170, 175, 180, 20 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740 or 750 or more consecutive nucleotides in length, or any range or value therein (e.g., a portion or 25 region of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241** (e.g., **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**).

In some embodiments, a "portion" or "region" of a *SHP* polypeptide sequence may be 30 about 5 to about 200 or more consecutive amino acid residues in length (e.g., about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87,

88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 176, 177, 178, 179, 180, 181, 812, 183, 184, 185, 186, 187, 188, 189, 190, 195, or 200, or more consecutive amino acid residues in length (e.g., a portion of any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, optionally **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**) (e.g., a C-terminal portion of a SHP polypeptide, e.g., **SEQ ID NOs:299, 301, 303, 305, 307, 309, or 317**).

As used herein with respect to nucleic acids, the term "functional fragment" refers to nucleic acid that encodes a functional fragment of a polypeptide. A "functional fragment" with respect to a polypeptide is a fragment of a polypeptide that retains one or more of the activities of the native reference polypeptide.

The term "gene," as used herein, refers to a nucleic acid molecule capable of being used to produce mRNA, antisense RNA, miRNA, anti-microRNA antisense oligodeoxyribonucleotide (AMO) and the like. Genes may or may not be capable of being used to produce a functional protein or gene product. Genes can include both coding and non-coding regions (e.g., introns, regulatory elements, promoters, enhancers, termination sequences and/or 5' and 3' untranslated regions). A gene may be "isolated" by which is meant a nucleic acid that is substantially or essentially free from components normally found in association with the nucleic acid in its natural state. Such components include other cellular material, culture medium from recombinant production, and/or various chemicals used in chemically synthesizing the nucleic acid.

The term "mutation" refers to point mutations (e.g., missense, or nonsense, or insertions or deletions of single base pairs that result in frame shifts), insertions, deletions, inversions and/or truncations. When the mutation is a substitution of a residue within an amino acid sequence with another residue, or a deletion or insertion of one or more residues within a sequence, the mutations are typically described by identifying the original residue followed by the position of the residue within the sequence and by the identity of the newly substituted residue. A truncation can include a truncation at the C-terminal end of a polypeptide or at the N-terminal end of a polypeptide. A truncation of a polypeptide can be the result of a deletion of the corresponding 5' end or 3' end of the gene encoding the polypeptide. A frameshift mutation can occur when deletions or insertions of one or more base pairs are introduced into a gene, optionally resulting in an out-of-frame mutation or an in-frame mutation. Frameshift mutations in a gene can result in the production of a polypeptide that is longer, shorter or the same length as the wild type polypeptide depending on when the first stop codon occurs following the

mutated region of the gene. As an example, an out-of-frame mutation that produces a premature stop codon can produce a polypeptide that is shorter than the wild type polypeptide, or, in some embodiments, the polypeptide may be absent/undetectable. A DNA inversion is the result of a rotation of a genetic fragment within a region of a chromosome.

5 The terms "complementary" or "complementarity," as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T" (5' to 3') binds to the complementary sequence "T-C-A" (3' to 5'). Complementarity between two single-stranded molecules may be "partial," in which only some of the nucleotides bind, or it may be complete when total complementarity exists
10 between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands.

 "Complement," as used herein, can mean 100% complementarity with the comparator nucleotide sequence or it can mean less than 100% complementarity (e.g., about 70%, 71%,
15 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, and the like, complementarity) to the comparator nucleotide sequence.

 Different nucleic acids or proteins having homology are referred to herein as "homologues." The term homologue includes homologous sequences from the same and from
20 other species and orthologous sequences from the same and other species. "Homology" refers to the level of similarity between two or more nucleic acid and/or amino acid sequences in terms of percent of positional identity (*i.e.*, sequence similarity or identity). Homology also refers to the concept of similar functional properties among different nucleic acids or proteins. Thus, the compositions and methods of the invention further comprise homologues to the nucleotide
25 sequences and polypeptide sequences of this invention. "Orthologous," as used herein, refers to homologous nucleotide sequences and/ or amino acid sequences in different species that arose from a common ancestral gene during speciation. A homologue of a nucleotide sequence of this invention has a substantial sequence identity (e.g., at least about 70%, 71%, 72%, 73%, 74%,
75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,
30 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100%) to said nucleotide sequence of the invention.

 As used herein "sequence identity" refers to the extent to which two optimally aligned polynucleotide or polypeptide sequences are invariant throughout a window of alignment of

components, *e.g.*, nucleotides or amino acids. "Identity" can be readily calculated by known methods including, but not limited to, those described in: *Computational Molecular Biology* (Lesk, A. M., ed.) Oxford University Press, New York (1988); *Biocomputing: Informatics and Genome Projects* (Smith, D. W., ed.) Academic Press, New York (1993); *Computer Analysis of Sequence Data, Part I* (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology* (von Heinje, G., ed.) Academic Press (1987); and *Sequence Analysis Primer* (Gribskov, M. and Devereux, J., eds.) Stockton Press, New York (1991).

As used herein, the term "percent sequence identity" or "percent identity" refers to the percentage of identical nucleotides in a linear polynucleotide sequence of a reference ("query") polynucleotide molecule (or its complementary strand) as compared to a test ("subject") polynucleotide molecule (or its complementary strand) when the two sequences are optimally aligned. In some embodiments, "percent sequence identity" can refer to the percentage of identical amino acids in an amino acid sequence as compared to a reference polypeptide.

As used herein, the phrase "substantially identical," or "substantial identity" in the context of two nucleic acid molecules, nucleotide sequences, or polypeptide sequences, refers to two or more sequences or subsequences that have at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100% nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. In some embodiments of the invention, the substantial identity exists over a region of consecutive nucleotides of a nucleotide sequence of the invention that is about 10 nucleotides to about 20 nucleotides, about 10 nucleotides to about 25 nucleotides, about 10 nucleotides to about 30 nucleotides, about 15 nucleotides to about 25 nucleotides, about 30 nucleotides to about 40 nucleotides, about 50 nucleotides to about 60 nucleotides, about 70 nucleotides to about 80 nucleotides, about 90 nucleotides to about 100 nucleotides, about 100 nucleotides to about 200 nucleotides, about 100 nucleotides to about 300 nucleotides, about 100 nucleotides to about 400 nucleotides, about 100 nucleotides to about 500 nucleotides, about 100 nucleotides to about 600 nucleotides, about 100 nucleotides to about 800 nucleotides, about 100 nucleotides to about 900 nucleotides, or more in length, or any range therein, up to the full length of the sequence. In some embodiments, nucleotide sequences can be substantially identical over at least about 20

nucleotides (e.g., about 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 50, 60, 70, or 80 nucleotides or more).

In some embodiments of the invention, the substantial identity exists over a region of consecutive amino acid residues of a polypeptide of the invention that is about 3 amino acid residues to about 20 amino acid residues, about 5 amino acid residues to about 25 amino acid residues, about 7 amino acid residues to about 30 amino acid residues, about 10 amino acid residues to about 25 amino acid residues, about 15 amino acid residues to about 30 amino acid residues, about 20 amino acid residues to about 40 amino acid residues, about 25 amino acid residues to about 40 amino acid residues, about 25 amino acid residues to about 50 amino acid residues, about 30 amino acid residues to about 50 amino acid residues, about 40 amino acid residues to about 50 amino acid residues, about 40 amino acid residues to about 70 amino acid residues, about 50 amino acid residues to about 70 amino acid residues, about 60 amino acid residues to about 80 amino acid residues, about 70 amino acid residues to about 80 amino acid residues, about 90 amino acid residues to about 100 amino acid residues, or more amino acid residues in length, and any range therein, up to the full length of the sequence. In some embodiments, polypeptide sequences can be substantially identical to one another over at least about 8 consecutive amino acid residues (e.g., about 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 130, 140, 150, 175, 200, 225, 250, 300, 350 or more amino acids in length or more consecutive amino acid residues). In some embodiments, two or more SHP polypeptides may be identical or substantially identical (e.g., at least 70% to 99.9% identical; e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% identical or any range or value therein) over at least 8 consecutive amino acids to about 350 consecutive amino acids. In some embodiments, two or more SHP polypeptides may be identical or substantially identical over at least 8, 9, 10, 11, 12, 13, 14, or 15 consecutive amino acids to about 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 consecutive amino acids).

For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference

sequences are entered into a computer, subsequence coordinates are designated if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

5 Optimal alignment of sequences for aligning a comparison window are well known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and optionally by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the
10 GCG® Wisconsin Package® (Accelrys Inc., San Diego, CA). An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components which are shared by the two aligned sequences divided by the total number of components in the reference sequence segment, *e.g.*, the entire reference sequence or a smaller defined part of the reference sequence. Percent sequence identity is represented as the identity fraction multiplied
15 by 100. The comparison of one or more polynucleotide sequences may be to a full-length polynucleotide sequence or a portion thereof, or to a longer polynucleotide sequence. For purposes of this invention "percent identity" may also be determined using BLASTX version 2.0 for translated nucleotide sequences and BLASTN version 2.0 for polynucleotide sequences.

Two nucleotide sequences may also be considered substantially complementary when
20 the two sequences hybridize to each other under stringent conditions. In some embodiments, two nucleotide sequences considered to be substantially complementary hybridize to each other under highly stringent conditions.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridizations
25 are sequence dependent and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes* part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York (1993). Generally, highly stringent hybridization and wash conditions are selected to
30 be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH.

The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected

to be equal to the T_m for a particular probe. An example of stringent hybridization conditions for hybridization of complementary nucleotide sequences which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formamide with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.1 5M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see*, Sambrook, *infra*, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example of a medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes. An example of a low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. For short probes (e.g., about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleotide sequences that do not hybridize to each other under stringent conditions are still substantially identical if the proteins that they encode are substantially identical. This can occur, for example, when a copy of a nucleotide sequence is created using the maximum codon degeneracy permitted by the genetic code.

A polynucleotide and/or recombinant nucleic acid construct of this invention (e.g., expression cassettes and/or vectors) may be codon optimized for expression. In some embodiments, the polynucleotides, nucleic acid constructs, expression cassettes, and/or vectors of the editing systems of the invention (e.g., comprising/encoding a sequence-specific nucleic acid binding domain (e.g., a sequence-specific nucleic acid binding domain (e.g., DNA binding domain) from a polynucleotide-guided endonuclease, a zinc finger nuclease, a transcription activator-like effector nuclease (TALEN), an Argonaute protein, and/or a CRISPR-Cas endonuclease (e.g., CRISPR-Cas effector protein) (e.g., a Type I CRISPR-Cas effector protein, a Type II CRISPR-Cas effector protein, a Type III CRISPR-Cas effector protein, a Type IV CRISPR-Cas effector protein, a Type V CRISPR-Cas effector protein or a Type VI CRISPR-Cas effector protein)), a nuclease (e.g., an endonuclease (e.g., FokI), a polynucleotide-guided endonuclease, a CRISPR-Cas endonuclease (e.g., CRISPR-Cas effector protein), a zinc finger nuclease, and/or a transcription activator-like effector nuclease (TALEN)), deaminase

proteins/domains (e.g., adenine deaminase, cytosine deaminase), a polynucleotide encoding a reverse transcriptase protein or domain, a polynucleotide encoding a 5'-3' exonuclease polypeptide, and/or affinity polypeptides, peptide tags, etc.) may be codon optimized for expression in a plant. In some embodiments, the codon optimized nucleic acids,

5 polynucleotides, expression cassettes, and/or vectors of the invention have about 70% to about 99.9% (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% or 100%) identity or more to the reference nucleic acids, polynucleotides, expression cassettes, and/or vectors that have not been codon optimized.

10 In any of the embodiments described herein, a polynucleotide or nucleic acid construct of the invention may be operatively associated with a variety of promoters and/or other regulatory elements for expression in a plant and/or a cell of a plant. Thus, in some embodiments, a polynucleotide or nucleic acid construct of this invention may further comprise one or more promoters, introns, enhancers, and/or terminators operably linked to one or more
15 nucleotide sequences. In some embodiments, a promoter may be operably associated with an intron (e.g., Ubi1 promoter and intron). In some embodiments, a promoter associated with an intron maybe referred to as a "promoter region" (e.g., Ubi1 promoter and intron).

By "operably linked" or "operably associated" as used herein in reference to polynucleotides, it is meant that the indicated elements are functionally related to each other and
20 are also generally physically related. Thus, the term "operably linked" or "operably associated" as used herein, refers to nucleotide sequences on a single nucleic acid molecule that are functionally associated. Thus, a first nucleotide sequence that is operably linked to a second nucleotide sequence means a situation when the first nucleotide sequence is placed in a functional relationship with the second nucleotide sequence. For instance, a promoter is
25 operably associated with a nucleotide sequence if the promoter effects the transcription or expression of said nucleotide sequence. Those skilled in the art will appreciate that the control sequences (e.g., promoter) need not be contiguous with the nucleotide sequence to which it is operably associated, as long as the control sequences function to direct the expression thereof. Thus, for example, intervening untranslated, yet transcribed, nucleic acid sequences can be
30 present between a promoter and the nucleotide sequence, and the promoter can still be considered "operably linked" to the nucleotide sequence.

As used herein, the term "linked," in reference to polypeptides, refers to the attachment of one polypeptide to another. A polypeptide may be linked to another polypeptide (at the N-

terminus or the C-terminus) directly (e.g., via a peptide bond) or through a linker.

The term "linker" is art-recognized and refers to a chemical group, or a molecule linking two molecules or moieties, e.g., two domains of a fusion protein, such as, for example, a nucleic acid binding polypeptide or domain and peptide tag and/or a reverse transcriptase and an affinity polypeptide that binds to the peptide tag; or a DNA endonuclease polypeptide or domain and peptide tag and/or a reverse transcriptase and an affinity polypeptide that binds to the peptide tag. A linker may be comprised of a single linking molecule or may comprise more than one linking molecule. In some embodiments, the linker can be an organic molecule, group, polymer, or chemical moiety such as a bivalent organic moiety. In some embodiments, the linker may be an amino acid or it may be a peptide. In some embodiments, the linker is a peptide.

In some embodiments, a peptide linker useful with this invention may be about 2 to about 100 or more amino acids in length, for example, about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more amino acids in length (e.g., about 2 to about 40, about 2 to about 50, about 2 to about 60, about 4 to about 40, about 4 to about 50, about 4 to about 60, about 5 to about 40, about 5 to about 50, about 5 to about 60, about 9 to about 40, about 9 to about 50, about 9 to about 60, about 10 to about 40, about 10 to about 50, about 10 to about 60, or about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 amino acids to about 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more amino acids in length (e.g., about 105, 110, 115, 120, 130, 140 150 or more amino acids in length). In some embodiments, a peptide linker may be a GS linker.

As used herein, the term "linked," or "fused" in reference to polynucleotides, refers to the attachment of one polynucleotide to another. In some embodiments, two or more polynucleotide molecules may be linked by a linker that can be an organic molecule, group, polymer, or chemical moiety such as a bivalent organic moiety. A polynucleotide may be linked or fused to another polynucleotide (at the 5' end or the 3' end) via a covalent or non-covalent linkage or binding, including e.g., Watson-Crick base-pairing, or through one or more linking nucleotides. In some embodiments, a polynucleotide motif of a certain structure may be inserted within another polynucleotide sequence (e.g., extension of the hairpin structure in the guide RNA). In

some embodiments, the linking nucleotides may be naturally occurring nucleotides. In some embodiments, the linking nucleotides may be non-naturally occurring nucleotides.

A "promoter" is a nucleotide sequence that controls or regulates the transcription of a nucleotide sequence (e.g., a coding sequence) that is operably associated with the promoter. The coding sequence controlled or regulated by a promoter may encode a polypeptide and/or a functional RNA. Typically, a "promoter" refers to a nucleotide sequence that contains a binding site for RNA polymerase II and directs the initiation of transcription. In general, promoters are found 5', or upstream, relative to the start of the coding region of the corresponding coding sequence. A promoter may comprise other elements that act as regulators of gene expression; e.g., a promoter region. These include a TATA box consensus sequence, and often a CAAT box consensus sequence (Breathnach and Chambon, (1981) *Annu. Rev. Biochem.* 50:349). In plants, the CAAT box may be substituted by the AGGA box (Messing et al., (1983) in *Genetic Engineering of Plants*, T. Kosuge, C. Meredith and A. Hollaender (eds.), Plenum Press, pp. 211-227).

Promoters useful with this invention can include, for example, constitutive, inducible, temporally regulated, developmentally regulated, chemically regulated, tissue-preferred and/or tissue-specific promoters for use in the preparation of recombinant nucleic acid molecules, e.g., "synthetic nucleic acid constructs" or "protein-RNA complex." These various types of promoters are known in the art.

The choice of promoter may vary depending on the temporal and spatial requirements for expression, and also may vary based on the host cell to be transformed. Promoters for many different organisms are well known in the art. Based on the extensive knowledge present in the art, the appropriate promoter can be selected for the particular host organism of interest. Thus, for example, much is known about promoters upstream of highly constitutively expressed genes in model organisms and such knowledge can be readily accessed and implemented in other systems as appropriate.

In some embodiments, a promoter functional in a plant may be used with the constructs of this invention. Non-limiting examples of a promoter useful for driving expression in a plant include the promoter of the RubisCo small subunit gene 1 (PrbcS1), the promoter of the actin gene (Pactin), the promoter of the nitrate reductase gene (Pnr) and the promoter of duplicated carbonic anhydrase gene 1 (Pdca1) (See, Walker et al. *Plant Cell Rep.* 23:727-735 (2005); Li et al. *Gene* 403:132-142 (2007); Li et al. *Mol Biol. Rep.* 37:1143-1154 (2010)). PrbcS1 and Pactin are constitutive promoters and Pnr and Pdca1 are inducible promoters. Pnr is induced by nitrate

and repressed by ammonium (Li et al. *Gene* 403:132-142 (2007)) and *Pdca1* is induced by salt (Li et al. *Mol Biol. Rep.* 37:1143-1154 (2010)). In some embodiments, a promoter useful with this invention is RNA polymerase II (Pol II) promoter. In some embodiments, a U6 promoter or a 7SL promoter from *Zea mays* may be useful with constructs of this invention. In some
5 embodiments, the U6c promoter and/or 7SL promoter from *Zea mays* may be useful for driving expression of a guide nucleic acid. In some embodiments, a U6c promoter, U6i promoter and/or 7SL promoter from *Glycine max* may be useful with constructs of this invention. In some embodiments, the U6c promoter, U6i promoter and/or 7SL promoter from *Glycine max* may be useful for driving expression of a guide nucleic acid.

10 Examples of constitutive promoters useful for plants include, but are not limited to, *cauliflower mosaic virus* promoter (*cmp*) (U.S. Patent No. 7,166,770), the rice actin 1 promoter (Wang et al. (1992) *Mol. Cell. Biol.* 12:3399-3406; as well as US Patent No. 5,641,876), *CaMV* 35S promoter (Odell et al. (1985) *Nature* 313:810-812), *CaMV* 19S promoter (Lawton et al. (1987) *Plant Mol. Biol.* 9:315-324), *nos* promoter (Ebert et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:5745-5749), *Adh* promoter (Walker et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:6624-6629),
15 sucrose synthase promoter (Yang & Russell (1990) *Proc. Natl. Acad. Sci. USA* 87:4144-4148), and the ubiquitin promoter. The constitutive promoter derived from ubiquitin accumulates in many cell types. Ubiquitin promoters have been cloned from several plant species for use in transgenic plants, for example, sunflower (Binet et al., 1991. *Plant Science* 79: 87-94), maize (Christensen et al., 1989. *Plant Molec. Biol.* 12: 619-632), and *Arabidopsis* (Norris et al. 1993. *Plant Molec. Biol.* 21:895-906). The maize ubiquitin promoter (UbiP) has been developed in
20 transgenic monocot systems and its sequence and vectors constructed for monocot transformation are disclosed in the patent publication EP 0 342 926. The ubiquitin promoter is suitable for the expression of the nucleotide sequences of the invention in transgenic plants,
25 especially monocotyledons. Further, the promoter expression cassettes described by McElroy et al. (*Mol. Gen. Genet.* 231: 150-160 (1991)) can be easily modified for the expression of the nucleotide sequences of the invention and are particularly suitable for use in monocotyledonous hosts.

30 In some embodiments, tissue specific/tissue preferred promoters can be used for expression of a heterologous polynucleotide in a plant cell. Tissue specific or preferred expression patterns include, but are not limited to, green tissue specific or preferred, root specific or preferred, stem specific or preferred, flower specific or preferred or pollen specific or preferred. Promoters suitable for expression in green tissue include many that regulate genes

involved in photosynthesis and many of these have been cloned from both monocotyledons and dicotyledons. In one embodiment, a promoter useful with the invention is the maize PEPC promoter from the phosphoenol carboxylase gene (Hudspeth & Grula, *Plant Molec. Biol.*

12:579-589 (1989)). Non-limiting examples of tissue-specific promoters include those

5 associated with genes encoding the seed storage proteins (such as β -conglycinin, cruciferin, napin and phaseolin), zein or oil body proteins (such as oleosin), or proteins involved in fatty acid biosynthesis (including acyl carrier protein, stearyl-ACP desaturase and fatty acid desaturases (fad 2-1)), and other nucleic acids expressed during embryo development (such as Bce4, see, e.g., Kridl et al. (1991) *Seed Sci. Res.* 1:209-219; as well as EP Patent No. 255378).

10 Tissue-specific or tissue-preferential promoters useful for the expression of the nucleotide sequences of the invention in plants, particularly maize, include but are not limited to those that direct expression in root, pith, leaf or pollen. Such promoters are disclosed, for example, in WO 93/07278, herein incorporated by reference in its entirety. Other non-limiting examples of tissue specific or tissue preferred promoters useful with the invention the cotton rubisco promoter

15 disclosed in US Patent 6,040,504; the rice sucrose synthase promoter disclosed in US Patent 5,604,121; the root specific promoter described by de Framond (FEBS 290:103-106 (1991); EP 0 452 269 to Ciba-Geigy); the stem specific promoter described in U.S. Patent 5,625,136 (to Ciba-Geigy) and which drives expression of the maize trpA gene; the cestrum yellow leaf curling virus promoter disclosed in WO 01/73087; and pollen specific or preferred promoters including, but not limited to, ProOsLPS10 and ProOsLPS11 from rice (Nguyen et al. *Plant Biotechnol. Reports* 9(5):297-306 (2015)), ZmSTK2_USP from maize (Wang et al. *Genome* 60(6):485-495 (2017)), LAT52 and LAT59 from tomato (Twell et al. *Development* 109(3):705-713 (1990)), Zm13 (U.S. Patent No. 10,421,972), PLA₂- δ promoter from arabidopsis (U.S. Patent No. 7,141,424), and/or the ZmC5 promoter from maize (International PCT Publication

25 No. WO1999/042587.

Additional examples of plant tissue-specific/tissue preferred promoters include, but are not limited to, the root hair-specific cis-elements (RHEs) (Kim et al. *The Plant Cell* 18:2958-2970 (2006)), the root-specific promoters RCc3 (Jeong et al. *Plant Physiol.* 153:185-197 (2010)) and RB7 (U.S. Patent No. 5459252), the lectin promoter (Lindstrom et al. (1990) *Der. Genet.* 11:160-167; and Vodkin (1983) *Prog. Clin. Biol. Res.* 138:87-98), corn alcohol dehydrogenase 1 promoter (Dennis et al. (1984) *Nucleic Acids Res.* 12:3983-4000), S-adenosyl-L-methionine synthetase (SAMS) (Vander Mijnsbrugge et al. (1996) *Plant and Cell Physiology*, 37(8):1108-1115), corn light harvesting complex promoter (Bansal et al. (1992) *Proc. Natl. Acad. Sci. USA*

89:3654-3658), corn heat shock protein promoter (O'Dell et al. (1985) *EMBO J.* 5:451-458; and Rochester et al. (1986) *EMBO J.* 5:451-458), pea small subunit RuBP carboxylase promoter (Cashmore, "Nuclear genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase" pp. 29-39 In: *Genetic Engineering of Plants* (Hollaender ed., Plenum Press 1983; and Poulsen et al. (1986) *Mol. Gen. Genet.* 205:193-200), Ti plasmid mannopine synthase promoter (Langridge et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3219-3223), Ti plasmid nopaline synthase promoter (Langridge et al. (1989), *supra*), petunia chalcone isomerase promoter (van Tunen et al. (1988) *EMBO J.* 7:1257-1263), bean glycine rich protein 1 promoter (Keller et al. (1989) *Genes Dev.* 3:1639-1646), truncated CaMV 35S promoter (O'Dell et al. (1985) *Nature* 313:810-812), potato patatin promoter (Wenzler et al. (1989) *Plant Mol. Biol.* 13:347-354), root cell promoter (Yamamoto et al. (1990) *Nucleic Acids Res.* 18:7449), maize zein promoter (Kriz et al. (1987) *Mol. Gen. Genet.* 207:90-98; Langridge et al. (1983) *Cell* 34:1015-1022; Reina et al. (1990) *Nucleic Acids Res.* 18:6425; Reina et al. (1990) *Nucleic Acids Res.* 18:7449; and Wandelt et al. (1989) *Nucleic Acids Res.* 17:2354), globulin-1 promoter (Belanger et al. (1991) *Genetics* 129:863-872), α -tubulin cab promoter (Sullivan et al. (1989) *Mol. Gen. Genet.* 215:431-440), PEPCase promoter (Hudspeth & Grula (1989) *Plant Mol. Biol.* 12:579-589), R gene complex-associated promoters (Chandler et al. (1989) *Plant Cell* 1:1175-1183), and chalcone synthase promoters (Franken et al. (1991) *EMBO J.* 10:2605-2612).

Useful for seed-specific expression is the pea vicilin promoter (Czako et al. (1992) *Mol. Gen. Genet.* 235:33-40; as well as the seed-specific promoters disclosed in U.S. Patent No. 5,625,136. Useful promoters for expression in mature leaves are those that are switched at the onset of senescence, such as the SAG promoter from *Arabidopsis* (Gan et al. (1995) *Science* 270:1986-1988).

In addition, promoters functional in chloroplasts can be used. Non-limiting examples of such promoters include the bacteriophage T3 gene 9 5' UTR and other promoters disclosed in U.S. Patent No. 7,579,516. Other promoters useful with the invention include but are not limited to the S-E9 small subunit RuBP carboxylase promoter and the Kunitz trypsin inhibitor gene promoter (Kti3).

Additional regulatory elements useful with this invention include, but are not limited to, introns, enhancers, termination sequences and/or 5' and 3' untranslated regions.

An intron useful with this invention can be an intron identified in and isolated from a plant and then inserted into an expression cassette to be used in transformation of a plant. As would be understood by those of skill in the art, introns can comprise the sequences required for

self-excision and are incorporated into nucleic acid constructs/expression cassettes in frame. An intron can be used either as a spacer to separate multiple protein-coding sequences in one nucleic acid construct, or an intron can be used inside one protein-coding sequence to, for example, stabilize the mRNA. If they are used within a protein-coding sequence, they are inserted "in-frame" with the excision sites included. Introns may also be associated with promoters to improve or modify expression. As an example, a promoter/intron combination useful with this invention includes but is not limited to that of the maize Ubi1 promoter and intron (see, e.g., **SEQ ID NO:21** and **SEQ ID NO:22**).

Non-limiting examples of introns useful with the present invention include introns from the ADHI gene (e.g., Adh1-S introns 1, 2 and 6), the ubiquitin gene (Ubi1), the RuBisCO small subunit (*rbcS*) gene, the RuBisCO large subunit (*rbcL*) gene, the actin gene (e.g., actin-1 intron), the pyruvate dehydrogenase kinase gene (*pdk*), the nitrate reductase gene (*nr*), the duplicated carbonic anhydrase gene 1 (*Tdca1*), the *psbA* gene, the *atpA* gene, or any combination thereof.

In some embodiments, a polynucleotide and/or a nucleic acid construct of the invention can be an "expression cassette" or can be comprised within an expression cassette. As used herein, "expression cassette" means a recombinant nucleic acid molecule comprising, for example, a one or more polynucleotides of the invention (e.g., a polynucleotide encoding a sequence-specific nucleic acid binding domain, a polynucleotide encoding a deaminase protein or domain, a polynucleotide encoding a reverse transcriptase protein or domain, a polynucleotide encoding a 5'-3' exonuclease polypeptide or domain, a guide nucleic acid and/or reverse transcriptase (RT) template), wherein polynucleotide(s) is/are operably associated with one or more control sequences (e.g., a promoter, terminator and the like). Thus, in some embodiments, one or more expression cassettes may be provided, which are designed to express, for example, a nucleic acid construct of the invention (e.g., a polynucleotide encoding a sequence-specific nucleic acid binding domain, a polynucleotide encoding a nuclease polypeptide/domain, a polynucleotide encoding a deaminase protein/domain, a polynucleotide encoding a reverse transcriptase protein/domain, a polynucleotide encoding a 5'-3' exonuclease polypeptide/domain, a polynucleotide encoding a peptide tag, and/or a polynucleotide encoding an affinity polypeptide, and the like, or comprising a guide nucleic acid, an extended guide nucleic acid, and/or RT template, and the like). When an expression cassette of the present invention comprises more than one polynucleotide, the polynucleotides may be operably linked to a single promoter that drives expression of all of the polynucleotides or the polynucleotides may be operably linked to one or more separate promoters (e.g., three polynucleotides may be

driven by one, two or three promoters in any combination). When two or more separate promoters are used, the promoters may be the same promoter, or they may be different promoters. Thus, a polynucleotide encoding a sequence specific nucleic acid binding domain, a polynucleotide encoding a nuclease protein/domain, a polynucleotide encoding a CRISPR-Cas effector protein/domain, a polynucleotide encoding an deaminase protein/domain, a polynucleotide encoding a reverse transcriptase polypeptide/domain (e.g., RNA-dependent DNA polymerase), and/or a polynucleotide encoding a 5'-3' exonuclease polypeptide/domain, a guide nucleic acid, an extended guide nucleic acid and/or RT template when comprised in a single expression cassette may each be operably linked to a single promoter, or separate promoters in any combination.

An expression cassette comprising a nucleic acid construct of the invention may be chimeric, meaning that at least one (e.g., one or more) of its components is heterologous with respect to at least one of its other components (e.g., a promoter from the host organism operably linked to a polynucleotide of interest to be expressed in the host organism, wherein the polynucleotide of interest is from a different organism than the host or is not normally found in association with that promoter). An expression cassette may also be one that is naturally occurring but has been obtained in a recombinant form useful for heterologous expression.

An expression cassette can optionally include a transcriptional and/or translational termination region (i.e., termination region) and/or an enhancer region that is functional in the selected host cell. A variety of transcriptional terminators and enhancers are known in the art and are available for use in expression cassettes. Transcriptional terminators are responsible for the termination of transcription and correct mRNA polyadenylation. A termination region and/or the enhancer region may be native to the transcriptional initiation region, may be native to, for example, a gene encoding a sequence-specific nucleic acid binding protein, a gene encoding a nuclease, a gene encoding a reverse transcriptase, a gene encoding a deaminase, and the like, or may be native to a host cell, or may be native to another source (e.g., foreign or heterologous to, for example, to a promoter, to a gene encoding a sequence-specific nucleic acid binding protein, a gene encoding a nuclease, a gene encoding a reverse transcriptase, a gene encoding a deaminase, and the like, or to the host cell, or any combination thereof).

An expression cassette of the invention also can include a polynucleotide encoding a selectable marker, which can be used to select a transformed host cell. As used herein, "selectable marker" means a polynucleotide sequence that when expressed imparts a distinct phenotype to the host cell expressing the marker and thus allows such transformed cells to be

distinguished from those that do not have the marker. Such a polynucleotide sequence may encode either a selectable or screenable marker, depending on whether the marker confers a trait that can be selected for by chemical means, such as by using a selective agent (e.g., an antibiotic and the like), or on whether the marker is simply a trait that one can identify through observation or testing, such as by screening (e.g., fluorescence). Many examples of suitable selectable markers are known in the art and can be used in the expression cassettes described herein.

In addition to expression cassettes, the nucleic acid molecules/constructs and polynucleotide sequences described herein can be used in connection with vectors. The term "vector" refers to a composition for transferring, delivering or introducing a nucleic acid (or nucleic acids) into a cell. A vector comprises a nucleic acid construct (e.g., expression cassette(s)) comprising the nucleotide sequence(s) to be transferred, delivered or introduced. Vectors for use in transformation of host organisms are well known in the art. Non-limiting examples of general classes of vectors include viral vectors, plasmid vectors, phage vectors, phagemid vectors, cosmid vectors, fosmid vectors, bacteriophages, artificial chromosomes, minicircles, or *Agrobacterium* binary vectors in double or single stranded linear or circular form which may or may not be self-transmissible or mobilizable. In some embodiments, a viral vector can include, but is not limited, to a retroviral, lentiviral, adenoviral, adeno-associated, or herpes simplex viral vector. A vector as defined herein can transform a prokaryotic or eukaryotic host either by integration into the cellular genome or exist extrachromosomally (e.g., autonomous replicating plasmid with an origin of replication). Additionally, included are shuttle vectors by which is meant a DNA vehicle capable, naturally or by design, of replication in two different host organisms, which may be selected from actinomycetes and related species, bacteria and eukaryotic (e.g., higher plant, mammalian, yeast or fungal cells). In some embodiments, the nucleic acid in the vector is under the control of, and operably linked to, an appropriate promoter or other regulatory elements for transcription in a host cell. The vector may be a bi-functional expression vector which functions in multiple hosts. In the case of genomic DNA, this may contain its own promoter and/or other regulatory elements and in the case of cDNA this may be under the control of an appropriate promoter and/or other regulatory elements for expression in the host cell. Accordingly, a nucleic acid or polynucleotide of this invention and/or expression cassettes comprising the same may be comprised in vectors as described herein and as known in the art.

As used herein, "contact," "contacting," "contacted," and grammatical variations thereof, refer to placing the components of a desired reaction together under conditions suitable for

carrying out the desired reaction (e.g., transformation, transcriptional control, genome editing, nicking, and/or cleavage). As an example, a target nucleic acid may be contacted with a sequence-specific nucleic acid binding protein (e.g., polynucleotide-guided endonuclease, a CRISPR-Cas endonuclease (e.g., CRISPR-Cas effector protein), a zinc finger nuclease, a transcription activator-like effector nuclease (TALEN) and/or an Argonaute protein)) and a deaminase or a nucleic acid construct encoding the same, under conditions whereby the sequence-specific nucleic acid binding protein, the reverse transcriptase and/or the deaminase are expressed and the sequence-specific nucleic acid binding protein binds to the target nucleic acid, and the reverse transcriptase and/or deaminase may be fused to either the sequence-specific nucleic acid binding protein or recruited to the sequence-specific nucleic acid binding protein (via, for example, a peptide tag fused to the sequence-specific nucleic acid binding protein and an affinity tag fused to the reverse transcriptase and/or deaminase) and thus, the deaminase and/or reverse transcriptase is positioned in the vicinity of the target nucleic acid, thereby modifying the target nucleic acid. Other methods for recruiting reverse transcriptase and/or deaminase may be used that take advantage of other protein-protein interactions, and also RNA-protein interactions and chemical interactions may be used for protein-protein and protein-nucleic acid recruitment.

As used herein, "modifying" or "modification" in reference to a target nucleic acid includes editing (e.g., mutating), covalent modification, exchanging/substituting nucleic acids/nucleotide bases, deleting, cleaving, nicking, and/or altering transcriptional control of a target nucleic acid. In some embodiments, a modification may include one or more single base changes (SNPs) of any type.

"Introducing," "introduce," "introduced" (and grammatical variations thereof) in the context of a polynucleotide of interest means presenting a nucleotide sequence of interest (e.g., polynucleotide, RT template, a nucleic acid construct, and/or a guide nucleic acid) to a plant, plant part thereof, or cell thereof, in such a manner that the nucleotide sequence gains access to the interior of a cell.

The terms "transformation" or "transfection" may be used interchangeably and as used herein refer to the introduction of a heterologous nucleic acid into a cell. Transformation of a cell may be stable or transient. Thus, in some embodiments, a host cell or host organism (e.g., a plant) may be stably transformed with a polynucleotide/nucleic acid molecule of the invention. In some embodiments, a host cell or host organism may be transiently transformed with a polynucleotide/nucleic acid molecule of the invention.

"Transient transformation" in the context of a polynucleotide means that a polynucleotide is introduced into the cell and does not integrate into the genome of the cell.

By "stably introducing" or "stably introduced" in the context of a polynucleotide introduced into a cell is intended that the introduced polynucleotide is stably incorporated into the genome of the cell, and thus the cell is stably transformed with the polynucleotide.

"Stable transformation" or "stably transformed" as used herein means that a nucleic acid molecule is introduced into a cell and integrates into the genome of the cell. As such, the integrated nucleic acid molecule is capable of being inherited by the progeny thereof, more particularly, by the progeny of multiple successive generations. "Genome" as used herein includes the nuclear and the plastid genome, and therefore includes integration of the nucleic acid into, for example, the chloroplast or mitochondrial genome. Stable transformation as used herein can also refer to a transgene that is maintained extrachromasomally, for example, as a minichromosome or a plasmid.

Transient transformation may be detected by, for example, an enzyme-linked immunosorbent assay (ELISA) or Western blot, which can detect the presence of a peptide or polypeptide encoded by one or more transgene introduced into an organism. Stable transformation of a cell can be detected by, for example, a Southern blot hybridization assay of genomic DNA of the cell with nucleic acid sequences which specifically hybridize with a nucleotide sequence of a transgene introduced into an organism (e.g., a plant). Stable transformation of a cell can be detected by, for example, a Northern blot hybridization assay of RNA of the cell with nucleic acid sequences which specifically hybridize with a nucleotide sequence of a transgene introduced into a host organism. Stable transformation of a cell can also be detected by, e.g., a polymerase chain reaction (PCR) or other amplification reactions as are well known in the art, employing specific primer sequences that hybridize with target sequence(s) of a transgene, resulting in amplification of the transgene sequence, which can be detected according to standard methods. Transformation can also be detected by direct sequencing and/or hybridization protocols well known in the art.

Accordingly, in some embodiments, nucleotide sequences, polynucleotides, nucleic acid constructs, and/or expression cassettes of the invention may be expressed transiently and/or they can be stably incorporated into the genome of the host organism. Thus, in some embodiments, a nucleic acid construct of the invention (e.g., one or more expression cassettes comprising polynucleotides for editing as described herein) may be transiently introduced into a cell with a guide nucleic acid and as such, no DNA is maintained in the cell.

A nucleic acid construct of the invention may be introduced into a plant cell by any method known to those of skill in the art. Non-limiting examples of transformation methods include transformation via bacterial-mediated nucleic acid delivery (e.g., via *Agrobacteria*), viral-mediated nucleic acid delivery, silicon carbide or nucleic acid whisker-mediated nucleic acid delivery, liposome mediated nucleic acid delivery, microinjection, microparticle bombardment, calcium-phosphate-mediated transformation, cyclodextrin-mediated transformation, electroporation, nanoparticle-mediated transformation, sonication, infiltration, PEG-mediated nucleic acid uptake, as well as any other electrical, chemical, physical (mechanical) and/or biological mechanism that results in the introduction of nucleic acid into the plant cell, including any combination thereof. Procedures for transforming both eukaryotic and prokaryotic organisms are well known and routine in the art and are described throughout the literature (See, for example, Jiang et al. 2013. *Nat. Biotechnol.* 31:233-239; Ran et al. *Nature Protocols* 8:2281-2308 (2013)). General guides to various plant transformation methods known in the art include Miki et al. ("Procedures for Introducing Foreign DNA into Plants" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B. R. and Thompson, J. E., Eds. (CRC Press, Inc., Boca Raton, 1993), pages 67-88) and Rakowoczy-Trojanowska (*Cell. Mol. Biol. Lett.* 7:849-858 (2002)).

In some embodiments of the invention, transformation of a cell may comprise nuclear transformation. In other embodiments, transformation of a cell may comprise plastid transformation (e.g., chloroplast transformation). In still further embodiments, nucleic acids of the invention may be introduced into a cell via conventional breeding techniques. In some embodiments, one or more of the polynucleotides, expression cassettes and/or vectors may be introduced into a plant cell via *Agrobacterium* transformation.

A polynucleotide therefore can be introduced into a plant, plant part, plant cell in any number of ways that are well known in the art. The methods of the invention do not depend on a particular method for introducing one or more nucleotide sequences into a plant, only that they gain access to the interior the cell. Where more than polynucleotide is to be introduced, they can be assembled as part of a single nucleic acid construct, or as separate nucleic acid constructs, and can be located on the same or different nucleic acid constructs. Accordingly, the polynucleotide can be introduced into the cell of interest in a single transformation event, or in separate transformation events, or, alternatively, a polynucleotide can be incorporated into a plant as part of a breeding protocol.

The present invention is directed to the modification of expression and protein production by genes that contribute to pod shattering in order to control seed dehiscence and improve yield and labor costs. The functionally redundant MADS-domain factors SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2) are required for both separation layer differentiation and to promote lignification of the lignified margin layer in Arabidopsis. Consequently, when *shp1 shp2* mutant fruit are mature, they fail to open, and the seeds are trapped inside. Accordingly, the present invention provides methods and compositions for modifying *SHATTERPROOF MADS-BOX (SHP)* genes in canola plants (e.g., an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene) to provide canola plants that exhibit reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin. In canola, *SHP* genes include *SHP1* (BnaA09g55330D), *SHP2* (BnaA07g18050D, BnaC06g16910D), *SHP3* (BnaA04g01810D, BnaC04g23360D), and/or *SHP4* (BnaA05g02990D) (Gene IDs from BrassicaEDB – a Gene Expression Database for *Brassica* Crops (brassica.biodb.org/analysis)), each of which may be targeted in a plant. Thus, an editing strategy useful for this invention can include generating a mutation in one or more than one *SHP* gene in a canola plant, e.g., a canola plant may comprise 1, 2, 3, 4, 5, and/or 6 or more *SHP* genes comprising a modification as described herein. In some embodiments, one or more than one mutation (optionally, a non-natural mutation) may be generated in a *SHP* gene of a plant. Mutations that may be useful for producing canola plants having reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin include, for example, substitutions, deletions, and/or insertions. In some aspects, a mutation generated by the editing technology can be a point mutation. In some embodiments, a mutation in one or more than one *SHP* gene as described herein results in knockdown of expression of the one or more than one *SHP* gene. In some embodiments, a mutation in one or more than one *SHP* gene as described herein results in production of a modified *SHP* polypeptide, optionally wherein the modified *SHP* polypeptide comprises a C-terminal truncation, optionally a truncation of about the last 65-80 consecutive amino acid residues (about 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or 80 consecutive amino acid residues) of the *SHP* polypeptide produced by the unmodified endogenous *SHP* gene.

In some embodiments, the invention provides a canola plant or plant part thereof comprising at least one mutation in at least one (e.g., one or more than one *SHP* gene) endogenous *SHATTERPROOF MADS-BOX (SHP)* gene encoding a Shatterproof MADS-box transcription factor (*SHP*) polypeptide, optionally wherein the at least one mutation may be a

non-natural mutation. In some embodiments, an endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, wherein the encoded SHP polypeptide is an SHP1 polypeptide, an SHP2 polypeptide, an SHP3 polypeptide, or an SHP4 polypeptide, respectively. An endogenous *SHP* gene may have the

5 gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*). In some embodiments, the at least one mutation in the canola plant may be a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation. In some embodiments, a mutation may be a knock-down mutation. As

10 used herein, a knock-down mutation results in a reduction in activity of at least 5% (e.g., a reduction in activity of about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or

15 100%, and any range or value therein).

In some embodiments, a canola plant cell is provided, the canola plant cell comprising an editing system, the editing system comprising: (a) a CRISPR-Cas effector protein; and (b) a guide nucleic acid (e.g., gRNA, gDNA, crRNA, crDNA, sgRNA, sgDNA) comprising a spacer sequence with complementarity to an endogenous target gene encoding a Shatterproof MADS-box transcription factor (SHP) polypeptide in the canola plant cell. The editing system may be

20 used to generate a mutation in the endogenous target gene encoding a SHP polypeptide. In some embodiments, the endogenous target gene is an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene (e.g., one or more than one endogenous *SHP* gene), optionally an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene,

25 and the SHP polypeptide is an SHP1 polypeptide, an SHP2 polypeptide, an SHP3 polypeptide, or an SHP4 polypeptide, respectively. In some embodiments, the mutation is a non-natural mutation. In some embodiments, the endogenous target gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence

30 identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99**,

145-147, 174-176, 203-205, 237-239 or 289-291. In some embodiments, a guide nucleic acid of an editing system may comprise the nucleotide sequence (a spacer sequence, e.g., one or more spacers) of any one of **SEQ ID NOs:292-297** (e.g., **SEQ ID NO:292** (PWsp236), **SEQ ID NO:293** (PWsp237), **SEQ ID NO:294** (PWsp238), **SEQ ID NO:295** (PWsp239), **SEQ ID NO:296** (PWsp240), and/or **SEQ ID NO:297** (PWsp241)) and/or **SEQ ID NOs:342-346** (e.g., **SEQ ID NO:342** (PWsp291), **SEQ ID NO:343** (PWsp292), **SEQ ID NO:344** (PWsp293), **SEQ ID NO:345** (PWsp294), and/or **SEQ ID NO:346** (PWsp294), or reverse complement thereof.

A mutation in an *SHP* gene of a canola plant, plant part thereof, or a canola plant cell useful for this invention may be any type of mutation, including a base substitution, a base deletion, and/or a base insertion. In some embodiments, the mutation may be a non-natural mutation. In some embodiments, the mutation may comprise a base substitution to an A, a T, a G, or a C. In some embodiments, the mutation may be a deletion (optionally, an out-of-frame deletion or an in-frame deletion) (e.g., of at least one base pair (e.g., 1 base pair to about 100 base pairs; e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 consecutive base pairs; e.g., 1 to about 50 consecutive base pairs, 1 to about 30 consecutive base pairs, 1 to about 15 consecutive base pairs) or an insertion of at least one base pair (e.g., 1 base pair to about 15 base pairs; e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 consecutive base pairs), optionally wherein the deletion is an out-of-frame deletion. For example, a deletion in an *SHP* gene may be about 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, or 45 consecutive base pairs, optionally about 7, 8, 10, 20, or 45 consecutive base pairs. In some embodiments, a non-natural mutation may be an insertion of at least one base pair (e.g., 1 base pair to about 100 consecutive base pairs; e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 consecutive base. In some embodiments, an insertion of 1 to about 100 base pairs is an out-of-frame insertion.

A mutation in an *SHP* gene may be located in the 3' region of the *SHP* gene (e.g., a *SHP1* gene, a *SHP2* gene, a *SHP3* gene, and/or a *SHP4* gene), optionally a non-natural mutation,

optionally in the 3' region of the *SHP* gene that encodes the C-terminal region of the encoded SHP polypeptide (e.g., the 3' coding regions (exons)). For example, a mutation in an *SHP* gene may be located (a) in the second to the last exon, (b) in the second to the last exon and the intron that is 3' to the second to the last exon and 5' to the last exon, and/or (c) in the last exon. In some embodiments, the mutation may be an out-of-frame deletion, an in-frame deletion, or an out-of-frame insertion. In some embodiments, the out-of-frame deletion, in-frame deletion, or out-of-frame insertion may result in a deletion of the last exon of the gene. In some embodiments, the out-of-frame deletion, in-frame deletion, or out-of-frame insertion results in a gene that encodes a truncated polypeptide, optionally a polypeptide having a C-terminal truncation resulting from a premature stop codon generated by the deletion or insertion. In some embodiments, the C-terminal truncation is a deletion of about 65 to about 80 consecutive amino acid residues (about 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or 80 consecutive amino acid residues) from the C-terminus of the SHP polypeptide (e.g., about the last 65 to 80 consecutive amino acid residues).

As used herein, a “non-natural mutation” refers to a mutation that is generated through human intervention and differs from mutations found in the same gene that have occurred in nature (e.g., occurred naturally).

In some embodiments, a mutation useful with this invention may be a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation.

The types of editing tools that may be used to generate these and other mutations in canola *SHP* genes include any base editors or cutters, which are guided to a target site using spacers having at least 80% (e.g., 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%, or any range or value therein), complementarity to a portion or a region of a *SHP* gene (e.g., one or more than one SHP gene, e.g., an SHP1 gene, an SHP2 gene, an SHP3 gene, and/or an SHP4 gene) as described herein.

In some embodiments, a mutation of a *SHP* gene is within a portion or region of the endogenous *SHP* gene having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally a portion or region of the endogenous *SHP* gene having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, 285, or 324-338**.

An endogenous *SHP* gene useful with this invention (e.g., an endogenous target gene) encodes a Shatterproof MADS-box transcription factor (SHP) polypeptide, and includes an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, or an endogenous *SHP4* gene, which encode an SHP1 polypeptide, an SHP2 polypeptide, an SHP3 polypeptide, or an SHP4 polypeptide, respectively. In some embodiments, an endogenous *SHP* gene (e.g., endogenous target gene) (1) may comprise a nucleic acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, (2) may comprise a region of a *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, (3) may encode a polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, and/or (4) may encode a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

In some embodiments, a canola plant comprising at least one (e.g., one or more, e.g., 1, 2, 3, 4, 5, or more) mutation in an endogenous *SHP* gene (in at least one endogenous *SHP* gene, e.g., in one or more *SHP* genes, e.g., *SHP1*, *SHP2*, *SHP3*, *SHP4*) exhibits reduced pod shattering (and/or reduced lignification (reduced lignin content) in the pod valve margin and/or increased harvestable seed) as compared to a canola plant devoid of the at least one mutation (e.g., an isogenic plant (e.g., wild type unedited plant or a null segregant)).

In some embodiments, a canola plant may be regenerated from a canola plant part and/or plant cell of the invention comprising a mutation in one or more than one endogenous *SHP* gene (an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene) as described herein, wherein the regenerated canola plant comprises the mutation in the one or more than one endogenous *SHP* gene and a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control canola plant devoid of the same mutation in the one or more than one *SHP* gene.

In some embodiments, a canola plant cell is provided, the canola plant cell comprising at least one (e.g., one or more) mutation (optionally a non-natural mutation) within an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the at least one mutation is a substitution, insertion, or deletion that is introduced using an editing system that comprises a nucleic acid binding domain that binds to a target site in the endogenous *SHP* gene. In some embodiments, the substitution, insertion, or deletion results in, for example, a premature stop codon. In some

embodiments, the substitution, insertion, or deletion results in, for example, a truncated SHP protein, optionally an SHP polypeptide having a C-terminal truncation. In some embodiments, the at least one mutation is a point mutation, optionally resulting in a premature stop codon, optionally a truncated SHP protein. In some embodiments, the at least one mutation within the

5 *SHP* gene is an insertion and/or a deletion, optionally the at least one mutation is an out-of-frame insertion or out-of-frame deletion. In some embodiments, the endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, or an endogenous *SHP4* gene.

In some embodiments, a target site in an *SHP* gene of a canola plant cell may be within a

10 region or portion of the endogenous *SHP* gene, the region having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257,**

15 **261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**. In some embodiments, the target site in the *SHP* gene is within a region of the endogenous *SHP* gene that encodes an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

In some embodiments, a mutation may be made following cleavage by an editing system

20 that comprises a nuclease and a nucleic acid binding domain that binds to a target site within: (a) a sequence having least 80% sequence identity to a sequence encoding of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, optionally within the 3' region of a sequence having least 80% sequence identity to a sequence encoding of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, optionally located (i) in the

25 second to the last exon, (ii) in the second to the last exon and the intron that is 3' to the second to the last exon and 5' to the last exon, and/or (iii) in the last exon of a sequence having least 80% sequence identity to a sequence encoding of any one of **SEQ ID NOs:69, 100, 148, 177, 206, or 240**, or (b) a sequence having at least 80% sequence identity to a sequence encoding any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally at

30 least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276,**

279, 280, 283, or 285, and the at least one mutation within an *SHP* gene is made following cleavage by the nuclease, optionally wherein the at least one mutation is a non-natural mutation.

In some embodiments, the at least one mutation may result in a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation.

5 In some embodiments, the canola plant cell may be regenerated into a canola plant that comprises the at least one mutation, optionally wherein the canola plant regenerated from the canola plant cell exhibits a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control plant devoid of the at least one mutation. In some embodiments, a canola plant comprising the at least one mutation in
10 an endogenous *SHP* gene is not regenerated.

In some embodiments, a method of producing/breeding a transgene-free edited canola plant is provided, the method comprising: crossing a canola plant of the present invention (e.g., a canola plant comprising one or more mutations (optionally, one or more non-natural mutations) in one or more *SHP* genes and having reduced pod shattering and/or reduced lignification
15 (reduced lignin content) in the pod valve margin) with a transgene free plant, thereby introducing the mutation into the canola plant that is transgene-free; and selecting a progeny canola plant that comprises the mutation and is transgene-free, thereby producing a transgene free edited canola plant.

Also provided herein is a method of providing a plurality of canola plants having
20 reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising planting two or more canola plants of the invention (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 1000, 2000, 3000, 400, 5000, or 10,000 or more canola plants comprising one or more mutations (optionally, one or more non-natural mutations) in one or more *SHP* genes and having reduced pod shattering and/or reduced
25 lignification (reduced lignin content) in the pod valve margin) in a growing area (e.g., a field (e.g., a cultivated field, an agricultural field), a growth chamber, a greenhouse, a recreational area, a lawn, and/or a roadside and the like), thereby providing a plurality of canola plants having reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a plurality of control canola plants devoid of the mutation.

30 In some embodiments, a method for editing a specific site in the genome of a canola plant cell is provided, the method comprising: cleaving, in a site-specific manner, a target site within an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene in the canola plant cell, the endogenous *SHP* gene: ((a) comprising a nucleotide sequence having at least 80% sequence

identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**,
 (b) comprising a region having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, (c) encoding an amino acid
 sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179,**
 5 **208, or 242**, (d) encoding a region having at least 80% sequence identity to an amino acid
 sequence of any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**,
 thereby generating an edit in the endogenous *SHP* gene of the canola plant cell and producing a
 canola plant cell comprising the edit in the endogenous *SHP* gene. In some embodiments, the
 endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous
 10 *SHP3* gene, or an endogenous *SHP4* gene, optionally wherein an edit is generated in two or
 more endogenous *SHP* genes (e.g., two or more of *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*).

In some embodiments, the edit in the endogenous *SHP* gene in a canola plant results in a
 mutation including, but not limited to, a base deletion, a base substitution, or a base insertion,
 optionally wherein the at least one mutation is a mutation. In some embodiments, the at least one
 15 mutation may be located in the 3' region of an *SHP* gene, for example, located in the second to
 the last exon, in the second to the last exon and the 3' adjacent intron, and/or in the last exon of
 an *SHP* genomic sequence. In some embodiments, the edit may result in at least one mutation
 that is an insertion of at least one base pair (e.g., 1 base pair to about 100 base pairs), optionally
 wherein the insertion is an out-of-frame insertion. In some embodiments, the edit may result in
 20 at least one mutation that is a deletion, optionally wherein the deletion is about 1 to about 100
 consecutive base pairs in length, e.g., about 1-50 consecutive base pairs, about 1-30 consecutive
 base pairs or about 1-15 consecutive base pairs in length, optionally about 1, 2, 3, 4, 5, 6, 7, 8, 9,
 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34,
 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 consecutive base pairs. A deletion
 25 or insertion useful with this invention may be an out-of-frame insertion or an out-of-frame
 deletion. In some embodiments, an out-of-frame insertion or out-of-frame deletion may result in
 a premature stop codon and truncated protein. In some embodiments, the edit in a *SHP* gene
 results in a truncated SHP polypeptide, optionally a C-terminal truncation of the SHP
 polypeptide, optionally wherein the C-terminal truncation is a deletion of about 65 to about 80
 30 consecutive amino acid residues (about 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79,
 or 80 consecutive amino acid residues) from the C-terminus of the SHP polypeptide (e.g., about
 the last 65 to 80 consecutive amino acid residues; at least all of the consecutive amino acid
 residues encoded by the last exon of the *SHP* genomic sequence, optionally all of the amino acid

residues encoded by the last exon of the *SHP* genomic sequence and at least a portion of the amino acid residues encoded by the second to the last exon of the *SHP* genomic sequence).

In some embodiments, a method of editing may further comprise regenerating a canola plant from the canola plant cell comprising the edit in the endogenous *SHP* gene, thereby
 5 producing a canola plant comprising the edit in its endogenous *SHP* gene (optionally in the 3' end of the *SHP* gene, optionally in the second to the last exon, the second to the last exon and the 3' adjacent intron, and/or in the last exon) and having a phenotype of reduce pod shattering when compared to a control canola plant that is devoid of the edit.

In some embodiments, a method for making a canola plant is provided, the method
 10 comprising (a) contacting a population of canola plant cells comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene with a nuclease linked to a nucleic acid binding domain (e.g., editing system) that binds to a sequence (i) having at least 80% sequence identity to a nucleotide sequence of any one **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, (ii) comprising a region having at least 80% identity to any one of **SEQ ID**
 15 **NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, and/or (iv) encoding a region having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**; (b) selecting a canola plant cell from the population of canola plant cells in which an endogenous *SHP* gene has been
 20 mutated, thereby producing a canola plant cell comprising a mutation in the endogenous *SHP* gene; and (c) growing the selected canola plant cell into a canola plant comprising a mutation in the endogenous *SHP* gene.

In some embodiments, a method for reducing pod shattering and/or reducing lignification (reduced lignin content) in the pod valve margin in a canola plant is provided, the
 25 method comprising (a) contacting a canola plant cell comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene with a nuclease targeting the endogenous *SHP* gene, wherein the nuclease is linked to a nucleic acid binding domain (e.g., editing system) that binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene: (i) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70,**
 30 **100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (ii) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (iv)

encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291** to produce a canola plant cell comprising a mutation in the endogenous *SHP* gene; and (b) growing the canola plant cell comprising a mutation in the endogenous *SHP* gene into a canola plant comprising the mutation in the endogenous *SHP* gene thereby producing a canola plant having a mutated endogenous *SHP* gene and reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

In some embodiments, a method for producing a canola plant or part thereof comprising at least one cell having a mutated endogenous *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., one or more mutated endogenous *SHP* gene), the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing the canola plant or part thereof comprising at least one cell having a mutation in the endogenous *SHP* gene (e.g., a mutation in one or more endogenous *SHP* genes).

Also provided herein is a method for producing a canola plant or part thereof comprising a mutated endogenous *SHATTERPROOF MADS-BOX (SHP)* gene and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of

SEQ ID NOs:71, 102, 150, 179, 208, or 242; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing the canola plant or part thereof comprising an endogenous *SHP* gene having a mutation (e.g., at least one endogenous *SHP* gene having a mutation) and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

In some embodiments, a nuclease may cleave an endogenous *SHP* gene, thereby introducing the mutation into the endogenous *SHP* gene. A nuclease useful with the invention may be any nuclease that can be utilized to edit/modify a target nucleic acid. Such nucleases include, but are not limited to a zinc finger nuclease, transcription activator-like effector nucleases (TALEN), endonuclease (e.g., FokI) and/or a CRISPR-Cas effector protein. Likewise, any nucleic acid binding domain useful with the invention may be any DNA binding domain or RNA binding domain that can be utilized to edit/modify a target nucleic acid. Such nucleic acid binding domains include, but are not limited to, a zinc finger, transcription activator-like DNA binding domain (TAL), an argonaute and/or a CRISPR-Cas effector DNA binding domain.

In some embodiments, a nucleic acid binding domain (e.g., DNA binding domain) is comprised in a nucleic acid binding polypeptide. A "nucleic acid binding protein" or "nucleic acid binding polypeptide" as used herein refers to a polypeptide that binds and/or is capable of binding a nucleic acid in a site- and/or sequence-specific manner. In some embodiments, a nucleic acid binding polypeptide may be a sequence-specific nucleic acid binding polypeptide (e.g., a sequence-specific DNA binding domain) such as, but not limited to, a sequence-specific binding polypeptide and/or domain from, for example, a polynucleotide-guided endonuclease, a CRISPR-Cas effector protein (e.g., a CRISPR-Cas endonuclease), a zinc finger nuclease, a transcription activator-like effector nuclease (TALEN) and/or an Argonaute protein. In some embodiments, a nucleic acid binding polypeptide comprises a cleavage polypeptide (e.g., a nuclease polypeptide and/or domain) such as, but not limited to, an endonuclease (e.g., FokI), a polynucleotide-guided endonuclease, a CRISPR-Cas endonuclease, a zinc finger nuclease, and/or a transcription activator-like effector nuclease (TALEN). In some embodiments, the nucleic acid binding polypeptide associates with and/or is capable of associating with (e.g., forms a complex with) one or more nucleic acid molecule(s) (e.g., forms a complex with a guide nucleic acid as described herein) that can direct or guide the nucleic acid binding polypeptide to a specific target nucleotide sequence (e.g., a gene locus of a genome) that is complementary to

the one or more nucleic acid molecule(s) (or a portion or region thereof), thereby causing the nucleic acid binding polypeptide to bind to the nucleotide sequence at the specific target site. In some embodiments, the nucleic acid binding polypeptide is a CRISPR-Cas effector protein as described herein. In some embodiments, reference is made to specifically to a CRISPR-Cas effector protein for simplicity, but a nucleic acid binding polypeptide as described herein may be used. In some embodiments, a polynucleotide and/or a nucleic acid construct of the invention can be an “expression cassette” or can be comprised within an expression cassette.

In some embodiments, a method of editing an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) in a canola plant or plant part is provided, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a cytosine base editing system comprising a cytosine deaminase and a nucleic acid binding domain that binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby editing the endogenous *SHP* gene in the canola plant or part thereof and producing a canola plant or part thereof comprising at least one cell having a mutation in the endogenous *SHP* gene.

In some embodiments, a method of editing an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) in a canola plant or plant part is provided, the method comprising contacting a target site in an *SHP* gene in the canola plant or plant part with an adenosine base editing system comprising an adenosine deaminase and a nucleic acid binding domain that binds to a target site in the *SHP* gene, wherein the *SHP* gene (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to

any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239** or **289-291**, thereby editing the endogenous *SHP* gene in the canola plant or part thereof and producing a canola plant or part thereof comprising at least one cell having a mutation in the endogenous *SHP* gene.

In some embodiments, a method of creating a mutation in a *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) in a canola plant is provided, comprising(a) targeting a gene editing system to a portion of the endogenous *SHP* gene that (i) comprises a sequence having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288** or **324-338**; and/or (ii) encodes a sequence having at least 80% identity to any one **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288** or **324-338**, and (b) selecting a canola plant that comprises a modified nucleic acid sequence in a region having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**. In some embodiments, the modification is a deletion or an insertion. In some embodiments, the modification is an out-of-frame deletion or out-of-frame insertion resulting in a truncated Shatterproof MADS-box transcription factor (SHP) polypeptide.

In some embodiments, a mutation provided by methods of the invention may a non-natural mutation. In some embodiments, the mutation may be a substitution, an insertion and/or a deletion, optionally wherein the insertion or deletion is an out-of-frame insertion or an out-of-frame deletion. In some embodiments, a mutation may be a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation. In some embodiments, a mutation may comprise a base substitution to an A, a T, a G, or a C. In some embodiments, the mutation may be a deletion (e.g., out-of-frame deletion) of about 1 base pair to about 100 consecutive base pairs, optionally, 1 to about 50 consecutive base pairs, 1 to about 30 consecutive base pairs, 1 to about 15 consecutive base pairs. In some embodiments, the mutation may be an insertion (e.g., an out-of-frame insertion) of at least one base pair (e.g., 1 base pair to about 100 consecutive base pairs). A mutation in a *SHP* gene may be located in the 3' region of the *SHP* gene, optionally wherein the mutation may be within a portion or region of the endogenous *SHP* gene that encodes the SHP polypeptide (e.g., the coding regions (exons), e.g., second to last exon and/or last exon). In some embodiments, the mutation in a *SHP* gene may be located in the intron located between the second to last exon and the last exon of the *SHP* gene. In some embodiments, the mutation may be located in a region of a *SHP* gene that

bridges between the second to last exon and the intron located between the second to last exon and the last exon of the *SHP* gene (e.g., the intron located immediately 3' to the second to the last exon). A mutation in an *SHP* gene can result in a polypeptide having a deletion of the amino acids encoded by the last exon, optionally a deletion of the amino acids encoded by the last exon and at least one amino acid (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 amino acids) encoded by the second to the last exon of the *SHP* gene. In some embodiments, a mutation of a *SHP* gene that is an out-of-frame deletion or an out-of-frame insertion may result in a premature stop codon and a truncated SHP polypeptide. In some embodiments, the out-of-frame deletion or out-of-frame insertion may be a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation.

In some embodiments, a method of detecting a mutant *SHATTERPROOF MADS-BOX* (*SHP*) gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) in a canola plant is provided, the method comprising detecting in the genome of a canola plant an endogenous *SHP* gene encoding a truncated SHP polypeptide, optionally wherein the mutation is located in the 3' region (optionally, second to last exon and/or last exon and/or the intron located between the second to last exon and the last exon) of the *SHP* gene, the 3' region having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**. In some embodiments, the mutation that is detected is an out-of-frame deletion or an out-of-frame insertion.

In some embodiments, the present invention provides a method of producing a canola plant comprising a mutation in an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) and at least one polynucleotide of interest, the method comprising crossing a canola plant of the invention comprising at least one mutation in an endogenous *SHP* gene (a first canola plant) with a second canola plant that comprises the at least one polynucleotide of interest to produce progeny canola plants; and selecting progeny canola plants comprising at least one mutation in the *SHP* gene and the at least one polynucleotide of interest, thereby producing the canola plant comprising a mutation in an endogenous *SHP* gene and at least one polynucleotide of interest.

The present invention further provides a method of producing a canola plant comprising a mutation in an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene (e.g., *SHP1*, *SHP2*,

SHP3, and/or *SHP4*) and at least one polynucleotide of interest, the method comprising introducing at least one polynucleotide of interest into a canola plant of the present invention comprising at least one mutation in a *SHP* gene, thereby producing a canola plant comprising at least one mutation in a *SHP* gene and at least one polynucleotide of interest.

5 In some embodiments, also provided is a method of producing a canola plant comprising a mutation in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) and exhibiting a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising crossing a first canola plant, which is a canola plant of the present invention comprising at least one
10 mutation in a *SHP* gene, with a second canola plant that exhibits a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin; and selecting progeny canola plants comprising the mutation in the *SHP* gene and a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, thereby producing the canola plant comprising a mutation in an endogenous *SHP* gene
15 and exhibiting a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control plant.

Further provided is a method of controlling weeds in a container (e.g., pot, or seed tray and the like), a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, the method comprising applying an herbicide to one or more (a plurality) canola plants
20 of the invention (e.g., a canola plant comprising at least one mutation in a *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) as described herein) growing in a container, a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, thereby controlling the weeds in the container, the growth chamber, the greenhouse, the field, the recreational area, the lawn, or on the roadside in which the one or more canola
25 plants are growing.

In some embodiments, a method of reducing insect predation on a canola plant is provided, the method comprising applying an insecticide to one or more canola plants of the invention, optionally, wherein the one or more canola plants are growing in a container, a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, thereby
30 reducing insect predation on the one or more canola plants.

In some embodiments, a method of reducing fungal disease on a canola plant is provided, the method comprising applying a fungicide to one or more canola plants of the invention, optionally, wherein the one or more canola plants are growing in a container, a

growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, thereby reducing fungal disease on the one or more canola plants.

Example endogenous SHP genes and encoded SHP polypeptides useful with this invention, as well as target regions for editing and example edited SHP genes and encoded polypeptides are provided in **Table 1**.

Table 1.

Endogenous SHP gene	Gene ID No.	Genomic SEQ ID NOs	Coding and polypeptide SEQ ID NOs	Genomic Regions	Poly-peptide regions	SEQ ID NOs for example edited genes and mutated protein
SHP1A	BnaA09g55330D	177	178, 179	180-202	203-205	
SHP2A	BnaA07g18050D	100	101, 102	103-144	145-147	Gene: 298, 300, 302, 306, 312, 314, 321 Protein: 299, 301, 303, 307, 313, 315
SHP2C	BnaC06g16910D	240	241, 242	243-288	289-291	Gene: 304, 308, 310, 316, 318, 319 Protein: 305, 309, 311, 317 Deleted portion of gene: 320
SHP3A	BnaA04g01810D	69	70, 71	72-96	97-99	
SHP3C	BnaC04g23360D	206	207, 208	209-236	237-239	
SHP4A	BnaA05g02990D	148	149, 150	151-173, 324-338	174-176	Gene: 322, 323

A polynucleotide of interest may be any polynucleotide that can confer a desirable phenotype or otherwise modify the phenotype or genotype of a plant. In some embodiments, a polynucleotide of interest may be polynucleotide that confers herbicide tolerance, insect resistance, nematode resistance, disease resistance, increased yield, increased nutrient use efficiency or abiotic stress resistance.

Thus, plants or plant cultivars which are to be treated with preference in accordance with the invention include all plants which, through genetic modification, received genetic material which imparts particular advantageous useful properties ("traits") to these plants. Examples of such properties are better plant growth, vigor, stress tolerance, standability, lodging resistance, nutrient uptake, plant nutrition, and/or yield, in particular improved growth, increased tolerance to high or low temperatures, increased tolerance to drought or to levels of water or soil salinity, enhanced flowering performance, easier harvesting, accelerated ripening, higher yields, higher

quality and/or a higher nutritional value of the harvested products, better storage life and/or processability of the harvested products.

Further examples of such properties are an increased resistance against animal and microbial pests, such as against insects, arachnids, nematodes, mites, slugs and snails owing, for example, to toxins formed in the plants. Among DNA sequences encoding proteins which confer properties of tolerance to such animal and microbial pests, in particular insects, mention will particularly be made of the genetic material from *Bacillus thuringiensis* encoding the Bt proteins widely described in the literature and well known to those skilled in the art. Mention will also be made of proteins extracted from bacteria such as *Photorhabdus* (WO97/17432 and WO98/08932). In particular, mention will be made of the Bt Cry or VIP proteins which include the CryIA, CryIAb, CryIAc, CryIIA, CryIIIA, CryIIIB2, Cry9c Cry2Ab, Cry3Bb and CryIF proteins or toxic fragments thereof and also hybrids or combinations thereof, especially the CryIF protein or hybrids derived from a CryIF protein (e.g. hybrid CryIA-CryIF proteins or toxic fragments thereof), the CryIA-type proteins or toxic fragments thereof, preferably the CryIAc protein or hybrids derived from the CryIAc protein (e.g. hybrid CryIAb-CryIAc proteins) or the CryIAb or Bt2 protein or toxic fragments thereof, the Cry2Ae, Cry2Af or Cry2Ag proteins or toxic fragments thereof, the CryIA.105 protein or a toxic fragment thereof, the VIP3Aa19 protein, the VIP3Aa20 protein, the VIP3A proteins produced in the COT202 or COT203 cotton events, the VIP3Aa protein or a toxic fragment thereof as described in Estruch et al. (1996), Proc Natl Acad Sci US A. 28;93(11):5389-94, the Cry proteins as described in WO2001/47952, the insecticidal proteins from *Xenorhabdus* (as described in WO98/50427), Serratia (particularly from *S. entomophila*) or *Photorhabdus* species strains, such as Tc-proteins from *Photorhabdus* as described in WO98/08932. Also any variants or mutants of any one of these proteins differing in some amino acids (1-10, preferably 1-5) from any of the above named sequences, particularly the sequence of their toxic fragment, or which are fused to a transit peptide, such as a plastid transit peptide, or another protein or peptide, is included herein.

Another and particularly emphasized example of such properties is conferred tolerance to one or more herbicides, for example imidazolinones, sulphonylureas, glyphosate or phosphinothricin. Among DNA sequences encoding proteins (i.e., polynucleotides of interest) which confer properties of tolerance to certain herbicides on the transformed plant cells and plants, mention will be particularly be made to the bar or PAT gene or the *Streptomyces coelicolor* gene described in WO2009/152359 which confers tolerance to glufosinate herbicides, a gene encoding a suitable EPSPS (5-Enolpyruvylshikimat-3-phosphat-Synthase) which confers

tolerance to herbicides having EPSPS as a target, especially herbicides such as glyphosate and its salts, a gene encoding glyphosate-n-acetyltransferase, or a gene encoding glyphosate oxidoreductase. Further suitable herbicide tolerance traits include at least one ALS (acetolactate synthase) inhibitor (e.g., WO2007/024782), a mutated Arabidopsis ALS/AHAS gene (e.g., U.S. Patent 6,855,533), genes encoding 2,4-D-monooxygenases conferring tolerance to 2,4-D (2,4-dichlorophenoxyacetic acid) and genes encoding Dicamba monooxygenases conferring tolerance to dicamba (3,6-dichloro-2-methoxybenzoic acid).

Further examples of such properties are increased resistance against phytopathogenic fungi, bacteria and/or viruses owing, for example, to systemic acquired resistance (SAR), systemin, phytoalexins, elicitors and also resistance genes and correspondingly expressed proteins and toxins.

Particularly useful transgenic events in transgenic plants or plant cultivars which can be treated with preference in accordance with the invention include Event 531/ PV-GHBK04 (cotton, insect control, described in WO2002/040677), Event 1143-14A (cotton, insect control, not deposited, described in WO2006/128569); Event 1143-51B (cotton, insect control, not deposited, described in WO2006/128570); Event 1445 (cotton, herbicide tolerance, not deposited, described in US-A 2002-120964 or WO2002/034946); Event 17053 (rice, herbicide tolerance, deposited as PTA-9843, described in WO2010/117737); Event 17314 (rice, herbicide tolerance, deposited as PTA-9844, described in WO2010/117735); Event 281-24-236 (cotton, insect control - herbicide tolerance, deposited as PTA-6233, described in WO2005/103266 or US-A 2005-216969); Event 3006-210-23 (cotton, insect control - herbicide tolerance, deposited as PTA-6233, described in US-A 2007-143876 or WO2005/103266); Event 3272 (corn, quality trait, deposited as PTA-9972, described in WO2006/098952 or US-A 2006-230473); Event 33391 (wheat, herbicide tolerance, deposited as PTA-2347, described in WO2002/027004), Event 40416 (corn, insect control - herbicide tolerance, deposited as ATCC PTA-11508, described in WO 11/075593); Event 43A47 (corn, insect control - herbicide tolerance, deposited as ATCC PTA-11509, described in WO2011/075595); Event 5307 (corn, insect control, deposited as ATCC PTA-9561, described in WO2010/077816); Event ASR-368 (bent grass, herbicide tolerance, deposited as ATCC PTA-4816, described in US-A 2006-162007 or WO2004/053062); Event B16 (corn, herbicide tolerance, not deposited, described in US-A 2003-126634); Event BPS-CV127- 9 (soybean, herbicide tolerance, deposited as NCIMB No. 41603, described in WO2010/080829); Event BLRI (oilseed rape, restoration of male sterility, deposited as NCIMB 41193, described in WO2005/074671), Event CE43-67B (cotton, insect

control, deposited as DSM ACC2724, described in US-A 2009-217423 or WO2006/128573);
 Event CE44-69D (cotton, insect control, not deposited, described in US-A 2010- 0024077);
 Event CE44-69D (cotton, insect control, not deposited, described in WO2006/128571); Event
 CE46-02A (cotton, insect control, not deposited, described in WO2006/128572); Event COT102
 5 (cotton, insect control, not deposited, described in US-A 2006-130175 or WO2004/039986);
 Event COT202 (cotton, insect control, not deposited, described in US-A 2007-067868 or
 WO2005/054479); Event COT203 (cotton, insect control, not deposited, described in
 WO2005/054480);); Event DAS21606-3 / 1606 (soybean, herbicide tolerance, deposited as
 PTA-11028, described in WO2012/033794), Event DAS40278 (corn, herbicide tolerance,
 10 deposited as ATCC PTA-10244, described in WO2011/022469); Event DAS-44406-6 /
 pDAB8264.44.06.1 (soybean, herbicide tolerance, deposited as PTA-11336, described in
 WO2012/075426), Event DAS-14536-7 /pDAB8291.45.36.2 (soybean, herbicide tolerance,
 deposited as PTA-11335, described in WO2012/075429), Event DAS-59122-7 (corn, insect
 control - herbicide tolerance, deposited as ATCC PTA 11384, described in US-A 2006-070139);
 15 Event DAS-59132 (corn, insect control - herbicide tolerance, not deposited, described in
 WO2009/100188); Event DAS68416 (soybean, herbicide tolerance, deposited as ATCC PTA-
 10442, described in WO2011/066384 or WO2011/066360); Event DP-098140-6 (corn, herbicide
 tolerance, deposited as ATCC PTA-8296, described in US-A 2009- 137395 or WO 08/112019);
 Event DP-305423-1 (soybean, quality trait, not deposited, described in US-A 2008-312082 or
 20 WO2008/054747); Event DP-32138-1 (corn, hybridization system, deposited as ATCC PTA-
 9158, described in US-A 2009-0210970 or WO2009/103049); Event DP-356043-5 (soybean,
 herbicide tolerance, deposited as ATCC PTA-8287, described in US-A 2010-0184079 or
 WO2008/002872); Event EE-I (brinjal, insect control, not deposited, described in WO
 07/091277); Event Fil 17 (corn, herbicide tolerance, deposited as ATCC 209031, described in
 25 US-A 2006-059581 or WO 98/044140); Event FG72 (soybean, herbicide tolerance, deposited as
 PTA-11041, described in WO2011/063413), Event GA21 (corn, herbicide tolerance, deposited
 as ATCC 209033, described in US-A 2005-086719 or WO 98/044140); Event GG25 (corn,
 herbicide tolerance, deposited as ATCC 209032, described in US-A 2005-188434 or
 WO98/044140); Event GHB119 (cotton, insect control - herbicide tolerance, deposited as ATCC
 30 PTA-8398, described in WO2008/151780); Event GHB614 (cotton, herbicide tolerance,
 deposited as ATCC PTA-6878, described in US-A 2010-050282 or W02007/017186); Event
 GJ11 (corn, herbicide tolerance, deposited as ATCC 209030, described in US-A 2005-188434 or
 WO98/044140); Event GM RZ13 (sugar beet, virus resistance, deposited as NCIMB-41601,

described in WO2010/076212); Event H7-l (sugar beet, herbicide tolerance, deposited as NCIMB 41158 or NCIMB 41159, described in US-A 2004-172669 or WO 2004/074492); Event JOPLINI (wheat, disease tolerance, not deposited, described in US-A 2008-064032); Event LL27 (soybean, herbicide tolerance, deposited as NCIMB41658, described in WO2006/108674 or US-A 2008-320616); Event LL55 (soybean, herbicide tolerance, deposited as NCIMB 41660, described in WO 2006/108675 or US-A 2008-196127); Event LLCotton25 (cotton, herbicide tolerance, deposited as ATCC PTA-3343, described in WO2003/013224 or US- A 2003-097687); Event LLRICE06 (rice, herbicide tolerance, deposited as ATCC 203353, described in US 6,468,747 or WO2000/026345); Event LLRice62 (rice, herbicide tolerance, deposited as ATCC 203352, described in WO2000/026345), Event LLRICE601 (rice, herbicide tolerance, deposited as ATCC PTA-2600, described in US-A 2008-2289060 or WO2000/026356); Event LY038 (corn, quality trait, deposited as ATCC PTA-5623, described in US-A 2007-028322 or WO2005/061720); Event MIR162 (corn, insect control, deposited as PTA-8166, described in US-A 2009-300784 or WO2007/142840); Event MIR604 (corn, insect control, not deposited, described in US-A 2008-167456 or WO2005/103301); Event MON15985 (cotton, insect control, deposited as ATCC PTA-2516, described in US-A 2004-250317 or WO2002/100163); Event MON810 (corn, insect control, not deposited, described in US-A 2002-102582); Event MON863 (corn, insect control, deposited as ATCC PTA-2605, described in WO2004/011601 or US-A 2006-095986); Event MON87427 (corn, pollination control, deposited as ATCC PTA-7899, described in WO2011/062904); Event MON87460 (corn, stress tolerance, deposited as ATCC PTA-8910, described in WO2009/111263 or US-A 2011-0138504); Event MON87701 (soybean, insect control, deposited as ATCC PTA- 8194, described in US-A 2009-130071 or WO2009/064652); Event MON87705 (soybean, quality trait - herbicide tolerance, deposited as ATCC PTA-9241, described in US-A 2010-0080887 or WO2010/037016); Event MON87708 (soybean, herbicide tolerance, deposited as ATCC PTA-9670, described in WO2011/034704); Event MON87712 (soybean, yield, deposited as PTA-10296, described in WO2012/051199), Event MON87754 (soybean, quality trait, deposited as ATCC PTA-9385, described in WO2010/024976); Event MON87769 (soybean, quality trait, deposited as ATCC PTA- 8911, described in US-A 2011-0067141 or WO2009/102873); Event MON88017 (corn, insect control - herbicide tolerance, deposited as ATCC PTA-5582, described in US-A 2008-028482 or WO2005/059103); Event MON88913 (cotton, herbicide tolerance, deposited as ATCC PTA-4854, described in WO2004/072235 or US-A 2006-059590); Event MON88302 (oilseed rape, herbicide tolerance, deposited as PTA-10955, described in WO2011/153186), Event

MON88701 (cotton, herbicide tolerance, deposited as PTA-11754, described in WO2012/134808), Event MON89034 (corn, insect control, deposited as ATCC PTA-7455, described in WO 07/140256 or US-A 2008-260932); Event MON89788 (soybean, herbicide tolerance, deposited as ATCC PTA-6708, described in US-A 2006-282915 or

5 WO2006/130436); Event MSI 1 (oilseed rape, pollination control - herbicide tolerance, deposited as ATCC PTA-850 or PTA-2485, described in WO2001/031042); Event MS8 (oilseed rape, pollination control - herbicide tolerance, deposited as ATCC PTA-730, described in WO2001/041558 or US-A 2003-188347); Event NK603 (corn, herbicide tolerance, deposited as ATCC PTA-2478, described in US-A 2007-292854); Event PE-7 (rice, insect control, not

10 deposited, described in WO2008/114282); Event RF3 (oilseed rape, pollination control - herbicide tolerance, deposited as ATCC PTA-730, described in WO2001/041558 or US-A 2003-188347); Event RT73 (oilseed rape, herbicide tolerance, not deposited, described in WO2002/036831 or US-A 2008-070260); Event SYHT0H2 / SYN-000H2-5 (soybean, herbicide tolerance, deposited as PTA-11226, described in WO2012/082548), Event T227-1 (sugar beet,

15 herbicide tolerance, not deposited, described in WO2002/44407 or US-A 2009-265817); Event T25 (corn, herbicide tolerance, not deposited, described in US-A 2001-029014 or WO2001/051654); Event T304-40 (cotton, insect control - herbicide tolerance, deposited as ATCC PTA-8171, described in US-A 2010-077501 or WO2008/122406); Event T342-142 (cotton, insect control, not deposited, described in WO2006/128568); Event TC1507 (corn,

20 insect control - herbicide tolerance, not deposited, described in US-A 2005-039226 or WO2004/099447); Event VIP1034 (corn, insect control - herbicide tolerance, deposited as ATCC PTA-3925, described in WO2003/052073), Event 32316 (corn, insect control-herbicide tolerance, deposited as PTA-11507, described in WO2011/084632), Event 4114 (corn, insect control-herbicide tolerance, deposited as PTA-11506, described in WO2011/084621), event EE-

25 GM3 / FG72 (soybean, herbicide tolerance, ATCC Accession N° PTA-11041) optionally stacked with event EE-GM1/LL27 or event EE-GM2/LL55 (WO2011/063413A2), event DAS-68416-4 (soybean, herbicide tolerance, ATCC Accession N° PTA-10442, WO2011/066360A1), event DAS-68416-4 (soybean, herbicide tolerance, ATCC Accession N° PTA-10442, WO2011/066384A1), event DP-040416-8 (corn, insect control, ATCC Accession N° PTA-

30 11508, WO2011/075593A1), event DP-043A47-3 (corn, insect control, ATCC Accession N° PTA-11509, WO2011/075595A1), event DP- 004114-3 (corn, insect control, ATCC Accession N° PTA-11506, WO2011/084621A1), event DP-032316-8 (corn, insect control, ATCC Accession N° PTA-11507, WO2011/084632A1), event MON-88302-9 (oilseed rape, herbicide

tolerance, ATCC Accession N° PTA-10955, WO2011/153186A1), event DAS-21606-3 (soybean, herbicide tolerance, ATCC Accession No. PTA-11028, WO2012/033794A2), event MON-87712-4 (soybean, quality trait, ATCC Accession N°. PTA-10296, WO2012/051199A2), event DAS-44406-6 (soybean, stacked herbicide tolerance, ATCC Accession N°. PTA-11336, WO2012/075426A1), event DAS-14536-7 (soybean, stacked herbicide tolerance, ATCC Accession N°. PTA-11335, WO2012/075429A1), event SYN-000H2-5 (soybean, herbicide tolerance, ATCC Accession N°. PTA-11226, WO2012/082548A2), event DP-061061-7 (oilseed rape, herbicide tolerance, no deposit N° available, WO2012071039A1), event DP-073496-4 (oilseed rape, herbicide tolerance, no deposit N° available, US2012131692), event 8264.44.06.1 (soybean, stacked herbicide tolerance, Accession N° PTA-11336, WO2012075426A2), event 8291.45.36.2 (soybean, stacked herbicide tolerance, Accession N°. PTA-11335, WO2012075429A2), event SYHT0H2 (soybean, ATCC Accession N°. PTA-11226, WO2012/082548A2), event MON88701 (cotton, ATCC Accession N° PTA-11754, WO2012/134808A1), event KK179-2 (alfalfa, ATCC Accession N° PTA-11833, WO2013/003558A1), event pDAB8264.42.32.1 (soybean, stacked herbicide tolerance, ATCC Accession N° PTA-11993, WO2013/010094A1), event MZDT09Y (corn, ATCC Accession N° PTA-13025, WO2013/012775A1).

The genes/events (e.g., polynucleotides of interest), which impart the desired traits in question, may also be present in combinations with one another in the transgenic plants.

Examples of transgenic plants which may be mentioned are the important crop plants, such as cereals (wheat, rice, triticale, barley, rye, oats), maize, soya beans, potatoes, sugar beet, sugar cane, tomatoes, peas and other types of vegetable, cotton, tobacco, oilseed rape and also fruit plants (with the fruits apples, pears, citrus fruits and grapes), with particular emphasis being given to maize, soya beans, wheat, rice, potatoes, cotton, sugar cane, tobacco and oilseed rape. Traits which are particularly emphasized are the increased resistance of the plants to insects, arachnids, nematodes and slugs and snails, as well as the increased resistance of the plants to one or more herbicides.

Commercially available examples of such plants, plant parts or plant seeds that may be treated with preference in accordance with the invention include commercial products, such as plant seeds, sold or distributed under the GENUITY®, DROUGHTGARD®, SMARTSTAX®, RIB COMPLETE®, ROUNDUP READY®, VT DOUBLE PRO®, VT TRIPLE PRO®, BOLLGARD II®, ROUNDUP READY 2 YIELD®, YIELDGARD®, ROUNDUP READY® 2 XTEND™, INTACTA RR2 PRO®, VISTIVE GOLD®, and/or XTENDFLEX™ trade names.

A *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHP1*, *SHP2*, *SHP3*, and/or *SHP4*) useful with this invention includes any canola *SHP* gene in which a mutation as described herein can confer reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin in a canola plant or part thereof comprising the mutation. In some

5 embodiments, an endogenous *SHP* gene (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of

10 **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

In some embodiments, the at least one mutation in an endogenous *SHP* gene in a canola plant may be a base substitution, a base deletion and/or a base insertion, optionally wherein the

15 at least one mutation may be a non-natural mutation. In some embodiments, the at least one mutation in an endogenous *SHP* gene in a canola plant may result in a canola plant having the phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control plant devoid of the edit/mutation. A canola plant of the invention having the phenotype of reduced pod shattering and/or reduced lignification

20 (reduced lignin content) in the pod valve margin may also exhibit/provide an increase in harvestable seed.

In some embodiments, a mutation in an endogenous *SHP* gene may be a base substitution, a base deletion and/or a base insertion of at least 1 base pair. In some

25 embodiments, a base deletion may be 1 nucleotide to about 100 nucleotides (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 base pairs, or any range or value therein, e.g., 1 to about 50 base pairs, 1 to about 30 base pairs, 1 to about 15 base pairs, or

30 any range or value therein), optionally where the mutation is at about 2 to about 100 consecutive nucleotides (e.g., 1 to about 50 consecutive base pairs, 1 to about 30 consecutive base pairs, 1 to about 15 consecutive base pairs). In some embodiments, a mutation in an endogenous *SHP* gene may be a base insertion of 1 to about 100 consecutive nucleotides of the *SHP* nucleic acid. In

some embodiments, a mutation in an endogenous *SHP* gene may be an out-of-frame insertion or an out-of-frame deletion that results in an SHP protein having a C-terminal truncation. In some embodiments, the at least one mutation may be a base substitution, optionally a substitution to an A, a T, a G, or a C. A mutation useful with this invention may be a point mutation. In some
 5 embodiments, the mutation may be a non-natural mutation.

In some embodiments, a mutation in an endogenous *SHP* gene may be made following cleavage by an editing system that comprises a nuclease and a nucleic acid binding domain that binds to a target site within a target nucleic acid (e.g., an endogenous *SHP* gene, e.g., *SHPI*, *SHPI*, *SHPI*, and/or *SHPI*), the target nucleic acid comprising a sequence having at least 80%
 10 sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, and/or encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, optionally wherein the target site is located in a region of the *SHP* gene: the region comprising a sequence having at least 80% identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202,**
 15 **209-236, 243-288 or 324-338** and/or encoding a sequence having at least 80% sequence identity to an amino acid sequence of any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

Further provided are guide nucleic acids (e.g., gRNA, gDNA, crRNA, crDNA) that bind to a target site in a *SHATTERPROOF MADS-BOX (SHP)* gene (e.g., *SHPI*, *SHPI*, *SHPI*, and/or
 20 *SHPI*), wherein the target site is in a region of the *SHP* gene having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275,**
 25 **276, 279, 280, 283, or 285**. In some embodiments, the guide nucleic acid comprises a spacer comprising any one of the nucleotide sequences of **SEQ ID NOs:292-297** and/or **SEQ ID NOs:342-346**.

In some embodiments, a canola plant or plant part thereof is provided that comprises at least one mutation in at least one endogenous *SHATTERPROOF MADS-BOX (SHP)* gene
 30 having a gene identification number (gene ID) of BnaA04g01810D (*SHPI*), BnaA07g18050D (*SHPI*), BnaA05g02990D (*SHPI*), BnaA09g55330D (*SHPI*), BnaC04g23360D (*SHPI*), and/or BnaC06g16910D (*SHPI*).

In some embodiments, a guide nucleic acid is provided that binds to a target nucleic acid in a *SHATTERPROOF MADS-BOX (SHP)* gene having a gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*).

5 In some embodiments, a system is provided comprising a guide nucleic acid comprising a spacer (e.g., one or more spacers) having the nucleotide sequence of any one of **SEQ ID NOs:292-297** and/or **SEQ ID NOs:342-346**, and a CRISPR-Cas effector protein that associates with the guide nucleic acid. In some embodiments, the system may further comprise a tracr nucleic acid that associates with the guide nucleic acid and a CRISPR-Cas effector protein,
10 optionally wherein the tracr nucleic acid and the guide nucleic acid are covalently linked.

As used herein, "a CRISPR-Cas effector protein in association with a guide nucleic acid" refers to the complex that is formed between a CRISPR-Cas effector protein and a guide nucleic acid in order to direct the CRISPR-Cas effector protein to a target site in a gene.

The invention further provides a gene editing system comprising a CRISPR-Cas effector
15 protein in association with a guide nucleic acid and the guide nucleic acid comprises a spacer sequence that binds to a *SHATTERPROOF MADS-BOX (SHP)* gene, optionally wherein the *SHP* gene (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240** or **241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID
20 NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288** or **324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239** or **289-291**. In some embodiments, a spacer sequence of the guide nucleic acid may comprise the
25 nucleotide sequence of any of **SEQ ID NOs:292-297** and/or **SEQ ID NOs:342-346**. In some embodiments, the gene editing system may further comprise a tracr nucleic acid that associates with the guide nucleic acid and a CRISPR-Cas effector protein, optionally wherein the tracr nucleic acid and the guide nucleic acid are covalently linked.

The present invention further provides a complex comprising a CRISPR-Cas effector
30 protein comprising a cleavage domain and a guide nucleic acid, wherein the guide nucleic acid binds to a target site in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene in canola, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207,**

240 or **241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288** or **324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide
 5 having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239** or **289-291**, and the cleavage domain cleaves a target strand in the *SHP* gene.

In some embodiments, an expression cassette(s) is/are provided that comprise (a) a polynucleotide encoding CRISPR-Cas effector protein comprising a cleavage domain and (b) a guide nucleic acid that binds to a target site in an endogenous *SHATTERPROOF MADS-BOX*
 10 (*SHP*) gene in canola, wherein the guide nucleic acid comprises a spacer sequence that is complementary to and binds to (i) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240** or **241**;
 (ii) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288** or **324-338**, optionally **SEQ ID**
 15 **NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**; (iii) a portion of a nucleic acid encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID**
NOs:71, 102, 150, 179, 208, or 242; and/or (iv) a portion of a nucleic acid encoding an amino
 20 acid sequence having at least 80% identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239** or **289-291**.

Also provided are nucleic acids encoding a Shatterproof MADS-box transcription factor (SHP) polypeptide (e.g., SHP1, SHP2, SHP3, SHP4), optionally wherein when present in a canola plant or plant part the mutated SHP polypeptide/mutated *SHP* gene results in the canola
 25 plant comprising a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control canola plant or plant part devoid of the mutation.

Nucleic acid constructs of the invention (e.g., a construct comprising a sequence specific nucleic acid binding domain (e.g., sequence specific DNA binding domain), a CRISPR-Cas
 30 effector domain, a deaminase domain, reverse transcriptase (RT), RT template and/or a guide nucleic acid, etc.) and expression cassettes/vectors comprising the same may be used as an editing system of this invention for modifying target nucleic acids (e.g., endogenous *SHP* genes,

e.g., endogenous *SHP1* gene, endogenous *SHP2* gene, endogenous *SHP3* gene, endogenous *SHP4* gene) and/or their expression.

Any canola plant comprising an endogenous *SHP* gene that is capable of conferring reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin when modified as described herein, may be modified (e.g., mutated, e.g., base edited, cleaved, nicked, etc.) as described herein (e.g., using the polypeptides, polynucleotides, RNPs, nucleic acid constructs, expression cassettes, and/or vectors of the invention) to reduce pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin in the canola plant.

An editing system useful with this invention can be any site-specific (sequence-specific) genome editing system now known or later developed, which system can introduce mutations in a target specific manner. For example, an editing system (e.g., site- or sequence-specific editing system) can include, but is not limited to, a CRISPR-Cas editing system, a meganuclease editing system, a zinc finger nuclease (ZFN) editing system, a transcription activator-like effector nuclease (TALEN) editing system, a base editing system and/or a prime editing system, each of which can comprise one or more polypeptides and/or one or more polynucleotides that when expressed as a system in a cell can modify (mutate) a target nucleic acid in a sequence specific manner. In some embodiments, an editing system (e.g., site- or sequence-specific editing system) can comprise one or more polynucleotides and/or one or more polypeptides, including but not limited to a nucleic acid binding domain (DNA binding domain), a nuclease, and/or other polypeptide, and/or a polynucleotide.

In some embodiments, an editing system can comprise one or more sequence-specific nucleic acid binding domains (DNA binding domains) that can be from, for example, a polynucleotide-guided endonuclease, a CRISPR-Cas endonuclease (e.g., CRISPR-Cas effector protein), a zinc finger nuclease, a transcription activator-like effector nuclease (TALEN) and/or an Argonaute protein. In some embodiments, an editing system can comprise one or more cleavage domains (e.g., nucleases) including, but not limited to, an endonuclease (e.g., FokI), a polynucleotide-guided endonuclease, a CRISPR-Cas endonuclease (e.g., CRISPR-Cas effector protein), a zinc finger nuclease, and/or a transcription activator-like effector nuclease (TALEN).

In some embodiments, an editing system can comprise one or more polypeptides that include, but are not limited to, a deaminase (e.g., a cytosine deaminase, an adenine deaminase), a reverse transcriptase, a Dna2 polypeptide, and/or a 5' flap endonuclease (FEN). In some embodiments, an editing system can comprise one or more polynucleotides, including, but is not limited to, a

CRISPR array (CRISPR guide) nucleic acid, extended guide nucleic acid, and/or a reverse transcriptase template.

In some embodiments, a method of modifying or editing *SHATTERPROOF MADS-BOX* (*SHP*) gene may comprise contacting a target nucleic acid (e.g., a nucleic acid encoding a Shatterproof MADS-box transcription factor (*SHP*) polypeptide, e.g., a *SHP1* polypeptide, a *SHP2* polypeptide, a *SHP3* polypeptide, a *SHP4* polypeptide) with a base-editing fusion protein (e.g., a sequence specific DNA binding protein (e.g., a CRISPR-Cas effector protein or domain) fused to a deaminase domain (e.g., an adenine deaminase and/or a cytosine deaminase) and a guide nucleic acid, wherein the guide nucleic acid is capable of guiding/targeting the base editing fusion protein to the target nucleic acid, thereby editing a locus within the target nucleic acid. In some embodiments, a base editing fusion protein and guide nucleic acid may be comprised in one or more expression cassettes. In some embodiments, the target nucleic acid may be contacted with a base editing fusion protein and an expression cassette comprising a guide nucleic acid. In some embodiments, the sequence-specific nucleic acid binding fusion proteins and guides may be provided as ribonucleoproteins (RNPs). In some embodiments, a cell may be contacted with more than one base-editing fusion protein and/or one or more guide nucleic acids that may target one or more target nucleic acids in the cell.

In some embodiments, a method of modifying or editing a *SHATTERPROOF MADS-BOX* (*SHP*) gene may comprise contacting a target nucleic acid (e.g., a nucleic acid encoding a *SHP* polypeptide) with a sequence-specific nucleic acid binding fusion protein (e.g., a sequence-specific DNA binding protein (e.g., a CRISPR-Cas effector protein or domain) fused to a peptide tag, a deaminase fusion protein comprising a deaminase domain (e.g., an adenine deaminase and/or a cytosine deaminase) fused to an affinity polypeptide that is capable of binding to the peptide tag, and a guide nucleic acid, wherein the guide nucleic acid is capable of guiding/targeting the sequence-specific nucleic acid binding fusion protein to the target nucleic acid and the sequence-specific nucleic acid binding fusion protein is capable of recruiting the deaminase fusion protein to the target nucleic acid via the peptide tag-affinity polypeptide interaction, thereby editing a locus within the target nucleic acid. In some embodiments, the sequence-specific nucleic acid binding fusion protein may be fused to the affinity polypeptide that binds the peptide tag and the deaminase may be fused to the peptide tag, thereby recruiting the deaminase to the sequence-specific nucleic acid binding fusion protein and to the target nucleic acid. In some embodiments, the sequence-specific binding fusion protein, deaminase fusion protein, and guide nucleic acid may be comprised in one or more expression cassettes. In

some embodiments, the target nucleic acid may be contacted with a sequence-specific binding fusion protein, deaminase fusion protein, and an expression cassette comprising a guide nucleic acid. In some embodiments, the sequence-specific nucleic acid binding fusion proteins, deaminase fusion proteins and guides may be provided as ribonucleoproteins (RNPs).

5 In some embodiments, methods such as prime editing may be used to generate a mutation in an endogenous *SHP* gene in a canola plant or part thereof. In prime editing, RNA-dependent DNA polymerase (reverse transcriptase, RT) and reverse transcriptase templates (RT template) are used in combination with sequence specific nucleic acid binding domains that confer the ability to recognize and bind the target in a sequence-specific manner, and which can
10 also cause a nick of the PAM-containing strand within the target. The nucleic acid binding domain may be a CRISPR-Cas effector protein and in this case, the CRISPR array or guide RNA may be an extended guide that comprises an extended portion comprising a primer binding site (PSB) and the edit to be incorporated into the genome (the template). Similar to base editing, prime editing can take advantages of the various methods of recruiting proteins for use
15 in the editing to the target site, such methods including both non-covalent and covalent interactions between the proteins and nucleic acids used in the selected process of genome editing.

As used herein, a "CRISPR-Cas effector protein" is a protein or polypeptide or domain thereof that cleaves or cuts a nucleic acid, binds a nucleic acid (e.g., a target nucleic acid and/or
20 a guide nucleic acid), and/or that identifies, recognizes, or binds a guide nucleic acid as defined herein. In some embodiments, a CRISPR-Cas effector protein may be an enzyme (e.g., a nuclease, endonuclease, nickase, etc.) or portion thereof and/or may function as an enzyme. In some embodiments, a CRISPR-Cas effector protein refers to a CRISPR-Cas nuclease polypeptide or domain thereof that comprises nuclease activity or in which the nuclease activity
25 has been reduced or eliminated, and/or comprises nickase activity or in which the nickase has been reduced or eliminated, and/or comprises single stranded DNA cleavage activity (ss DNase activity) or in which the ss DNase activity has been reduced or eliminated, and/or comprises self-processing RNase activity or in which the self-processing RNase activity has been reduced or eliminated. A CRISPR-Cas effector protein may bind to a target nucleic acid.

30 In some embodiments, a sequence-specific nucleic acid binding domain may be a CRISPR-Cas effector protein. In some embodiments, a CRISPR-Cas effector protein may be from a Type I CRISPR-Cas system, a Type II CRISPR-Cas system, a Type III CRISPR-Cas system, a Type IV CRISPR-Cas system, Type V CRISPR-Cas system, or a Type VI CRISPR-

Cas system. In some embodiments, a CRISPR-Cas effector protein of the invention may be from a Type II CRISPR-Cas system or a Type V CRISPR-Cas system. In some embodiments, a CRISPR-Cas effector protein may be Type II CRISPR-Cas effector protein, for example, a Cas9 effector protein. In some embodiments, a CRISPR-Cas effector protein may be Type V CRISPR-Cas effector protein, for example, a Cas12 effector protein.

In some embodiments, a CRISPR-Cas effector protein may include, but is not limited to, a Cas9, C2c1, C2c3, Cas12a (also referred to as Cpf1), Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c, Cas13d, Cas1, Cas1B, Cas2, Cas3, Cas3', Cas3", Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Cse1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 (dinG), and/or Csf5 nuclease, optionally wherein the CRISPR-Cas effector protein may be a Cas9, Cas12a (Cpf1), Cas12b, Cas12c (C2c3), Cas12d (CasY), Cas12e (CasX), Cas12g, Cas12h, Cas12i, C2c4, C2c5, C2c8, C2c9, C2c10, Cas14a, Cas14b, and/or Cas14c effector protein.

In some embodiments, a CRISPR-Cas effector protein useful with the invention may comprise a mutation in its nuclease active site (e.g., RuvC, HNH, e.g., RuvC site of a Cas12a nuclease domain; e.g., RuvC site and/or HNH site of a Cas9 nuclease domain). A CRISPR-Cas effector protein having a mutation in its nuclease active site, and therefore, no longer comprising nuclease activity, is commonly referred to as "dead," e.g., dCas. In some embodiments, a CRISPR-Cas effector protein domain or polypeptide having a mutation in its nuclease active site may have impaired activity or reduced activity as compared to the same CRISPR-Cas effector protein without the mutation, e.g., a nickase, e.g., Cas9 nickase, Cas12a nickase.

A CRISPR Cas9 effector protein or CRISPR Cas9 effector domain useful with this invention may be any known or later identified Cas9 nuclease. In some embodiments, a CRISPR Cas9 polypeptide can be a Cas9 polypeptide from, for example, *Streptococcus* spp. (e.g., *S. pyogenes*, *S. thermophilus*), *Lactobacillus* spp., *Bifidobacterium* spp., *Kandleria* spp., *Leuconostoc* spp., *Oenococcus* spp., *Pediococcus* spp., *Weissella* spp., and/or *Olsenella* spp. Example Cas9 sequences include, but are not limited to, the amino acid sequences of **SEQ ID NO:56** and **SEQ ID NO:57** or the nucleotide sequences of **SEQ ID NOs:58-68**.

In some embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide derived from *Streptococcus pyogenes* and recognizes the PAM sequence motif NGG, NAG, NGA (Mali et al, Science 2013; 339(6121): 823-826). In some embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide derived from *Streptococcus thermophiles* and

recognizes the PAM sequence motif NGGNG and/or NNAGAAW (W = A or T) (See, e.g., Horvath et al, Science, 2010; 327(5962): 167-170, and Deveau et al, J Bacteriol 2008; 190(4): 1390-1400). In some embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide derived from *Streptococcus mutans* and recognizes the PAM sequence motif NGG and/or NAAR (R = A or G) (See, e.g., Deveau et al, J BACTERIOL 2008; 190(4): 1390-1400). In some
 5 embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide derived from *Streptococcus aureus* and recognizes the PAM sequence motif NNGRR (R = A or G). In some
 10 embodiments, the CRISPR-Cas effector protein may be a Cas9 protein derived from *S. aureus*, which recognizes the PAM sequence motif N GRRT (R = A or G). In some
 15 embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide derived from *S. aureus*, which recognizes the PAM sequence motif N GRRV (R = A or G). In some
 20 embodiments, the CRISPR-Cas effector protein may be a Cas9 polypeptide that is derived from *Neisseria meningitidis* and recognizes the PAM sequence motif N GATT or N GCTT (R = A or G, V = A, G or C) (See, e.g., Hou et al, PNAS 2013, 1-6). In the aforementioned embodiments, N can be
 25 any nucleotide residue, e.g., any of A, G, C or T. In some embodiments, the CRISPR-Cas
 30 effector protein may be a Cas13a protein derived from *Leptotrichia shahii*, which recognizes a protospacer flanking sequence (PFS) (or RNA PAM (rPAM)) sequence motif of a single 3' A, U, or C, which may be located within the target nucleic acid.

In some embodiments, the CRISPR-Cas effector protein may be derived from Cas12a,
 20 which is a Type V Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas
 25 nuclease, see, e.g., amino acid sequences of **SEQ ID NOs:1-17**, nucleic acid sequences of **SEQ ID NOs:18-20**. Cas12a differs in several respects from the more well-known Type II CRISPR
 30 Cas9 nuclease. For example, Cas9 recognizes a G-rich protospacer-adjacent motif (PAM) that is
 3' to its guide RNA (gRNA, sgRNA, crRNA, crDNA, CRISPR array) binding site (protospacer,
 target nucleic acid, target DNA) (3'-NGG), while Cas12a recognizes a T-rich PAM that is
 located 5' to the target nucleic acid (5'-TTN, 5'-TTTN. In fact, the orientations in which Cas9
 and Cas12a bind their guide RNAs are very nearly reversed in relation to their N and C termini.
 Furthermore, Cas12a enzymes use a single guide RNA (gRNA, CRISPR array, crRNA) rather
 than the dual guide RNA (sgRNA (e.g., crRNA and tracrRNA)) found in natural Cas9 systems,
 and Cas12a processes its own gRNAs. Additionally, Cas12a nuclease activity produces
 staggered DNA double stranded breaks instead of blunt ends produced by Cas9 nuclease
 activity, and Cas12a relies on a single RuvC domain to cleave both DNA strands, whereas Cas9
 utilizes an HNH domain and a RuvC domain for cleavage.

A CRISPR Cas12a effector protein/domain useful with this invention may be any known or later identified Cas12a polypeptide (previously known as Cpf1) (see, e.g., U.S. Patent No. 9,790,490, which is incorporated by reference for its disclosures of Cpf1 (Cas12a) sequences). The term "Cas12a", "Cas12a polypeptide" or "Cas12a domain" refers to an RNA-guided
5 nuclease comprising a Cas12a polypeptide, or a fragment thereof, which comprises the guide nucleic acid binding domain of Cas12a and/or an active, inactive, or partially active DNA cleavage domain of Cas12a. In some embodiments, a Cas12a useful with the invention may comprise a mutation in the nuclease active site (e.g., RuvC site of the Cas12a domain). A Cas12a domain or Cas12a polypeptide having a mutation in its nuclease active site, and
10 therefore, no longer comprising nuclease activity, is commonly referred to as deadCas12a (e.g., dCas12a). In some embodiments, a Cas12a domain or Cas12a polypeptide having a mutation in its nuclease active site may have impaired activity, e.g., may have nickase activity.

Any deaminase domain/polypeptide useful for base editing may be used with this invention. In some embodiments, the deaminase domain may be a cytosine deaminase domain
15 or an adenine deaminase domain. A cytosine deaminase (or cytidine deaminase) useful with this invention may be any known or later identified cytosine deaminase from any organism (see, e.g., U.S. Patent No. 10,167,457 and Thuronyi et al. *Nat. Biotechnol.* 37:1070–1079 (2019), each of which is incorporated by reference herein for its disclosure of cytosine deaminases). Cytosine deaminases can catalyze the hydrolytic deamination of cytidine or deoxycytidine to uridine or
20 deoxyuridine, respectively. Thus, in some embodiments, a deaminase or deaminase domain useful with this invention may be a cytidine deaminase domain, catalyzing the hydrolytic deamination of cytosine to uracil. In some embodiments, a cytosine deaminase may be a variant of a naturally occurring cytosine deaminase, including but not limited to a primate (e.g., a human, monkey, chimpanzee, gorilla), a dog, a cow, a rat or a mouse. Thus, in some
25 embodiments, a cytosine deaminase useful with the invention may be about 70% to about 100% identical to a wild type cytosine deaminase (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, and any range or value therein, to a naturally occurring cytosine deaminase).

30 In some embodiments, a cytosine deaminase useful with the invention may be an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase. In some embodiments, the cytosine deaminase may be an APOBEC1 deaminase, an APOBEC2 deaminase, an APOBEC3A deaminase, an APOBEC3B deaminase, an APOBEC3C deaminase, an

APOBEC3D deaminase, an APOBEC3F deaminase, an APOBEC3G deaminase, an APOBEC3H deaminase, an APOBEC4 deaminase, a human activation induced deaminase (hAID), an rAPOBEC1, FERNY, and/or a CDA1, optionally a pmCDA1, an atCDA1 (e.g., At2g19570), and evolved versions of the same (e.g., **SEQ ID NO:27**, **SEQ ID NO:28** or **SEQ ID NO:29**). In some embodiments, the cytosine deaminase may be an APOBEC1 deaminase having the amino acid sequence of **SEQ ID NO:23**. In some embodiments, the cytosine deaminase may be an APOBEC3A deaminase having the amino acid sequence of **SEQ ID NO:24**. In some embodiments, the cytosine deaminase may be an CDA1 deaminase, optionally a CDA1 having the amino acid sequence of **SEQ ID NO:25**. In some embodiments, the cytosine deaminase may be a FERNY deaminase, optionally a FERNY having the amino acid sequence of **SEQ ID NO:26**. In some embodiments, a cytosine deaminase useful with the invention may be about 70% to about 100% identical (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100% identical) to the amino acid sequence of a naturally occurring cytosine deaminase (e.g., an evolved deaminase). In some embodiments, a cytosine deaminase useful with the invention may be about 70% to about 99.5% identical (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical) to the amino acid sequence of **SEQ ID NO:23**, **SEQ ID NO:24**, **SEQ ID NO:25** or **SEQ ID NO:26** (e.g., at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of **SEQ ID NO:23**, **SEQ ID NO:24**, **SEQ ID NO:25**, **SEQ ID NO:26**, **SEQ ID NO:27**, **SEQ ID NO:28** or **SEQ ID NO:29**). In some embodiments, a polynucleotide encoding a cytosine deaminase may be codon optimized for expression in a plant and the codon optimized polypeptide may be about 70% to 99.5% identical to the reference polynucleotide.

In some embodiments, a nucleic acid construct of this invention may further encode a uracil glycosylase inhibitor (UGI) (e.g., uracil-DNA glycosylase inhibitor) polypeptide/domain. Thus, in some embodiments, a nucleic acid construct encoding a CRISPR-Cas effector protein and a cytosine deaminase domain (e.g., encoding a fusion protein comprising a CRISPR-Cas effector protein domain fused to a cytosine deaminase domain, and/or a CRISPR-Cas effector protein domain fused to a peptide tag or to an affinity polypeptide capable of binding a peptide tag and/or a deaminase protein domain fused to a peptide tag or to an affinity polypeptide capable of binding a peptide tag) may further encode a uracil-DNA glycosylase inhibitor (UGI),

optionally wherein the UGI may be codon optimized for expression in a plant. In some embodiments, the invention provides fusion proteins comprising a CRISPR-Cas effector polypeptide, a deaminase domain, and a UGI and/or one or more polynucleotides encoding the same, optionally wherein the one or more polynucleotides may be codon optimized for expression in a plant. In some embodiments, the invention provides fusion proteins, wherein a CRISPR-Cas effector polypeptide, a deaminase domain, and a UGI may be fused to any combination of peptide tags and affinity polypeptides as described herein, thereby recruiting the deaminase domain and UGI to the CRISPR-Cas effector polypeptide and a target nucleic acid. In some embodiments, a guide nucleic acid may be linked to a recruiting RNA motif and one or more of the deaminase domain and/or UGI may be fused to an affinity polypeptide that is capable of interacting with the recruiting RNA motif, thereby recruiting the deaminase domain and UGI to a target nucleic acid.

A "uracil glycosylase inhibitor" useful with the invention may be any protein that is capable of inhibiting a uracil-DNA glycosylase base-excision repair enzyme. In some embodiments, a UGI domain comprises a wild type UGI or a fragment thereof. In some embodiments, a UGI domain useful with the invention may be about 70% to about 100% identical (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100% identical and any range or value therein) to the amino acid sequence of a naturally occurring UGI domain. In some embodiments, a UGI domain may comprise the amino acid sequence of **SEQ ID NO:41** or a polypeptide having about 70% to about 99.5% sequence identity to the amino acid sequence of **SEQ ID NO:41** (e.g., at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of **SEQ ID NO:41**). For example, in some embodiments, a UGI domain may comprise a fragment of the amino acid sequence of **SEQ ID NO:41** that is 100% identical to a portion of consecutive nucleotides (e.g., 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 consecutive nucleotides; e.g., about 10, 15, 20, 25, 30, 35, 40, 45, to about 50, 55, 60, 65, 70, 75, 80 consecutive nucleotides) of the amino acid sequence of **SEQ ID NO:41**. In some embodiments, a UGI domain may be a variant of a known UGI (e.g., **SEQ ID NO:41**) having about 70% to about 99.5% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% sequence identity, and any range or value therein) to the known UGI. In some embodiments, a

polynucleotide encoding a UGI may be codon optimized for expression in a plant (e.g., a plant) and the codon optimized polypeptide may be about 70% to about 99.5% identical to the reference polynucleotide.

An adenine deaminase (or adenosine deaminase) useful with this invention may be any known or later identified adenine deaminase from any organism (see, e.g., U.S. Patent No. 10,113,163, which is incorporated by reference herein for its disclosure of adenine deaminases). An adenine deaminase can catalyze the hydrolytic deamination of adenine or adenosine. In some embodiments, the adenine deaminase may catalyze the hydrolytic deamination of adenosine or deoxyadenosine to inosine or deoxyinosine, respectively. In some embodiments, the adenosine deaminase may catalyze the hydrolytic deamination of adenine or adenosine in DNA. In some embodiments, an adenine deaminase encoded by a nucleic acid construct of the invention may generate an A→G conversion in the sense (e.g., "+"; template) strand of the target nucleic acid or a T→C conversion in the antisense (e.g., "-", complementary) strand of the target nucleic acid.

In some embodiments, an adenosine deaminase may be a variant of a naturally occurring adenine deaminase. Thus, in some embodiments, an adenosine deaminase may be about 70% to 100% identical to a wild type adenine deaminase (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, and any range or value therein, to a naturally occurring adenine deaminase). In some embodiments, the deaminase or deaminase does not occur in nature and may be referred to as an engineered, mutated or evolved adenosine deaminase. Thus, for example, an engineered, mutated or evolved adenine deaminase polypeptide or an adenine deaminase domain may be about 70% to 99.9% identical to a naturally occurring adenine deaminase polypeptide/domain (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8% or 99.9% identical, and any range or value therein, to a naturally occurring adenine deaminase polypeptide or adenine deaminase domain). In some embodiments, the adenosine deaminase may be from a bacterium, (e.g., *Escherichia coli*, *Staphylococcus aureus*, *Haemophilus influenzae*, *Caulobacter crescentus*, and the like). In some embodiments, a polynucleotide encoding an adenine deaminase polypeptide/domain may be codon optimized for expression in a plant.

In some embodiments, an adenine deaminase domain may be a wild type tRNA-specific adenosine deaminase domain, e.g., a tRNA-specific adenosine deaminase (TadA) and/or a mutated/evolved adenosine deaminase domain, e.g., mutated/evolved tRNA-specific adenosine deaminase domain (TadA*). In some embodiments, a TadA domain may be from *E. coli*. In some embodiments, the TadA may be modified, e.g., truncated, missing one or more N-terminal and/or C-terminal amino acids relative to a full-length TadA (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 N-terminal and/or C terminal amino acid residues may be missing relative to a full length TadA. In some embodiments, a TadA polypeptide or TadA domain does not comprise an N-terminal methionine. In some embodiments, a wild type *E. coli* TadA comprises the amino acid sequence of **SEQ ID NO:30**. In some embodiments, a mutated/evolved *E. coli* TadA* comprises the amino acid sequence of **SEQ ID NOs:31-40** (e.g., **SEQ ID NOs: 31, 32, 33, 34, 35, 36, 37, 38, 39** or **40**). In some embodiments, a polynucleotide encoding a TadA/TadA* may be codon optimized for expression in a plant.

A cytosine deaminase catalyzes cytosine deamination and results in a thymidine (through a uracil intermediate), causing a C to T conversion, or a G to A conversion in the complementary strand in the genome. Thus, in some embodiments, the cytosine deaminase encoded by the polynucleotide of the invention generates a C→T conversion in the sense (e.g., "+"; template) strand of the target nucleic acid or a G→A conversion in antisense (e.g., "-", complementary) strand of the target nucleic acid.

In some embodiments, the adenine deaminase encoded by the nucleic acid construct of the invention generates an A→G conversion in the sense (e.g., "+"; template) strand of the target nucleic acid or a T→C conversion in the antisense (e.g., "-", complementary) strand of the target nucleic acid.

The nucleic acid constructs of the invention encoding a base editor comprising a sequence-specific nucleic acid binding protein and a cytosine deaminase polypeptide, and nucleic acid constructs/expression cassettes/vectors encoding the same, may be used in combination with guide nucleic acids for modifying target nucleic acid including, but not limited to, generation of C→T or G→A mutations in a target nucleic acid including, but not limited to, a plasmid sequence; generation of C→T or G→A mutations in a coding sequence to alter an amino acid identity; generation of C→T or G→A mutations in a coding sequence to generate a stop codon; generation of C→T or G→A mutations in a coding sequence to disrupt a start codon; generation of point mutations in genomic DNA to disrupt function; and/or generation of point mutations in genomic DNA to disrupt splice junctions.

The nucleic acid constructs of the invention encoding a base editor comprising a sequence-specific nucleic acid binding protein and an adenine deaminase polypeptide, and expression cassettes and/or vectors encoding the same may be used in combination with guide nucleic acids for modifying a target nucleic acid including, but not limited to, generation of A→G or T→C mutations in a target nucleic acid including, but not limited to, a plasmid sequence; generation of A→G or T→C mutations in a coding sequence to alter an amino acid identity; generation of A→G or T→C mutations in a coding sequence to generate a stop codon; generation of A→G or T→C mutations in a coding sequence to disrupt a start codon; generation of point mutations in genomic DNA to disrupt function; and/or generation of point mutations in genomic DNA to disrupt splice junctions.

The nucleic acid constructs of the invention comprising a CRISPR-Cas effector protein or a fusion protein thereof may be used in combination with a guide RNA (gRNA, CRISPR array, CRISPR RNA, crRNA), designed to function with the encoded CRISPR-Cas effector protein or domain, to modify a target nucleic acid. A guide nucleic acid useful with this invention comprises at least one spacer sequence and at least one repeat sequence. The guide nucleic acid is capable of forming a complex with the CRISPR-Cas nuclease domain encoded and expressed by a nucleic acid construct of the invention and the spacer sequence is capable of hybridizing to a target nucleic acid, thereby guiding the complex (e.g., a CRISPR-Cas effector fusion protein (e.g., CRISPR-Cas effector domain fused to a deaminase domain and/or a CRISPR-Cas effector domain fused to a peptide tag or an affinity polypeptide to recruit a deaminase domain and optionally, a UGI) to the target nucleic acid, wherein the target nucleic acid may be modified (e.g., cleaved or edited) or modulated (e.g., modulating transcription) by the deaminase domain.

As an example, a nucleic acid construct encoding a Cas9 domain linked to a cytosine deaminase domain (e.g., fusion protein) may be used in combination with a Cas9 guide nucleic acid to modify a target nucleic acid, wherein the cytosine deaminase domain of the fusion protein deaminates a cytosine base in the target nucleic acid, thereby editing the target nucleic acid. In a further example, a nucleic acid construct encoding a Cas9 domain linked to an adenine deaminase domain (e.g., fusion protein) may be used in combination with a Cas9 guide nucleic acid to modify a target nucleic acid, wherein the adenine deaminase domain of the fusion protein deaminates an adenosine base in the target nucleic acid, thereby editing the target nucleic acid.

Likewise, a nucleic acid construct encoding a Cas12a domain (or other selected CRISPR-Cas nuclease, e.g., C2c1, C2c3, Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c, Cas13d, Cas1, Cas1B, Cas2, Cas3, Cas3', Cas3'', Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Cse3, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 (dinG), and/or Csf5) linked to a cytosine deaminase domain or adenine deaminase domain (e.g., fusion protein) may be used in combination with a Cas12a guide nucleic acid (or the guide nucleic acid for the other selected CRISPR-Cas nuclease) to modify a target nucleic acid, wherein the cytosine deaminase domain or adenine deaminase domain of the fusion protein deaminates a cytosine base in the target nucleic acid, thereby editing the target nucleic acid.

A "guide nucleic acid," "guide RNA," "gRNA," "CRISPR RNA/DNA" "crRNA" or "crDNA" as used herein means a nucleic acid that comprises at least one spacer sequence, which is complementary to (and hybridizes to) a target DNA (e.g., protospacer), and at least one repeat sequence (e.g., a repeat of a Type V Cas12a CRISPR-Cas system, or a fragment or portion thereof; a repeat of a Type II Cas9 CRISPR-Cas system, or fragment thereof; a repeat of a Type V C2c1 CRISPR Cas system, or a fragment thereof; a repeat of a CRISPR-Cas system of, for example, C2c3, Cas12a (also referred to as Cpf1), Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c, Cas13d, Cas1, Cas1B, Cas2, Cas3, Cas3', Cas3'', Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Cse3, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 (dinG), and/or Csf5, or a fragment thereof), wherein the repeat sequence may be linked to the 5' end and/or the 3' end of the spacer sequence. The design of a gRNA of this invention may be based on a Type I, Type II, Type III, Type IV, Type V, or Type VI CRISPR-Cas system.

In some embodiments, a Cas12a gRNA may comprise, from 5' to 3', a repeat sequence (full length or portion thereof ("handle"); e.g., pseudoknot-like structure) and a spacer sequence.

In some embodiments, a guide nucleic acid may comprise more than one repeat sequence-spacer sequence (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, or more repeat-spacer sequences) (e.g., repeat-spacer-repeat, e.g., repeat-spacer-repeat-spacer-repeat-spacer-repeat-spacer-repeat-spacer, and the like). The guide nucleic acids of this invention are synthetic, human-made, and not found in nature. A gRNA can be quite long and may be used as an aptamer (like in the MS2 recruitment strategy) or other RNA structures hanging off the spacer.

A "repeat sequence" as used herein, refers to, for example, any repeat sequence of a wild-type CRISPR Cas locus (e.g., a Cas9 locus, a Cas12a locus, a C2c1 locus, etc.) or a repeat sequence of a synthetic crRNA that is functional with the CRISPR-Cas effector protein encoded by the nucleic acid constructs of the invention. A repeat sequence useful with this invention can be any known or later identified repeat sequence of a CRISPR-Cas locus (e.g., Type I, Type II, Type III, Type IV, Type V or Type VI) or it can be a synthetic repeat designed to function in a Type I, II, III, IV, V or VI CRISPR-Cas system. A repeat sequence may comprise a hairpin structure and/or a stem loop structure. In some embodiments, a repeat sequence may form a pseudoknot-like structure at its 5' end (i.e., "handle"). Thus, in some embodiments, a repeat sequence can be identical to or substantially identical to a repeat sequence from wild-type Type I CRISPR-Cas loci, Type II, CRISPR-Cas loci, Type III, CRISPR-Cas loci, Type IV CRISPR-Cas loci, Type V CRISPR-Cas loci and/or Type VI CRISPR-Cas loci. A repeat sequence from a wild-type CRISPR-Cas locus may be determined through established algorithms, such as using the CRISPRfinder offered through CRISPRdb (see, Grissa et al. *Nucleic Acids Res.* 35(Web Server issue):W52-7). In some embodiments, a repeat sequence or portion thereof is linked at its 3' end to the 5' end of a spacer sequence, thereby forming a repeat-spacer sequence (e.g., guide nucleic acid, guide RNA/DNA, crRNA, crDNA).

In some embodiments, a repeat sequence comprises, consists essentially of, or consists of at least 10 nucleotides depending on the particular repeat and whether the guide nucleic acid comprising the repeat is processed or unprocessed (e.g., about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 to 100 or more nucleotides, or any range or value therein). In some embodiments, a repeat sequence comprises, consists essentially of, or consists of about 10 to about 20, about 10 to about 30, about 10 to about 45, about 10 to about 50, about 15 to about 30, about 15 to about 40, about 15 to about 45, about 15 to about 50, about 20 to about 30, about 20 to about 40, about 20 to about 50, about 30 to about 40, about 40 to about 80, about 50 to about 100 or more nucleotides.

A repeat sequence linked to the 5' end of a spacer sequence can comprise a portion of a repeat sequence (e.g., 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or more contiguous nucleotides of a wild type repeat sequence). In some embodiments, a portion of a repeat sequence linked to the 5' end of a spacer sequence can be about five to about ten consecutive nucleotides in length (e.g., about 5, 6, 7, 8, 9, 10 nucleotides) and have at least 90% sequence identity (e.g., at least about 90%, 91%, 92%,

93%, 94%, 95%, 96%, 97%, 98%, 99%, or more (e.g., 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, or 100%)) to the same region (e.g., 5' end) of a wild type CRISPR Cas repeat nucleotide sequence. In some embodiments, a portion of a repeat sequence may comprise a pseudoknot-like structure at its 5' end (e.g., "handle").

5 A "spacer sequence" as used herein is a nucleotide sequence that is complementary to a target nucleic acid (e.g., target DNA) (e.g., protospacer) (e.g., a portion of consecutive nucleotides of a sequence that (a) comprises a sequence having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% identity to any one of **SEQ ID**
 10 **NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**; (c) encodes an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179,**
 15 **208, or 242**; and/or (d) encodes an amino acid sequence comprising a region having at least 80% identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**. In some embodiments, a spacer sequence (e.g., one or more spacers) may include, but is not limited to, the nucleotide sequences of any one of **SEQ ID NOs:292-297** and/or **SEQ ID NOs:342-346**. The spacer sequence can be fully complementary or substantially complementary (e.g., at least
 20 about 70% complementary (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more (e.g., 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, or 100%)) to a target nucleic acid. Thus, in some embodiments, the spacer sequence can have one, two, three, four, or five mismatches as compared to the target nucleic acid, which mismatches
 25 can be contiguous or noncontiguous. In some embodiments, the spacer sequence can have 70% complementarity to a target nucleic acid. In other embodiments, the spacer nucleotide sequence can have 80% complementarity to a target nucleic acid. In still other embodiments, the spacer nucleotide sequence can have 85%, 90%, 95%, 96%, 97%, 98%, 99% or 99.5% complementarity, and the like, to the target nucleic acid (protospacer). In some embodiments,
 30 the spacer sequence is 100% complementary to the target nucleic acid. A spacer sequence may have a length from about 15 nucleotides to about 30 nucleotides (e.g., 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides, or any range or value therein). Thus, in some embodiments, a spacer sequence may have complete complementarity or substantial

complementarity over a region of a target nucleic acid (e.g., protospacer) that is at least about 15 nucleotides to about 30 nucleotides in length. In some embodiments, the spacer is about 20 nucleotides in length. In some embodiments, the spacer is about 21, 22, or 23 nucleotides in length.

5 In some embodiments, the 5' region of a spacer sequence of a guide nucleic acid may be identical to a target DNA, while the 3' region of the spacer may be substantially complementary to the target DNA (see, for example, a spacer sequence of a Type V CRISPR-Cas system), or the 3' region of a spacer sequence of a guide nucleic acid may be identical to a target DNA, while the 5' region of the spacer may be substantially complementary to the target DNA (see, for
10 example, a spacer sequence of a Type II CRISPR-Cas system), and therefore, the overall complementarity of the spacer sequence to the target DNA may be less than 100%. Thus, for example, in a guide for a Type V CRISPR-Cas system, the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 nucleotides in the 5' region (i.e., seed region) of, for example, a 20 nucleotide spacer sequence may be 100% complementary to the target DNA, while the remaining nucleotides in the 3'
15 region of the spacer sequence are substantially complementary (e.g., at least about 70% complementary) to the target DNA. In some embodiments, the first 1 to 8 nucleotides (e.g., the first 1, 2, 3, 4, 5, 6, 7, 8, nucleotides, and any range therein) of the 5' end of the spacer sequence may be 100% complementary to the target DNA, while the remaining nucleotides in the 3' region of the spacer sequence are substantially complementary (e.g., at least about 50%
20 complementary (e.g., 50%, 55%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more)) to the target DNA.

As a further example, in a guide for a Type II CRISPR-Cas system, the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 nucleotides in the 3' region (i.e., seed region) of, for example, a 20 nucleotide
25 spacer sequence may be 100% complementary to the target DNA, while the remaining nucleotides in the 5' region of the spacer sequence are substantially complementary (e.g., at least about 70% complementary) to the target DNA. In some embodiments, the first 1 to 10 nucleotides (e.g., the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 nucleotides, and any range therein) of the 3' end of the spacer sequence may be 100% complementary to the target DNA, while the
30 remaining nucleotides in the 5' region of the spacer sequence are substantially complementary (e.g., at least about 50% complementary (e.g., at least about 50%, 55%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,

88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more or any range or value therein)) to the target DNA.

In some embodiments, a seed region of a spacer may be about 8 to about 10 nucleotides in length, about 5 to about 6 nucleotides in length, or about 6 nucleotides in length.

5 As used herein, a "target nucleic acid", "target DNA," "target nucleotide sequence," "target region," or a "target region in the genome" refers to a region of a plant's genome that is fully complementary (100% complementary) or substantially complementary (e.g., at least 70% complementary (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 10 98%, 99%, or more)) to a spacer sequence in a guide nucleic acid of this invention. A target region useful for a CRISPR-Cas system may be located immediately 3' (e.g., Type V CRISPR-Cas system) or immediately 5' (e.g., Type II CRISPR-Cas system) to a PAM sequence in the genome of the organism (e.g., a plant genome). A target region may be selected from any region of at least 15 consecutive nucleotides (e.g., 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 15 30 nucleotides, and the like) located immediately adjacent to a PAM sequence.

A "protospacer sequence" refers to the target double stranded DNA and specifically to the portion of the target DNA (e.g., or target region in the genome) that is fully or substantially complementary (and hybridizes) to the spacer sequence of the CRISPR repeat-spacer sequences (e.g., guide nucleic acids, CRISPR arrays, crRNAs).

20 In the case of Type V CRISPR-Cas (e.g., Cas12a) systems and Type II CRISPR-Cas (Cas9) systems, the protospacer sequence is flanked by (e.g., immediately adjacent to) a protospacer adjacent motif (PAM). For Type IV CRISPR-Cas systems, the PAM is located at the 5' end on the non-target strand and at the 3' end of the target strand (see below, as an example).

25 5'-NNNNNNNNNNNNNNNNNNNN-3' RNA Spacer
 | | | | | | | | | | | | | | | |
 3'AAANNNNNNNNNNNNNNNNNNN-5' Target strand
 | | | |
 5'TTTNNNNNNNNNNNNNNNNNN-3' Non-target strand
 30

In the case of Type II CRISPR-Cas (e.g., Cas9) systems, the PAM is located immediately 3' of the target region. The PAM for Type I CRISPR-Cas systems is located 5' of the target strand. There is no known PAM for Type III CRISPR-Cas systems. Makarova et al. describes the nomenclature for all the classes, types and subtypes of CRISPR systems (*Nature Reviews* 35 *Microbiology* 13:722–736 (2015)). Guide structures and PAMs are described in by R. Barrangou

(*Genome Biol.* 16:247 (2015)).

Canonical Cas12a PAMs are T rich. In some embodiments, a canonical Cas12a PAM sequence may be 5'-TTN, 5'-TTTN, or 5'-TTTV. In some embodiments, canonical Cas9 (e.g., *S. pyogenes*) PAMs may be 5'-NGG-3'. In some embodiments, non-canonical PAMs may be used
5 but may be less efficient.

Additional PAM sequences may be determined by those skilled in the art through established experimental and computational approaches. Thus, for example, experimental approaches include targeting a sequence flanked by all possible nucleotide sequences and identifying sequence members that do not undergo targeting, such as through the transformation
10 of target plasmid DNA (Esvelt et al. 2013. *Nat. Methods* 10:1116-1121; Jiang et al. 2013. *Nat. Biotechnol.* 31:233-239). In some aspects, a computational approach can include performing BLAST searches of natural spacers to identify the original target DNA sequences in bacteriophages or plasmids and aligning these sequences to determine conserved sequences adjacent to the target sequence (Briner and Barrangou. 2014. *Appl. Environ. Microbiol.* 80:994-
15 1001; Mojica et al. 2009. *Microbiology* 155:733-740).

In some embodiments, the present invention provides expression cassettes and/or vectors comprising the nucleic acid constructs of the invention (e.g., one or more components of an editing system of the invention). In some embodiments, expression cassettes and/or vectors comprising the nucleic acid constructs of the invention and/or one or more guide nucleic acids
20 may be provided. In some embodiments, a nucleic acid construct of the invention encoding a base editor (e.g., a construct comprising a CRISPR-Cas effector protein and a deaminase domain (e.g., a fusion protein)) or the components for base editing (e.g., a CRISPR-Cas effector protein fused to a peptide tag or an affinity polypeptide, a deaminase domain fused to a peptide tag or an affinity polypeptide, and/or a UGI fused to a peptide tag or an affinity polypeptide), may be
25 comprised on the same or on a separate expression cassette or vector from that comprising the one or more guide nucleic acids. When the nucleic acid construct encoding a base editor or the components for base editing is/are comprised on separate expression cassette(s) or vector(s) from that comprising the guide nucleic acid, a target nucleic acid may be contacted with (e.g., provided with) the expression cassette(s) or vector(s) encoding the base editor or components for
30 base editing in any order from one another and the guide nucleic acid, e.g., prior to, concurrently with, or after the expression cassette comprising the guide nucleic acid is provided (e.g., contacted with the target nucleic acid).

Fusion proteins of the invention may comprise sequence-specific nucleic acid binding domains (e.g., sequence-specific DNA binding domains), CRISPR-Cas polypeptides, and/or deaminase domains fused to peptide tags or affinity polypeptides that interact with the peptide tags, as known in the art, for use in recruiting the deaminase to the target nucleic acid. Methods of recruiting may also comprise guide nucleic acids linked to RNA recruiting motifs and deaminases fused to affinity polypeptides capable of interacting with RNA recruiting motifs, thereby recruiting the deaminase to the target nucleic acid. Alternatively, chemical interactions may be used to recruit polypeptides (e.g., deaminases) to a target nucleic acid.

A peptide tag (e.g., epitope) useful with this invention may include, but is not limited to, a GCN4 peptide tag (e.g., Sun-Tag), a c-Myc affinity tag, an HA affinity tag, a His affinity tag, an S affinity tag, a methionine-His affinity tag, an RGD-His affinity tag, a FLAG octapeptide, a strep tag or strep tag II, a V5 tag, and/or a VSV-G epitope. Any epitope that may be linked to a polypeptide and for which there is a corresponding affinity polypeptide that may be linked to another polypeptide may be used with this invention as a peptide tag. In some embodiments, a peptide tag may comprise 1 or 2 or more copies of a peptide tag (e.g., repeat unit, multimerized epitope (e.g., tandem repeats)) (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more repeat units. In some embodiments, an affinity polypeptide that interacts with/binds to a peptide tag may be an antibody. In some embodiments, the antibody may be a scFv antibody. In some embodiments, an affinity polypeptide that binds to a peptide tag may be synthetic (e.g., evolved for affinity interaction) including, but not limited to, an affibody, an anticalin, a monobody and/or a DARPin (see, e.g., Sha et al., *Protein Sci.* 26(5):910-924 (2017)); Gilbreth (*Curr Opin Struc Biol* 22(4):413-420 (2013)), U.S. Patent No. 9,982,053, each of which are incorporated by reference in their entireties for the teachings relevant to affibodies, anticalins, monobodies and/or DARPins. Example peptide tag sequences and their affinity polypeptides include, but are not limited to, the amino acid sequences of **SEQ ID NOs:42-44**.

In some embodiments, a guide nucleic acid may be linked to an RNA recruiting motif, and a polypeptide to be recruited (e.g., a deaminase) may be fused to an affinity polypeptide that binds to the RNA recruiting motif, wherein the guide binds to the target nucleic acid and the RNA recruiting motif binds to the affinity polypeptide, thereby recruiting the polypeptide to the guide and contacting the target nucleic acid with the polypeptide (e.g., deaminase). In some embodiments, two or more polypeptides may be recruited to a guide nucleic acid, thereby contacting the target nucleic acid with two or more polypeptides (e.g., deaminases). Example

RNA recruiting motifs and their affinity polypeptides include, but are not limited to, the sequences of **SEQ ID NOs:45-55**.

In some embodiments, a polypeptide fused to an affinity polypeptide may be a reverse transcriptase and the guide nucleic acid may be an extended guide nucleic acid linked to an RNA recruiting motif. In some embodiments, an RNA recruiting motif may be located on the 3' end of the extended portion of an extended guide nucleic acid (e.g., 5'-3', repeat-spacer-extended portion (RT template-primer binding site)-RNA recruiting motif). In some embodiments, an RNA recruiting motif may be embedded in the extended portion.

In some embodiments of the invention, an extended guide RNA and/or guide RNA may be linked to one or to two or more RNA recruiting motifs (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more motifs; e.g., at least 10 to about 25 motifs), optionally wherein the two or more RNA recruiting motifs may be the same RNA recruiting motif or different RNA recruiting motifs. In some embodiments, an RNA recruiting motif and corresponding affinity polypeptide may include, but is not limited, to a telomerase Ku binding motif (e.g., Ku binding hairpin) and the corresponding affinity polypeptide Ku (e.g., Ku heterodimer), a telomerase Sm7 binding motif and the corresponding affinity polypeptide Sm7, an MS2 phage operator stem-loop and the corresponding affinity polypeptide MS2 Coat Protein (MCP), a PP7 phage operator stem-loop and the corresponding affinity polypeptide PP7 Coat Protein (PCP), an SfMu phage Com stem-loop and the corresponding affinity polypeptide Com RNA binding protein, a PUF binding site (PBS) and the affinity polypeptide Pumilio/fem-3 mRNA binding factor (PUF), and/or a synthetic RNA-aptamer and the aptamer ligand as the corresponding affinity polypeptide. In some embodiments, the RNA recruiting motif and corresponding affinity polypeptide may be an MS2 phage operator stem-loop and the affinity polypeptide MS2 Coat Protein (MCP). In some embodiments, the RNA recruiting motif and corresponding affinity polypeptide may be a PUF binding site (PBS) and the affinity polypeptide Pumilio/fem-3 mRNA binding factor (PUF).

In some embodiments, the components for recruiting polypeptides and nucleic acids may those that function through chemical interactions that may include, but are not limited to, rapamycin-inducible dimerization of FRB – FKBP; Biotin-streptavidin; SNAP tag; Halo tag; CLIP tag; DmrA-DmrC heterodimer induced by a compound; bifunctional ligand (e.g., fusion of two protein-binding chemicals together, e.g., dihydrofolate reductase (DHFR)).

In some embodiments, the nucleic acid constructs, expression cassettes or vectors of the invention that are optimized for expression in a plant may be about 70% to 100% identical (e.g., about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%,

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100%) to the nucleic acid constructs, expression cassettes or vectors comprising the same polynucleotide(s) but which have not been codon optimized for expression in a plant.

Further provided herein are cells comprising one or more polynucleotides, guide nucleic acids, nucleic acid constructs, expression cassettes or vectors of the invention.

The nucleic acid constructs of the invention (e.g., a construct comprising a sequence specific DNA binding domain, a CRISPR-Cas effector domain, a deaminase domain, reverse transcriptase (RT), RT template and/or a guide nucleic acid, etc.) and expression cassettes/vectors comprising the same may be used as an editing system of this invention for modifying target nucleic acids and/or their expression.

A target nucleic acid of any plant or plant part (or groupings of plants, for example, into a genus or higher order classification) may be modified (e.g., mutated, e.g., base edited, cleaved, nicked, etc.) using the polypeptides, polynucleotides, ribonucleoproteins (RNPs), nucleic acid constructs, expression cassettes, and/or vectors of the invention including an angiosperm, a gymnosperm, a monocot, a dicot, a C3, C4, CAM plant, a bryophyte, a fern and/or fern ally, a microalgae, and/or a macroalgae. A plant and/or plant part that may be modified as described herein may be a plant and/or plant part of any plant species/variety/cultivar. In some embodiments, a plant that may be modified as described herein is a monocot. In some embodiments, a plant that may be modified as described herein is a dicot.

The term "plant part," as used herein, includes but is not limited to reproductive tissues (e.g., petals, sepals, stamens, pistils, receptacles, anthers, pollen, flowers, fruits, flower bud, ovules, seeds, embryos, nuts, kernels, ears, cobs and husks); vegetative tissues (e.g., petioles, stems, roots, root hairs, root tips, pith, coleoptiles, stalks, shoots, branches, bark, apical meristem, axillary bud, cotyledon, hypocotyls, and leaves); vascular tissues (e.g., phloem and xylem); specialized cells such as epidermal cells, parenchyma cells, chlorenchyma cells, schlerenchyma cells, stomates, guard cells, cuticle, mesophyll cells; callus tissue; and cuttings. The term "plant part" also includes plant cells, including plant cells that are intact in plants and/or parts of plants, plant protoplasts, plant tissues, plant organs, plant cell tissue cultures, plant calli, plant clumps, and the like. As used herein, "shoot" refers to the above ground parts including the leaves and stems. As used herein, the term "tissue culture" encompasses cultures of tissue, cells, protoplasts and callus.

As used herein, "plant cell" refers to a structural and physiological unit of the plant, which typically comprise a cell wall but also includes protoplasts. A plant cell of the present

invention can be in the form of an isolated single cell or can be a cultured cell or can be a part of a higher-organized unit such as, for example, a plant tissue (including callus) or a plant organ. In some embodiments, a plant cell can be an algal cell. A "protoplast" is an isolated plant cell without a cell wall or with only parts of the cell wall. Thus, in some embodiments of the invention, a transgenic cell comprising a nucleic acid molecule and/or nucleotide sequence of the invention is a cell of any plant or plant part including, but not limited to, a root cell, a leaf cell, a tissue culture cell, a seed cell, a flower cell, a fruit cell, a pollen cell, and the like. In some aspects of the invention, the plant part can be a plant germplasm. In some aspects, a plant cell can be non-propagating plant cell that does not regenerate into a plant.

"Plant cell culture" means cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes and embryos at various stages of development.

As used herein, a "plant organ" is a distinct and visibly structured and differentiated part of a plant such as a root, stem, leaf, flower bud, or embryo.

"Plant tissue" as used herein means a group of plant cells organized into a structural and functional unit. Any tissue of a plant *in planta* or in culture is included. This term includes, but is not limited to, whole plants, plant organs, plant seeds, tissue culture and any groups of plant cells organized into structural and/or functional units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

In some embodiments of the invention, a transgenic tissue culture or transgenic plant cell culture is provided, wherein the transgenic tissue or cell culture comprises a nucleic acid molecule/nucleotide sequence of the invention. In some embodiments, transgenes may be eliminated from a plant developed from the transgenic tissue or cell by breeding of the transgenic plant with a non-transgenic plant and selecting among the progeny for the plants comprising the desired gene edit and not the transgenes used in producing the edit.

Any canola plant comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene may be modified as described herein to improve one or more yield traits. Non-limiting examples of canola plant species that may be modified as described herein may include, but are not limited to, *Brassica napus*, *Brassica rapa*, *Brassica juncea* and/or *Brassica rapa subsp. Oleifera* (syn. *B. campestris* L.).

The invention will now be described with reference to the following examples. It should be appreciated that these examples are not intended to limit the scope of the claims to the

invention but are rather intended to be exemplary of certain embodiments. Any variations in the exemplified methods that occur to the skilled artisan are intended to fall within the scope of the invention.

EXAMPLES

Example 1. Modification of *SHATTERPROOF MADS-BOX (SHP)* genes in canola

A strategy to generate edits in the canola *SHP* genes of BnaA07g18050D (SEQ ID NO:100) (SHP2) and BnaC06g16910D (SHP2) (SEQ ID NO:240) was developed to decrease activity of the MADS domain transcription factor encoded by the *SHP* genes. To generate a range of alleles with edits in the C-terminal region of the target *SHP* genes, multiple CRISPR guide nucleic acids comprising spacers (SEQ ID NOs:292-297) (see Table 2) having complementarity to targets within the *SHP* genes were designed and placed into a construct.

Lines carrying edits in the *SHP* genes were screened and those that showed about 10% of the sequencing reads having edits in the targeted gene were advanced to the next generation.

Table 2. Spacers for guide nucleic acids

Spacer (SEQ ID NO:)	Target gene
PWsp236 (292)	BnaA07g18050D (SHP2) (SEQ ID NO:100), BnaC06g16910D (SHP2) (SEQ ID NO:240)
PWsp237 (293)	BnaC06g16910D (SHP2) (SEQ ID NO:240)
PWsp238 (294)	BnaC06g16910D (SHP2) (SEQ ID NO:240), BnaA07g18050D (SHP2) (SEQ ID NO:100)
PWsp239 (295)	BnaC06g16910D (SHP2) (SEQ ID NO:240)
PWsp240 (296)	BnaC06g16910D (SHP2) (SEQ ID NO:240), BnaA07g18050D (SHP2) (SEQ ID NO:100)
PWsp241 (297)	BnaC06g16910D (SHP2) (SEQ ID NO:240), BnaA07g18050D (SHP2) (SEQ ID NO:100)

Example 2.

Edited plants were sequenced with NGS sequencing techniques to further characterize the edit alleles generated. A range of alleles for the target genes were generated as further described in Table 3.

Table 3. Edited alleles

E1 plant	Edit
CE69293	The BnaA07g18050D allele (SEQ ID NO:100) (SHP2) in CE69293 is heterozygous and the edited copy contains two deletions, one of 8 bp and the other of 10 bp, which results in a frame shift mutation leading to the insertion of 6 novel amino acids at position 180 in the protein sequence before reaching a premature stop codon
CE69294	The BnaA07g18050D allele (SEQ ID NO:100) (SHP2) in CE69294 is homozygous for a 10 bp deletion. The deletion results in a frame shift mutation such that 9 novel amino acids are added to the terminal end of the translated protein and an early stop codon is generated. The resulting protein is missing the amino acids encoded by the last exon of the gene and is also missing a portion of the amino acids encoded by the second to last exon.
CE69295	The BnaA07g18050D allele (SEQ ID NO:100) (SHP2) in CE69295 is homozygous for a 45 bp deletion. The deletion removes the intron-exon boundary of the final exon of the gene and generates an altered protein which does not contain the amino acids encoded by the last exon. The BnaC06g16910D (SHP2) (SEQ ID NO:240) allele in CE69295 is homozygous for a 7 bp deletion. The deletion of the 7 bp results in a frame shift mutation such that a portion of the second to last exon is not translated into the protein and none of the last exon is included in the protein. The edit also results in the addition of 10 novel amino acids to the C-terminal end of the protein before reaching a stop codon.
CE69307	The BnaA07g18050D allele (SEQ ID NO:100) (SHP2) of CE69307 is homozygous for a 20 bp deletion. The deletion results in an immediate stop codon which truncates the resulting protein such that a portion of the amino acids encoded by the second to last exon are missing and all of the amino acids from the last exon are missing. The BnaC06g16910D (SHP2) (SEQ ID NO:240) allele in CE69307 is homozygous for a 7 bp deletion which results in a frame shift leading to a stop codon at position 180 in the amino acid sequence. This frame shift results in a portion of the amino acids in the second to last exon, and all of the amino acids encoded by the last exon are missing from the edited allele.

Example 3. Evaluation of lignin content of edited plants

Plants carrying edits as provided above in **Table 3**, above, were evaluated by staining for lignin content in the pod valve margin. The results are provided in **Table 4** for two of the edited plants.

Table 4. Lignin staining

Plant line	Lignin staining
CE69293	+++
CE69294	+
Wild type	+++++

Example 4: Edited alleles

Canola lines were generated as described in **Example 1** and several lines were recovered which contained a range of edited alleles of the SHP genes BnaA07g18050D (**SEQ ID NO:100**) and BnaC06g16910D (**SEQ ID NO:240**). The SHP genes were sequenced by next generation sequencing and the edited alleles identified are further described in **Table 5** and **Table 6**.

Table 5. Edited alleles of BnaA07g18050D

Allele name (SEQ ID NO)	Description	Notes
Allele C SEQ ID NO:321	7 bp deletion (AGATATC) starting at position 6240 of SEQ ID NO:100	Out of frame deletion
Allele D SEQ ID NO:312	11 bp deletion (CAACTTGTCTA SEQ ID NO:339) starting at position 6422 of SEQ ID NO:100	Out-of-frame deletion which removes the last 3 amino acids of the wild type polypeptide sequence (SEQ ID NO:102) and replaces them with new amino acids (SEQ ID NO:313) (encoded by SEQ ID NO:312)
Allele F SEQ ID NO:314	12 bp deletion (AACTTGTCTAAG SEQ ID NO:340) starting at position 6423 of SEQ ID NO:100	Out-of-frame deletion which removes the last 3 amino acids of the wild type polypeptide sequence (SEQ ID NO:102) and replaces them with new amino acids (SEQ ID NO:315) (encoded by SEQ ID NO:314)

Table 6. Edited alleles of BnaC06g16910D

Allele name (SEQ ID NO)	Description	Notes
Allele A SEQ ID NO:318	420 bp deleted starting at position 3349 of SEQ ID NO:240	Deletes 42 amino acids from C terminal end of the protein
Allele B SEQ ID NO:310	10 bp deletion (TTCAACTTGT SEQ ID NO:341) starting at position 3387 of SEQ ID NO:240	Removes the last four amino acids from SEQ ID NO:242 and replaces them with the amino acids PKF (SEQ ID NO:311)
Allele E SEQ ID NO:319	70 bp deletion (SEQ ID NO:320) starting at position 3391 of SEQ ID NO:240	Deletes the last 3 amino acids from SEQ ID NO:242 (SEQ ID NO:313)
Allele G SEQ ID NO:316	8 bp deletion (TACCTGCG) starting at position 3105 of SEQ ID NO:240 ; and 6 bp deletion (CCTCTT) starting at position 3383 of SEQ ID NO:240	Replaces the last 68 amino acids of SEQ ID NO:242 with an S residue (SEQ ID NO:317)

Example 5. Phenotype analysis

Pod shatter was evaluated by harvesting canola pods when fully mature. The harvested pods were fully dried in a seed dryer (e.g., 48-72 hours). Ten canola pods were selected at

random for evaluation and placed into an empty plastic pipette tip box along with 2 – 9 mm steel balls. The box with the pods and steel balls was placed into a Geno/Grinder automated plant tissue homogenizer set to a speed of 600 for 45 seconds. The contents of the box were inspected, and the number of unbroken pods recorded. The foregoing was repeated five more times for each sample evaluated and a statistical analysis of the results is provided below in

Table 7.

Table 7. Pod shatter results

Genotype (number of plants)	Mean for number of unbroken pods (standard deviation)
Wild type control (24)	3.9 (1.2)
Wild type control expressing GUS (24)	5.0 (1.4)
Null segregant control (27)	5.1 (2.0)
Allele A (16)	3.1 (1.5)
Allele B homozygous (36)	6.1 (1.7) *
Allele C (27)	3.0 (1.8)
Allele D homozygous	5.4 (1.6) *
Allele E (27)	
Allele F homozygous	7.1 (1.3) *
Allele G homozygous (12)	
Allele F homozygous	5.7 (1.9) *
Allele B homozygous (28)	

* indicates significance to the $p < 0.05$ when compared to wild type

The canola lines with edited Allele A of BnaC06g16910D and no other SHP gene edit; as well as canola lines with edited Allele C of BnaA07g18050D and no other SHP gene edit were not significantly different from the wild type control in the number of unshattered pods.

The canola lines with the combination of the edited Allele F/Allele G genotype showed an increase in the number of unshattered pods that was significantly different from the wild type control. The canola lines with edited Allele B and no other SHP gene edit also exhibited significantly less shattering when compared to the wild type control. Additionally, the canola lines with the combination of the edited Allele D and edited Allele E, and the canola lines having the combination of the edited Allele F and edited Allele B were significantly improved when compared to the wild type control lines; but were not significantly different from the null segregant control suggesting that there may be a transformation affect that is contributing to the observed phenotype.

A strong statistical evidence of decreased pod shatter in the canola lines was observed with the combination of edited Allele F/Allele G ($p\text{-value} < 0.003$). Additionally, the genotype having the edited SHP Allele B alone also showed a strong decrease in pod shatter when

compared to the wild type control line and showed a moderate decrease in pod shatter when compared to the null segregant control (p-values < 0.07). The genotypes having the edited SHP allele combination of Allele D/Alele E and the genotype having the combination of Allele F/Alele B were both found to be significantly improved for the presence of intact pods (reduced pod shattering) when compared to the wild type control line but not when compared to the null segregant control (p-values > 0.58).

Example 6. Modification of *SHATTERPROOF MADS-BOX (SHP)* genes in canola

A strategy to generate edits in the canola *SHP* gene of BnaA05g02990D (**SEQ ID NO:148**) (*SHP4*, *SHP4A*) was developed to decrease activity of the MADS domain transcription factor encoded by the *SHP* gene. To generate a range of alleles with edits in the C-terminal region of the target *SHP* gene, multiple CRISPR guide nucleic acids comprising spacers (**SEQ ID NOs:342-346**) (see **Table 8**) having complementarity to targets within the SHP gene were designed and placed into a construct.

Table 8. Spacer sequences for guides for canola SHP gene having the gene ID number of BnaA05g02990D (*SHP4A*, **SEQ ID NO:148**)

Spacer name	SEQ ID NO	sequence
PWsp291	342	TATAATTCTACAGATTAATGAAA
PWsp292	343	CGAGTCATCTTCTCATCAGTCGG
PWsp293	344	CTTAAACAAGTTGGAGAGGTGGT
PWsp294	345	ACCATAATTGTAACATGAATAGA
PWsp295	346	AGAGGGACCTAAGTTACATATGT

Lines carrying edits in the *SHP4* gene were screened and those that showed about 10% of the sequencing reads having edits in the targeted gene were advanced to the next generation. Edited plants were sequenced with NGS sequencing techniques to further characterize the edit alleles generated.

Two edited alleles of the BnaA05g02990D gene were identified. Allele H of Bna05g02990D contains a 2 bp deletion (AC) at position 3472 of **SEQ ID NO:148** giving rise to the edited genomic sequence of **SEQ ID NO:322**. The deletion in Allele H does not affect the coding region of BnaA05g02990D (**SEQ ID NO:148**) but alters the 3' UTR of the gene.

Allele I of Bna05g02990D contains a 7 bp deletion (CTATTCA) at position 3401 of **SEQ ID NO:148** giving rise to the edited genomic sequence of **SEQ ID NO:323**. The deletion in Allele I does not affect the coding region of BnaA05g02990D (**SEQ ID NO:148**) but alters the 3' UTR of the gene.

Plants carrying the edited alleles Allele H and/or Allele I were evaluated by staining for lignin content in the pod valve margin. The results are provided in **Table9**.

Table 9.

Genotype	Lignin Staining
Null segregant	+++++
Heterozygous Allele I	+++++
Homozygous Allele I	+++++
Homozygous Allele H	+++

5 A considerable amount of variation was observed in the intensity of lignin staining within each genotype, likely due to the difference in developmental stages of collected pods. Pods from SHP4a Allele I (both Hom and Het) stained just as strongly as the null segregant samples. The pods from SHP4a – Homozygous Allele H did not stain as strongly as the null segregant. This data suggest there's a moderate reduction in lignin deposition associated with at
10 least Allele H.

The foregoing is illustrative of the present invention and is not to be construed as limiting thereof. The invention is defined by the following claims, with equivalents of the claims to be included therein.

THAT WHICH IS CLAIMED IS:

1. A canola plant or plant part thereof comprising at least one mutation in at least one (e.g., one or more) endogenous *SHATTERPROOF MADS-BOX (SHP)* gene encoding a Shatterproof
5 MADS-box transcription factor (SHP) polypeptide.
2. The canola plant or plant part thereof of claim 1, wherein the endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally wherein the endogenous SHP gene has a gene identification
10 number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*).
3. The canola plant or plant part thereof of claim 1 or claim 2, wherein the at least one mutation results in a hypomorphic mutation, a dominant negative mutation, or a dominant
15 negative hypomorphic mutation.
4. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation is a base substitution, a base deletion and/or a base insertion.
- 20 5. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation comprises a base substitution to an A, a T, a G, or a C.
6. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation is a deletion of at least one base pair, optionally one base pair to about 100
25 consecutive base pairs.
7. The canola plant or part thereof of any one of claims 1 to 5, wherein the at least one mutation is an insertion of at least one base pair.
- 30 8. The canola plant or part thereof of any of the preceding claims, wherein the at least one mutation is located in the 3' end of the endogenous *SHP* gene.

9. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation in the endogenous *SHP* gene is located in the second to last exon, in the second to last exon and in the intron adjacent to the 3' end of the second to last exon (e.g., the mutation spans the second to the last exon and the intron adjacent to the 3' end of the second to last exon) and/or in the last exon of the endogenous *SHP* gene.

10. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation in an endogenous *SHP* gene results in a mutated *SHP* gene that produces a truncated protein.

11. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation is an out-of-frame insertion or an out-of-frame deletion.

12. The canola plant or part thereof of claim 11, wherein the out-of-frame insertion or out-of-frame deletion produces/results in a premature stop codon, optionally resulting in a C-terminal truncation of the encoded polypeptide (e.g., a truncated SHP polypeptide).

13. The canola plant or part thereof of any one of the preceding claims, wherein the at least one mutation in an endogenous *SHP* gene results in a C-terminal truncation of the encoded SHP polypeptide.

14. The canola plant or part thereof of claim 13, wherein the C-terminal truncation is deletion of one amino acid residue to about 100 consecutive amino acid residues from the encoded SHP polypeptide, optionally a deletion of about 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or 80 consecutive amino acid residues from the C-terminus of the encoded SHP polypeptide.

15. The canola plant or part thereof of any one of the preceding claims, wherein the canola plant or plant part thereof comprises at least one mutation in two or more (e.g., 2, 3, 4, 5, or 6 or more) endogenous *SHP* genes (e.g., in two or more of a *SHP1* gene, a *SHP2* gene, a *SHP3* gene, and/or a *SHP4* gene).

16. The canola plant or part thereof of any one of the preceding claims, wherein the endogenous *SHP* gene: (a) comprises a nucleotide sequence having at least 80% sequence

identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;
 (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any
 one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c)
 encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID**
 5 **NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at
 least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205,**
237-239 or 289-291.

17. The canola plant or part thereof of any one of the preceding claims, wherein the at least
 10 one mutation results in a deletion or insertion of one or more base pairs located in a region
 having at least 80% sequence identity to any one of the nucleotide sequences of **SEQ ID**
NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338.

18. The canola plant or part thereof of any one of the preceding claims, wherein the canola
 15 plant comprising the at least one mutation exhibits reduced pod shattering and/or reduced
 lignification (reduced lignin content) in the pod valve margin when compared to a control canola
 plant devoid of the at least one mutation.

19. The canola plant or part thereof of any one of the preceding claims, wherein the at least
 20 one mutation results in a mutated *SHP* gene comprising a nucleotide sequence having at least
 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314,**
316, 318, 319, 321, 322, or 323.

20. The canola plant or part thereof of any one of the preceding claims, wherein the at least
 25 one mutation results in a mutated *SHP* gene that encodes a SHP polypeptide having at least 90%
 sequence identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or**
317.

21. The canola plant or part thereof of any one of the preceding claims, wherein the mutation
 30 is a non-natural mutation.

22. A canola plant cell, comprising an editing system, the editing system comprising:
 (a) a CRISPR-Cas effector protein; and

(b) a guide nucleic acid (e.g., gRNA, gDNA, crRNA, crDNA, sgRNA, sgDNA) comprising a spacer sequence with complementarity to an endogenous target gene encoding a Shatterproof MADS-box transcription factor (SHP) polypeptide in the canola plant cell.

5

23. The canola plant cell of claim 22, wherein the SHP polypeptide is a SHP1 polypeptide, a SHP2 polypeptide, a SHP3 polypeptide, and/or a SHP4 polypeptide, optionally a SHP1A polypeptide, a SHP2A polypeptide, a SHP2C polypeptide, a SHP3A polypeptide, a SHP3C polypeptide, or a SHP4A polypeptide.

10

24. The canola plant cell of claim 22 or claim 23, wherein the editing system generates a mutation in the endogenous target gene encoding the SHP polypeptide.

25. The canola plant cell of any one of claims 22-24, wherein the endogenous target gene comprises two or more endogenous *SHP* genes (e.g., two or more of an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally two or more of an endogenous *SHP1A* gene, an endogenous *SHP2A* gene, an endogenous *SHP2C* gene, an endogenous *SHP3A* gene, an endogenous *SHP3C* gene, or an endogenous *SHP4A* gene).

20

26. The canola plant cell of any one of claims 22-25, wherein the endogenous target gene: (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**; (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

25

27. The canola plant cell any one of claims 22-26, wherein the guide nucleic acid comprises any one of the nucleotide sequences of **SEQ ID NOs:292-297 or 342-346**.

30

28. A canola plant regenerated from the plant part of any one of claims 1-21 or the canola plant cell of any one of claims 22-27.

29. The canola plant of claim 28, wherein the canola plant comprises a mutated *SHP* gene comprising a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323** and/or encodes a SHP polypeptide having at least 90% sequence identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317**.

30. The canola plant of claim 28 or claim 29, wherein the canola plant comprises a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

31. The canola plant of any one of claims 28-30, wherein the mutated *SHP* gene comprises a non-natural mutation.

32. A canola plant cell comprising at least one mutation within an endogenous *SHP* gene, wherein the at least one mutation is a base substitution, a base insertion or a base deletion that is introduced using an editing system that comprises a nucleic acid binding domain that binds to a target site in the endogenous *SHP* gene.

33. The canola plant cell of claim 32, wherein the endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally an endogenous *SHP1A* gene, an endogenous *SHP2A* gene, an endogenous *SHP2C* gene, an endogenous *SHP3A* gene, an endogenous *SHP3C* gene, or an endogenous *SHP4A* gene.

34. The canola plant cell of claim 32 or claim 33, wherein endogenous *SHP* gene is two or more endogenous *SHP* genes and the canola plant cell comprises at least one mutation in the two or more endogenous *SHP* genes.

35. The canola plant cell of any one of claims 32-34, wherein the at least one mutation is a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation.

5 36. The canola plant cell of any one of claims 32-35, wherein the target site is located in the second to last exon, in the last exon, and/or in the second to last exon and in the intron adjacent to the 3' end of the second to last exon (e.g., the mutation spans the second to the last exon and the intron adjacent to the 3' end of the second to last exon) of the endogenous *SHP* gene.

10 37. The canola plant cell of any one of claims 32-36, wherein the target site is within a region of the endogenous *SHP* gene, the region having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**.

15 38. The canola plant cell of any one of claims 32-37, wherein the target site is within a region of the endogenous *SHP* gene that encodes an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, optionally within a region of the endogenous *SHP* gene the region encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

20 39. The canola plant cell of any one of claims 32-38, wherein the editing system further comprises a nuclease, the nucleic acid binding domain binds to a target site in a sequence having least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241** and/or in a sequence having at least 80% sequence identity to a sequence
25 of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, and the at least one mutation within a *SHP* gene is made following cleavage by the nuclease.

40. The canola plant cell of claim 39, wherein the nuclease is a zinc finger nuclease, transcription activator-like effector nucleases (TALEN), endonuclease (e.g., FokI) or a
30 CRISPR-Cas effector protein.

41. The canola plant cell of any one of claims 32-40, wherein the nucleic acid binding domain is a zinc finger, transcription activator-like DNA binding domain (TAL), argonaute, or a CRISPR-Cas effector nucleic acid binding domain.

5 42. The canola plant cell of any one of claims 32-41, wherein the at least one mutation within the *SHP* gene is an insertion and/or a deletion, optionally the at least one mutation is an out-of-frame insertion or an out-of-frame deletion.

10 43. The canola plant cell of any one of claims 32-42, wherein the at least one mutation within the *SHP* gene is an insertion and/or a deletion that results in a premature stop codon, optionally results in an SHP polypeptide with a C-terminal truncation.

15 44. The canola plant cell of any one of claims 32-43, wherein the at least one mutation within the *SHP* gene comprises a point mutation.

20 45. The canola plant cell of any one of claims 32-44, wherein the at least one mutation results in a mutated *SHP* gene comprising a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323.**

25 46. The canola plant cell of any one of claims 32-45, wherein the at least one mutation results in a mutated *SHP* gene that encodes a SHP polypeptide having at least 90% sequence identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317.**

30 47. The canola plant cell of any one of claims 32-46, wherein the at least one mutation is a non-natural mutation.

48. A canola plant regenerated from the canola plant cell of any one of claims 32-47 comprising the at least one mutation within the endogenous *SHP* gene.

49. The canola plant of claim 48, wherein the canola plant comprising the at least one mutation has a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin

content) in the pod valve margin as compared to a canola plant devoid of the at least one mutation (e.g., an isogenic plant (e.g., wild type unedited plant or a null segregant).

50. A method of producing/breeding a transgene-free edited canola plant, comprising:

5 crossing the canola plant of any one of the claims 1-21, 28-31, 48, or 49 with a transgene free plant, thereby introducing the at least one mutation into the canola plant that is transgene-free; and

selecting a progeny canola plant that comprises the at least one mutation and is transgene-free, thereby producing a transgene free edited canola plant.

10 51. A method of providing a plurality of canola plants having phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising planting two or more canola plants of any one of claims 1-21, 28-31, 48, or 49 in a growing area, thereby providing the plurality of canola plants having a phenotype of
15 reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a plurality of control canola plants devoid of the at least one mutation.

52. A method of creating a mutation in an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene in a canola plant, comprising:

20 (a) targeting a gene editing system to a portion of the endogenous *SHP* gene that

(i) comprises a sequence having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; and/or

(ii) encodes a sequence having at least 80% identity to any one **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, and

25 (b) selecting a canola plant that comprises a modification located in a region of the endogenous *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**, optionally located in a region of the endogenous *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285**.

30

53. A method of generating variation in a Shatterproof MADS-box transcription factor (SHP) polypeptide in a canola plant cell, comprising:

introducing an editing system into a canola plant cell, wherein the editing system is targeted to a region of *SHATTERPROOF MADS-BOX (SHP)* gene; and

5 contacting the region of the *SHP* gene with the editing system, thereby introducing a mutation into the *SHP* gene and generating variation in the SHP polypeptide in the canola plant cell.

54. The method of claim 53, wherein the *SHP* gene comprises a nucleotide sequence having
10 at least 80% sequence identity to **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, and/or the region of the *SHP* gene in which variation is generated comprises a nucleotide sequence having at least 80% sequence identity to of any one **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**.

15 55. The method of claim 53 or claim 54, wherein the SHP polypeptide comprises an amino acid sequence having at least 80% sequence identity to **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, and/or the region of the SHP polypeptide in which variation is generated comprises an amino acid sequence having at least 80% sequence identity to of any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

20 56. The method of any one of claims 53-55, wherein generating variation in the SHP polypeptide in a canola plant results in the canola plant having reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

25 57. The method of any one of claims 53-56, wherein contacting the region of the endogenous *SHP* gene in the canola plant cell with the editing system produces a canola plant cell comprising in its genome an edited endogenous *SHP* gene, the method further comprising (a) regenerating a canola plant from the canola plant cell; (b) selfing the canola plant to produce progeny canola plants (E1); (c) assaying the progeny canola plants of (b) for reduced pod
30 shattering and/or reduced lignification (reduced lignin content) in the pod valve margin; and (d) selecting the progeny canola plants exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin to produce selected progeny canola plants

exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control canola plant.

58. The method of claim 57, further comprising (e) selfing the selected progeny canola plants of (d) to produce progeny canola plants (E2); (f) assaying the progeny canola plants of (e) for reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin; and (g) selecting the progeny canola plants exhibiting in reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin to produce selected progeny canola plants exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin as compared to a control canola plant, optionally repeating (e) through (g) one or more additional times.

59. A method of detecting a mutant *SHP* gene (a mutation in an endogenous *SHP* gene) in a canola plant is provided, the method comprising detecting in the genome of a canola plant a nucleic acid sequence of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, the nucleic acid sequence having at least one mutation in the 3' region of the gene, optionally having at least one mutation located in the second to last exon, in the last exon, and/or in the second to last exon and in the intron adjacent to the 3' end of the second to last exon (e.g., the mutation spans the second to the last exon and the intron adjacent to the 3' end of the second to last exon) of the endogenous *SHP* gene.

60. The method of claim 59, wherein the mutant *shp* gene that is detected comprises a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323**.

61. The method of claim 59 or claim 60, wherein the mutant *shp* gene encodes a polypeptide having at least 90% identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317**.

62. A method for editing a specific site in the genome of a canola plant cell, the method comprising: cleaving, in a site-specific manner, a target site within an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene in the canola plant cell, the endogenous *SHP* gene:

(a) comprising a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**,

(b) comprising a region having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**,

(c) encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**,

5 (d) encoding a region having at least 80% sequence identity to an amino acid sequence of any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby generating an edit in the endogenous *SHP* gene of the canola plant cell and producing a canola plant cell comprising the edit in the endogenous *SHP* gene.

10 63. The method of claim 62, wherein the endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally an endogenous *SHP1A* gene, an endogenous *SHP2A* gene, an endogenous *SHP2C* gene, an endogenous *SHP3A* gene, an endogenous *SHP3C* gene, or an endogenous *SHP4A* gene.

15 64. The method of claim 62 or claim 63, wherein an edit is generated in two or more endogenous *SHP* genes.

65. The method of any one of claims 62-64, further comprising regenerating a canola plant from the canola plant cell comprising the edit in the endogenous *SHP* gene to produce a canola
20 plant comprising the edit in its endogenous *SHP* gene.

66. The method of any one of claims 62-65, wherein the edit results in a non-natural mutation.

25 67. The method of any one of claims 62-66, wherein the edit is located in the 3' end of the endogenous *SHP* gene, optionally located in the second to last exon, in the last exon, and/or in the second to last exon and in the intron adjacent to the 3' end of the second to last exon (e.g., the mutation spans the second to the last exon and the intron adjacent to the 3' end of the second to last exon) of the endogenous *SHP* gene.

30 68. The method of any one of claims 62-67, wherein the canola plant comprising the edit in its endogenous *SHP* gene exhibits a phenotype of reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

69. The method of any one of claims 62-68, wherein the endogenous *SHP* gene encodes a Shatterproof MADS-box transcription factor (SHP) polypeptide and the edit results in a truncated SHP polypeptide, optionally a C-terminal truncation of the SHP polypeptide,
 5 optionally wherein the C-terminal truncation results in a deletion of 1 amino acid residue to about 100 consecutive amino acid residues from the C-terminus of the SHP polypeptide.

70. The method of any one of claims 62-69, wherein the edit results in a mutated *SHP* gene comprising a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID**
 10 **NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323.**

71. The method of any one of claims 62-70, wherein the edit results in a mutated *SHP* gene that encodes a SHP polypeptide having at least 90% sequence identity to any one of **SEQ ID**
NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317.

72. A method for making a canola plant, comprising:

(a) contacting a population of canola plant cells comprising an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene with a nuclease linked to a nucleic acid binding domain (e.g., editing system) that binds to a sequence (i) having at least 80% sequence identity
 20 to a nucleotide sequence of any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**, (ii) comprising a region having at least 80% identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**; (iii) encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**, and/or (iv) encoding a region having at least 80% sequence identity to any one
 25 of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**;

(b) selecting a canola plant cell from the population of canola plant cells in which an endogenous *SHP* gene has been mutated, thereby producing a canola plant cell comprising a mutation in the endogenous *SHP* gene; and

(c) growing the selected canola plant cell into a canola plant comprising a mutation in the
 30 endogenous *SHP* gene.

73. A method for reducing pod shattering and/or reducing lignification (reduced lignin content) in the pod valve margin in a canola plant, comprising

(a) contacting a canola plant cell comprising an endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene with a nuclease targeting the endogenous *SHP* gene, wherein the nuclease is linked to a nucleic acid binding domain (e.g., editing system) that binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene:

(i) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;

(ii) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;

(iii) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or

(iv) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291** to produce a canola plant cell comprising a mutation in the endogenous *SHP* gene; and

(b) growing the canola plant cell comprising a mutation in the endogenous *SHP* gene into a canola plant comprising the mutation in the endogenous *SHP* gene thereby producing a canola plant having a mutated endogenous *SHP* gene and reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

74. A method for producing a canola plant or part thereof comprising at least one cell having a mutated endogenous *SHATTERPROOF MADS-BOX* (*SHP*) gene, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene

(a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;

(b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;

(c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or

(d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing

the canola plant or part thereof comprising at least one cell having a mutation in the endogenous *SHP* gene.

75. A method for producing a canola plant or part thereof comprising a mutated endogenous *SHATTERPROOF MADS-BOX (SHP)* gene and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin, the method comprising contacting a target site in an endogenous *SHP* gene in the canola plant or plant part with a nuclease comprising a cleavage domain and a nucleic acid binding domain, wherein the nucleic acid binding domain binds to a target site in the endogenous *SHP* gene, wherein the endogenous *SHP* gene:
- (a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;
 - (b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;
 - (c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or
 - (d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, thereby producing the canola plant or part thereof comprising an endogenous *SHP* gene having a mutation and exhibiting reduced pod shattering and/or reduced lignification (reduced lignin content) in the pod valve margin.

76. The method of any one of claims 72-75, wherein the endogenous *SHP* gene is an endogenous *SHP1* gene, an endogenous *SHP2* gene, an endogenous *SHP3* gene, and/or an endogenous *SHP4* gene, optionally an endogenous *SHP1A* gene, an endogenous *SHP2A* gene, an endogenous *SHP2C* gene, an endogenous *SHP3A* gene, an endogenous *SHP3C* gene, or an endogenous *SHP4A* gene.

77. The method of any one of claims 72-76, wherein the nuclease cleaves the endogenous *SHP* gene, thereby introducing the mutation into the endogenous *SHP* gene.

78. The method of any one of claims 72-77 wherein the mutation is a non-natural mutation.

79. The method of any one of claims 72-78, wherein the mutation is a substitution, an insertion and/or a deletion.

80. The method of any one of claims 72-79, wherein the mutation is a hypomorphic mutation, a dominant negative mutation, or a dominant negative hypomorphic mutation.

81. The method of any one of claims 72-80, wherein the mutation is a substitution, an insertion, and/or a deletion.

82. The method of any one of claims 72-81, wherein the mutation comprises a point mutation.

83. The method of any one of claims 72-82, wherein the mutation is an insertion of about one base pair to about 100 base pairs.

84. The method of any one of claims 72-83, wherein the mutation is a deletion of one base pair to about 100 base pairs.

85. The method of any one of claims 72-84, wherein the mutation is an out-of-frame insertion or an out-of-frame deletion that results in a premature stop codon.

86. The method of any one of claims 72-85, wherein the mutation results in a mutated SHP gene that produces a mutated SHP protein, optionally wherein the mutated SHP polypeptide is truncated at its C-terminus.

87. The method of claim 86, wherein the mutated SHP polypeptide comprises a truncation at its C-terminus of one amino acid residue to about 100 consecutive amino acid residues, optionally a C-terminal truncation about 70 consecutive amino acid residues to about 80 consecutive amino acid residues.

88. The method of claim 86 or claim 87, wherein the mutated *SHP* gene comprises a nucleotide sequence having at least 90% sequence identity to any one of **SEQ ID NOs:298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 319, 321, 322, or 323**.

89. The method of any one of claims 86-88, wherein the mutated SHP polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of **SEQ ID NOs:299, 301, 303, 305, 307, 309, 311, 313, 315, or 317.**

5

90. The method of any one of claims 72-89, wherein the nuclease is a zinc finger nuclease, transcription activator-like effector nucleases (TALEN), endonuclease or a CRISPR-Cas effector protein.

10 91. The method of any one of claims 72-90, wherein the nucleic acid binding domain is a zinc finger, transcription activator-like DNA binding domain (TAL), argonaute or a CRISPR-Cas effector DNA binding domain.

92. A plant produced by any one of the methods of claims 72-91.

15

93. A guide nucleic acid that binds to a target site in a *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the target site is in a region of the *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338,** optionally a region of the *SHP* gene having at least 80% sequence identity to any one of **SEQ ID NOs:75-82, 85-92, 107-112, 116-120, 124-127, 129, 135, 136, 139, 140, 156, 157, 159-161, 164-166, 181-184, 187-190, 195, 196, 212-219, 222-224, 229, 230, 246-248, 251-253, 255-257, 261-264, 267, 268, 271, 272, 275, 276, 279, 280, 283, or 285.**

20

94. The guide nucleic acid of claim 93, wherein the guide nucleic acid comprises a spacer comprising a nucleotide sequence of any one of **SEQ ID NOs:292-279.**

25

95. A system comprising the guide nucleic acid of claim 93 or claim 94 and a CRISPR-Cas effector protein that associates with the guide nucleic acid.

30 96. The system of claim 95, further comprising a tracr nucleic acid that associates with the guide nucleic acid and a CRISPR-Cas effector protein, optionally wherein the tracr nucleic acid and the guide nucleic acid are covalently linked.

97. A gene editing system comprising a CRISPR-Cas effector protein in association with a guide nucleic acid, wherein the guide nucleic acid comprises a spacer sequence that binds to an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene.

98. The gene editing system of claim 97, wherein the *SHP* gene:

(a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;

(b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;

(c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or

(d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

99. The gene editing system of claim 97 or claim 98, wherein the guide nucleic acid comprises a spacer sequence comprising a nucleotide sequence of any one of **SEQ ID NOs:292-297 or 342-346**.

100. The gene editing system of any one of claims 97-99, further comprising a tracr nucleic acid that associates with the guide nucleic acid and a CRISPR-Cas effector protein, optionally wherein the tracr nucleic acid and the guide nucleic acid are covalently linked.

101. A complex comprising a guide nucleic acid and a CRISPR-Cas effector protein comprising a cleavage domain, wherein the guide nucleic acid binds to a target site in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the endogenous *SHP* gene:

(a) comprises a nucleotide sequence having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;

(b) comprises a region having at least 80% sequence identity to a nucleotide sequence of any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;

(c) encodes a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or

(d) encodes a region of a SHP polypeptide having at least 80% sequence identity to any one of **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**, and the cleavage domain cleaves a target strand in the *SHP* gene.

5 102. An expression cassette comprising (a) a polynucleotide encoding CRISPR-Cas effector protein comprising a cleavage domain and (b) a guide nucleic acid that binds to a target site in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene, wherein the guide nucleic acid comprises a spacer sequence that is complementary to and binds to

10 (i) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:69, 70, 100, 101, 148, 149, 177, 178, 206, 207, 240 or 241**;

(ii) a portion of a nucleic acid having at least 80% sequence identity to any one of **SEQ ID NOs:72-96, 103-144, 151-173, 180-202, 209-236, 243-288 or 324-338**;

(iii) a portion of a nucleic acid encoding an amino acid sequence having at least 80% sequence identity to any one of **SEQ ID NOs:71, 102, 150, 179, 208, or 242**; and/or

15 (iv) a portion of a nucleic acid encoding an amino acid sequence having at least 80% identity to any one **SEQ ID NOs:97-99, 145-147, 174-176, 203-205, 237-239 or 289-291**.

103. A canola plant or plant part thereof comprising at least one mutation in at least one endogenous *SHATTERPROOF MADS-BOX (SHP)* gene having a gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*), optionally wherein the at least one mutation is a non-natural mutation.

25 104. A guide nucleic acid that binds to a target nucleic acid in a *SHATTERPROOF MADS-BOX (SHP)* gene having a gene identification number (gene ID) of BnaA04g01810D (*SHP3*), BnaA07g18050D (*SHP2*), BnaA05g02990D (*SHP4*), BnaA09g55330D (*SHP1*), BnaC04g23360D (*SHP3*), and/or BnaC06g16910D (*SHP2*).

30 105. A method of producing a canola plant comprising a mutation in an endogenous *SHATTERPROOF MADS-BOX (SHP)* gene and at least one polynucleotide of interest, the method comprising

crossing a first canola plant, which is the canola plant of any one of claims 1-21, 28-31, 48, 49, 92 or 103, with a second canola plant that comprises the at least one polynucleotide of interest to produce progeny canola plants; and

5 selecting progeny canola plants comprising the mutation in the *SHP* gene and the at least one polynucleotide of interest, thereby producing the canola plant comprising a mutation in an endogenous *SHP* gene and at least one polynucleotide of interest.

106. A method of producing a canola plant comprising a mutation in an endogenous *SHP* gene and at least one polynucleotide of interest, the method comprising
10 introducing at least one polynucleotide of interest into a canola plant of any one of claims 1-21, 28-31, 48, 49, 92 or 103, thereby producing a canola plant comprising a mutation in a *SHP* gene and at least one polynucleotide of interest.

107. A method of producing a canola plant comprising a mutation in an endogenous *SHP* gene and exhibiting a phenotype of improved yield traits, improved plant architecture and/or improved defense traits, comprising
15 crossing a first plant, which is the plant of any one of claims 1-21, 28-31, 48, 49, 92 or 103, with a second plant that exhibits a phenotype of improved yield traits, improved plant architecture and/or improved defense traits; and
20 selecting progeny plants comprising the mutation in the *SHP* gene and a phenotype of improved yield traits, improved plant architecture and/or improved defense traits, thereby producing the plant comprising a mutation in an endogenous *SHP* gene and exhibiting a phenotype of improved yield traits, improved plant architecture and/or improved defense traits as compared to a control plant.

25 108. A method of controlling weeds in a container (e.g., pot, or seed tray and the like), a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, comprising
applying an herbicide to one or more (a plurality) canola plants of any one of claims 1-21, 28-31, 48, 49, 92 or 103 growing in a container, a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside, thereby controlling the weeds in the container, the growth chamber, the greenhouse, the field, the recreational area, the lawn, or on the roadside in which the one or more canola plants are growing.

109. A method of reducing insect predation on a canola plant, comprising applying an insecticide to one or more canola plants of any one of claims 1-21, 28-31, 48, 49, 92 or 103, thereby reducing insect predation on the one or more plants.

5 110. A method of reducing fungal disease on a canola plant, comprising applying a fungicide to one or more canola plants of any one of claims 1-21, 28-31, 48, 49, 92 or 103, thereby reducing fungal disease on the one or more canola plants.

10 111. The method of claim 109 or claim 1010, wherein the one or more canola plants are growing in a container, a growth chamber, a greenhouse, a field, a recreational area, a lawn, or on a roadside.

15 112. The method of any one of claims 105-111, wherein the polynucleotide of interest is a polynucleotide that confers herbicide tolerance, insect resistance, disease resistance, increased yield, increased nutrient use efficiency or abiotic stress resistance.